DFID Department for
International
Development

Guidance on
Evaluation
and
Review
for
DFID Staff

Guidance on Evaluation and Review for DFID Staff

Evaluation Department
July 2005

Abbreviations and Acronyms

CDF	Comprehensive Development Framework
CIDA	Canadian International Development Agency
DAC	Development Assistance Committee
Danida	Danish International Development Assistance
DFID	UK Department for International Development
EOI	Expression of interest
EU	European Uniona
Eval	Evaluation
EvD	DFID Evaluation Department
EVSUM	Evaluation Summary
ILO	International Labour Organisation
JICA	Japanese International Co-operation Agency
Logframe	Logical Framework
M&E	Monitoring and Evaluation
MDG	Millennium Development Goals
MTR	Mid Term Review
NGO	Non Governmental Organisation
Norad	Norwegian Agency for Development Co-operation
ODA	Official Development Assistance
OECD	Organisation for Economic Co-operation and Development
OPR	Output to Purpose Review
PARC	Performance Assessment Resource Centre
PCR	Project Completion Report
PRA	Participatory Rural Appraisal
PRISM	Performance Related Information System for Management
PRSP	Poverty Reduction Strategy Paper
SIDA	Swedish International Development Co-operation Agency
ToR	Terms of Reference
UN	United Nations
UNHCR	United Nations High Commissioner for Refugees
UNICEF	United Nations Children's Fund

Contents

Introduction

Using DFID evaluation guidance

This guidance will take you step by step through the practice of designing, managing, reporting on and responding to an evaluation.

Good evaluation practice depends on a solid partnership between those commissioning and managing evaluation studies on behalf of their governments and agencies, and the consultants carrying out the work and producing the reports. This guidance is geared towards DFID staff but also attempt to assist all partners in thinking through their roles in the process.

Whilst Evaluation Department (EvD) is responsible for an annual programme of centrally mandated evaluations agreed with the Management Board, the main aim of this guidance material is to improve the quality of decentralised evaluation.

**Want to learn more about evaluation in international
development in the current context?
This guidance is a good starting point for you…**

Chapters 1–3 provide a useful introduction and the annexes provide more detail and references to further reading.

**New to evaluation?
This guidance is for you…**

If you are about to commission or plan an evaluation, or as an evaluator you want to understand what will be expected of you, then go straight to **Chapter 4**. You will be taken step by step on a quick tour of the evaluation process that will answer the following questions:

- What kind of evaluation or review am I planning?
- How can I use the review or evaluation to improve development effectiveness?
- What do DFID's corporate systems require?
- Whom should I involve?
- How should I plan and carry out this review or evaluation?
- How can I reduce the risk of things going wrong?

**Some experience of evaluation, but facing some tricky Issues?
This guidance is for you…**

If you have some experience of project reviews and evaluations but you need to commission and manage a more complex, large-scale, joint evaluation we offer a step by step approach to such evaluations in the practical guidance in **Chapter 4, section 4.3.**

**Want a simple guide to evaluation terms and international standards?
This guidance is for you…**

If you want a simple guide to current definitions and standards in international development evaluations, go straight to the **Annexes**.

Chapter 1

The changing context of international development and evaluation

1.1   Development effectiveness

Well-planned and managed evaluation exercises can improve the quality of future development activities. Evaluations can also challenge accepted thinking to improve the overall development effectiveness.

Agreement on a common agenda for development in the form of the UN Millennium Development Goals (MDGs) has encouraged the different stakeholders in the international development community to come together to support the main drivers of change in a more coherent way. These drivers of change are global and regional as much as national and involve work with many different stakeholders. Evaluating this level of development activity requires increasingly complex, multi-stakeholder, thematic evaluations and covers areas such as:

- conflict reduction and prevention
- fair international trade
- environment agreements
- gender, human rights and democracy
- a new international financial architecture
- more effective working arrangements amongst the UN family of institutions and other global institutions
- a stronger national and international civil society
- country ownership and leadership of poverty reduction processes.
- attempts to harmonise activities amongst donors

1.2 Donor harmonisation

In order to work together more effectively, donors are trying to harmonise their monitoring and evaluation practices, linking these activities more closely with partner governments and civil society in developing countries. This requires common understanding of terminology and common standards for monitoring and evaluation in the context of international development. The Organisation for Economic Co-operation and Development (OECD) Development Assistance Committee's Network on Development Evaluation has been developing these definitions and standards. They lie at the heart of this guidance and are explained in detail in the annexes. Applied consistently, they will be an important stepping stone on the road to harmonisation.

The MDGs provoke renewed urgency around efforts towards global poverty reduction, for which all development partners will be held accountable. High level meetings of multilateral funding agencies and bilateral donor representatives have produced important agreements on how funders and donors should harmonise their practices to reduce the transaction burdens on developing country governments and to increase the overall effectiveness of aid flows.

1.3 National poverty reduction strategies

The World Bank has encouraged developing country governments to chart their own course towards poverty reduction through developing the process of Comprehensive Development Frameworks (CDFs) which external funders can support[1]. Debt relief is being made

available to heavily indebted poor countries on the basis of locally owned and credible Poverty Reduction Strategy Papers (PRSPs). As a result, country-led poverty reduction plans are being supported more directly by donors, in the form of sector wide approaches or poverty related direct budget support. To reduce transaction costs, and to strengthen local systems of accountability, these plans will be increasingly monitored and evaluated by country governments supported by external funders.

Whilst the gains in terms of reaching the MDGs are attractive, the risks for external donors posed by budget support are also significant. Government budgets lie outside the control of external funders and may go off-track in countries with high levels of poverty, through lack of political will or weak governance. The assessment of risk and risk management therefore feature more strongly in the design of development programmes. Risk assessment and management practices will need to be evaluated closely as an aspect of evaluating sustainability.

The monitoring and evaluation of PRSPs mean international evaluators need to work closely with local counterparts to support the development of local monitoring and evaluation capacity. Initially the monitoring and evaluation of PRSPs may be a joint evaluation process, led by country governments and supported by the external funding agencies. In time country governments may prefer to conduct their own evaluations, examining the effectiveness of external support for their poverty reduction efforts.

1.4   Joint evaluation

<div style="border:2px solid teal; background:#cfe5e5; padding:1em;">

Joint Evaluation

Purpose

The aim is to ensure that the evaluation becomes an efficient learning tool, helps promote good governance, focuses on mutual obligations, enables the partners to be fully accountable, and lowers the transaction costs of separate evaluations.

Challenges

Co-ordination of aid interventions among donors is notoriously difficult since each agency has to give up its own procedures. The same applies to joint evaluations. A key challenge will be to secure ownership for such evaluations by focusing on what is relevant for each partner.

Opportunities

Greater credibility and broader ownership of findings among the donor community than would be the case with single donor evaluations.

</div>

A joint evaluation is only realistic where partners are committed to the same joint strategy and are prepared to minimise their distinct institutional evaluation requirements. The main challenge is how to involve all the different partners appropriately at the key stages from evaluation design to reporting.

In joint evaluations with a large steering group of stakeholders, the involvement of the majority is typically limited to a few key stages in the evaluation process, such as the approval of the evaluation design and the team and the chance to comment on interim and final reports. The management and co-ordination of the evaluation process is typically delegated to one donor agency or a management group. This type of evaluation is described in Chapter 4, section 4.3

---

[1] *The poverty reduction Strategy Initiative – An independent Evaluation of the World Bank's Support Through 2003* (The International Bank for Reconstruction and Development/The World Bank, © 2004).

1.5 Organisational effectiveness

The UK Department for International Development (DFID), in line with other donors, places stronger emphasis on results based management. The DFID Management Board is seeking feedback on DFID's performance through more regular and timely reviews and evaluations. Secretaries of State and the Development Committee Members are keen to know what mixture of aid instruments work best in what circumstances. As a result, DFID senior managers are tightening up corporate performance systems and strengthening DFID's Evaluation resource to increase the amount of evidence available to make these judgements.

Whilst evaluating outcomes and processes remains important, results based management in the context of the MDGs puts a greater emphasis on measuring impact. As the technical understanding of monitoring and evaluation processes grows and spreads, there is more widespread public interest, internationally, in the impact of development interventions and the efficient use of funding.

Evaluation strengthens DFID's accountability by documenting the allocation, use and results of its development assistance and by calling those responsible for policy and implementation to account for performance. It is recognised that DFID is part of a collective effort, which makes attribution of performance difficult. Evaluation can help to clarify where accountability rests and to confirm achievement. DFID's evaluation studies strengthen its accountability not only to the British Parliament, but also to the government and public in partner countries, and inform other development partners.

Increasing global communication and information exchange, and in the UK the Freedom of Information Act, mean greater scrutiny of the effectiveness of development activities in both donor and recipient countries.

Commitments to increased transparency are leading to more evaluation reports being published on the internet. This provides greater scope for meta-evaluation, the evaluation of a range of evaluations, and wider opportunities for learning from evaluation.

In recent years DFID, like many other bilateral donors, has decentralised operations to country level offices encouraging local staff to work with greater degrees of autonomy. New staff with limited training in monitoring and evaluation are therefore assuming responsibility for conducting (increasingly complex) joint evaluation studies. This guidance is intended to support them.

The changing context and trends in monitoring and evaluating in international development

The changing context and trends in monitoring and evaluating in international development

From	To
Interim project reviews involving **independent teams of consultants** with representatives of implementing agencies	Interim project reviews carried out as audited **internal processes**
Mostly **project** level evaluations	**More larger scale** sector-wide, thematic **country level and synthesis evaluations**
Mostly commissioned and managed by the **individual** funding agencies	More **joint evaluation** (jointly funded, jointly commissioned and jointly managed)
Development partners in country formally involved in specifying the terms of reference	Development partners in country more involved in the **specification and management of the evaluation**
Donors primarily interested in the results and recommendations **of the evaluations they fund**	Country-led **poverty reduction** processes become the locus for evaluation effort, with partner countries commissioning and increasingly interested in the results and the recommendations
	Partner countries beginning to evaluate the **contribution of the donors** to their development priorities
Poor dissemination of evaluation reports	Greater **transparency and** dissemination **of evaluation reports through the Internet**
Agencies more focussed on the efficiency of inputs in relation to outputs	**More concern with outcomes impact and sustainability as part of a** greater interest in results based management
Acceptance of evaluation as central to development effectiveness, less concern about proportional costs and uses of evaluation	**More concern about the** cost-effectiveness of evaluation in relation to performance management **as a route to greater development effectiveness**
Cross-cutting themes of gender, environment, human rights and democracy **emerging**	**The role of** cross-cutting themes **of gender, environment, conflict prevention, human rights and democracy is** better understood and highlighted more in evaluation exercises
Accountability primarily 'upwards' towards the source of the financial flows	**Increasingly mutual accountability** to partners involved in the partnership, and 'downwards' to recipients of aid flows and primary beneficiaries
Limited use of local perspectives and other partner insights	Greater recognition **and use of partner expertise and insight**
Limited focus on evaluating the **process of change**	More **interest in evaluating processes as well as results**
Limited attempts to harmonise donor practices **in evaluation terminology and methodology and reporting requirements**	Greater commitment to harmonising **donor practices**

Chapter 2

Basic ideas about evaluation in international development

2.1   What is evaluation?

Evaluation of development assistance describes what has happened and why, using reliable and transparent methods of observation and analysis. All major funding agencies involved in international development assistance undertake formal evaluations of part of their development activities each year. OECD's Development Assistance Committee (DAC), has recommended the definition quoted below for its member countries. This definition has been adopted by all major donor agencies[2]:

> 'An evaluation is an assessment, as systematic and objective as possible, of ongoing or completed aid activities, their design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, developmental efficiency, effectiveness, impact and sustainability'.

Even though evaluation makes use of and informs internal monitoring and performance assessment, it is distinct from these activities in that it involves usually external, impartial views. The term 'evaluation' is normally reserved in DFID for impartial, independent and in-depth evaluations[3] leading to an independent report. In DFID, internal evaluative activities, taking place during the lifetime of the project or programme, may still involve external consultants. These are usually referred to as 'reviews'.

2.2   Evaluation and monitoring

Monitoring and evaluation are often mentioned jointly under the acronym M&E. Although monitoring and evaluation are complementary analytical functions, they have clearly distinguishable characteristics and vary in depth and purpose. Monitoring is a continuous internal process, conducted by managers, to check on the progress of development interventions against pre-defined objectives and plans — 'keeping the ship on course'. In DFID, evaluations normally take place when a project or programme has finished. The evaluation will answer specific questions related to the relevance, effectiveness, efficiency, impact and sustainability of the completed development activity. Impact is assessed in terms of intended and unintended, positive and negative effects, and whether these can be attributed to the project or other forces operating in the same context. Unintended negative effects would also be noted where these are attributable to the project or programme. Evaluations uncover wider lessons that can be applied elsewhere.

---

[2] Source: Organisation for Economic Cooperation and Development's Development Assistance Committee, *'Principles for evaluation of development assistance' (Paris, 1991)*.

[3] In practice, the word evaluation is often used in slightly different ways by different organisations, to cover both internal reviews and external evaluations. An evaluation is considered to be independent when responsibility for initiating it and carrying it out lies outside the activities of line-staff responsible for programme implementation. In some contexts the term evaluation is also applied to internal reviews conducted during the lifetime of the project or programme by agency staff who may not be directly responsible for the work being evaluated.

Reviews lie somewhere between monitoring and full evaluations. They take place at key intervals in the lifetime of a project or programme, check that the overall direction is still the right one and ask whether the development activities are likely to meet the purpose for which they were planned. In DFID, such reviews would also note and score any changes in the original risk assessment.

> **Monitoring and Evaluation – complementary but different**
> - **Inclusive monitoring** practices linked to decision-making are part of a good performance system — they help to improve the quality of the particular development intervention at the time.
> - **Good evaluation** helps everyone to learn and to improve the quality of future development assistance.

2.3   Why are evaluations done?

In the context of international development, evaluations have two main objectives:

- Lesson learning: to help development partners to learn from experience through discovering whether particular development interventions have worked or not, and through understanding why they have been relatively successful or unsuccessful in particular contexts; and

- Accountability: to account for the use of resources to funders, such as the funding agencies themselves, the parliaments and taxpayers who vote their funds and the beneficiaries in whose name the funds are raised and used.

> **Accountability**
> 'In general terms, what an evaluation for accountability seeks to find out is whether the organisations that are responsible for the evaluated intervention have done as good a job as possible under the circumstances. This means trying to find out if and to what extent the intervention has achieved the results that it was intended to achieve or that it could reasonably have been expected to achieve. It is also likely to involve an assessment of the quality of the processes of planning and implementation. When results are difficult to measure — a common situation in development co-operation — an evaluation initiated for purposes of accountability may focus entirely on questions about process'.
> *Source:* 'Looking back, moving forward', Sida Evaluation Manual

Although accountability and lesson learning aim to be two sides of the same coin, there are sometimes tensions between these twin objectives. Particular evaluations may emphasise one aspect more than another. Separating or combining these objectives affects the evaluation design and the composition of the evaluation team. For instance, it is customary to involve representatives of the implementing agencies more fully in evaluation processes where agency lesson learning is considered more important than the accountability aspect. The lessons learning process involves openly documenting and disseminating experience for all development partners, so that this information can be transformed into knowledge over time, and then inform the design of future activities.

In order to improve the collective performance of the international development community, lessons learned may inform future strategic planning and performance management internally, or inter-organisational planning. In the short-term, lessons learned may help development partners to make specific decisions, such as whether to extend a project, or how to amend a particular policy that may be inhibiting the sustainability of the activity.

The over-arching goal for evaluation in international development is to foster a transparent, inquisitive and self-critical organisational culture across the whole international development community, so we can all learn to do better.

> ## Summary of the main uses of the lessons learnt from evaluation
> To inform urgent managerial decisions within the funding agency
> To inform future planning in a particular sector or country context
> - To inform policy in the donor or partner country government
> - To learn lessons from the work evaluated that can be applied elsewhere
> - To learn from the process of conducting the evaluation

Chapter 3

Types of evaluation

3.1　Broad categories of evaluation

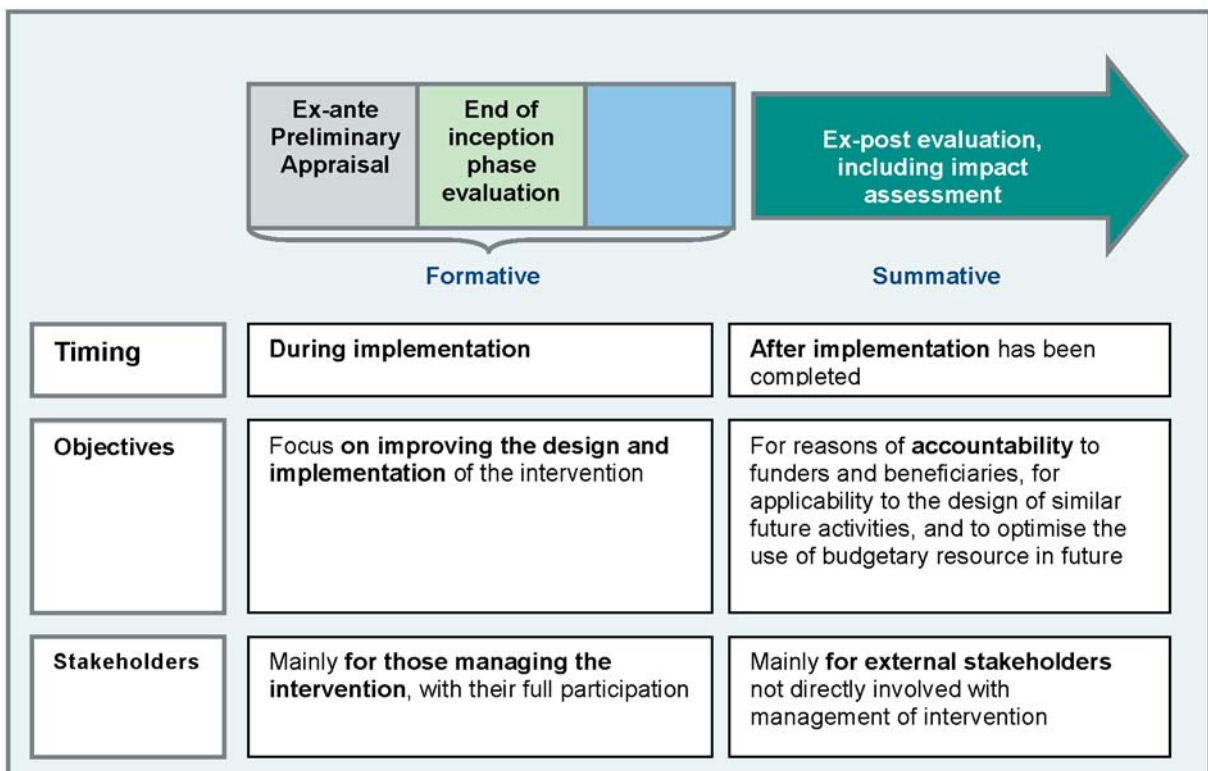Evaluations can be categorised in several different ways:
- by when they take place;
- by the processes used; and
- by where they focus — since evaluation teams do not have the time and resources to look at everything, those commissioning evaluations need to be clear about limiting or focusing the scope of any evaluation exercise, within the resources available.

3.1.1　Type of evaluations by when they take place:

Evaluative activities take place at different times throughout the design and implementation of a development intervention. Broadly they may be formative or summative.

- **Formative evaluation (called 'review' in DFID)** is undertaken during implementation to gain a better understanding of what is being achieved and to identify how the programme or project can be improved.
- **Summative evaluation** is carried out after implementation to assess effectiveness, and to determine results and overall value.

In international development evaluation, the ideas of formative and summative evaluation are reflected in the project design and implementation cycle. The timing, objectives and stakeholders of both types of evaluation are shown in the following diagram.

	Formative	Summative
Timing	**During implementation**	**After implementation** has been completed
Objectives	Focus **on improving the design and implementation** of the intervention	For reasons of **accountability** to funders and beneficiaries, for applicability to the design of similar future activities, and to optimise the use of budgetary resource in future
Stakeholders	Mainly **for those managing the intervention**, with their full participation	Mainly **for external stakeholders** not directly involved with management of intervention

Preliminary, ex-ante, evaluative assessments or appraisals

These are undertaken at the end of the design stage before the project is given permission to proceed. The appraisal may examine the project details, the appropriateness of any co-operation required from partner governments, and would check that the results to be expected and the means to achieve them are realistic.

End of phase evaluations

These are carried out at the completion of a particular phase during multi—phase initiatives. For large process programmes in countries with difficult environments, there may be an inception phase when the implementation team has a chance to engage on the ground and pilot the programme design. At the end of this phase, an external independent evaluation may be commissioned to see if the programme is likely to succeed and to learn lessons from the pilot, before more funds are committed. End of phase evaluations provide recommendations for consideration in the implementation of subsequent phases, or they may recommend the termination of project funding.

Annual Reviews

In DFID all projects and programmes of over £1 million have to be reviewed annually. These are usually internal reviews involving the DFID project manager and representatives in the project, who review progress and make an annual assessment of whether the project is likely to achieve its purpose and whether the risk status of the project has changed. This information is recorded by completing the form on DFID's Performance Related Information System for Management (PRISM).

Interim evaluative reviews

These are normally called Output to Purpose Reviews in DFID. They are usually undertaken at around the mid-point of the implementation of an intervention. They measure and report on performance to date and indicate adjustments that may need to be made to ensure the successful implementation of the project or programme. These adjustments may include adding to, or changing, the outputs in the project's logical framework. Amending the purpose statement in a logical framework would normally require the permission of a senior officer in DFID, since this may affect the contracted responsibilities of the implementers or the agreement with country governments. Information from the review is recorded on PRISM.

Ex-post evaluations and impact assessments

These are undertaken at the end of an intervention. In DFID these evaluations enable project completion reports to be submitted onto PRISM. Ex-post evaluations focus on the DAC criteria of relevance, effectiveness, efficiency, impact and sustainability. Impact assessments report on the development results achieved and focus on the intended and unintended, positive and negative outcomes and impacts.

3.1.2 Type by evaluation process:
Self-evaluation or auto-evaluation

This is similar to monitoring, encouraging operational staff and beneficiaries to assess their own activities. This may be the first part of a wider monitoring and review process, presenting an internally derived assessment as part of the data collection for the full external independent evaluation.

Participatory evaluation

This emphasises the participation of the key stakeholders, especially the intended beneficiaries. This may involve special focus groups of women, children and other vulnerable groups or specialist participatory techniques such as Participatory Rural Appraisal (PRA) to ensure a fair representation of different views.

Process evaluation

This aspect of evaluation may be emphasised where process is a significant aspect of the intervention's design, for instance where processes of institutional or cultural change involve local ownership and are necessary for development to occur and to be sustained. Normally a process approach will involve those affected by the change process in setting their own milestones, and reflecting on progress at agreed intervals. Each occasion for joint reflection provides an opportunity to learn new lessons about how to proceed to the next stage. This is largely an internal process but may involve an external facilitator/evaluator e.g. DFID's Country Assistance Planning annual review process makes use of a review of the agreed indicators or milestones in the Change Impact Monitoring Table.

Synthesis evaluation

This brings together a range of separate evaluation reports, of separate case studies or country examples, into a single report, to tease out overall findings and conclusions, e.g. a synthesis evaluation of health sector programmes in different countries.

Meta-evaluations

These are evaluations of evaluations to bring out overall lessons and examine the assumptions and evaluation methodologies applied.

3.1.3 Type by subject

Evaluations can also be categorised by subject, as shown in the following table. Each category in the table is explained further in the subsequent sections.

➔ Project

Evaluation of individually planned undertakings designed to achieve specific objectives within a given budget and time period.

➔ **Programme**

Evaluation of a coherent set of activities in terms of policies, institutions or finances usually covering a number of related projects or activities in one country.

➔ **Sector**

Evaluation of a single sector or sub-sector such as health or education, or primary education.

➔ **Country programme (all types of development assistance to one country)**

Evaluation of the combined cross-sectoral support provided by a single funding agency to a partner country. This could be done as a joint multi-stakeholder evaluation.

➔ **Country development (all types of development assistance to a country from a partner country perspective)**

Evaluation of the combined cross-sectoral support of all funding agencies to a partner country. Including trade, donor and policy coherence, and often in relation to the country's poverty reduction strategy.

➔ **Aid instruments**

Evaluation of a specific instrument or channel for development aid funding, e.g. through research, through NGOs, through humanitarian assistance, through balance of payment support, through general budget support for poverty reduction, through multilateral agencies, through technical assistance, or through bilateral donor projects or programmes.

➔ **Partnerships and global Funding Mechanisms and institutions**

Include the evaluation of NGO partnership schemes, global funds, global public-private partnerships and global institutions such as the UN agencies.

➔ **Thematic (at global level or across several different agencies work at national level)**

Evaluation of selected aspects of different types of development aid instruments, e.g. influence on trade negotiations, environment, gender, HIV/AIDS or evaluating a range of sector programmes in different countries.

3.2   Project evaluations

Purpose	Challenges	Opportunities
Focuses on a pre-defined cause-effect chain, leading from project inputs to outputs and fulfilment of objectives.	Objectives are often complex, unrealistic, and ill-defined, or may have changed over time. Baseline studies may not have been done.	The more limited scope offered by a single project makes it relatively easy to focus the evaluation, if the objectives against which the project should be evaluated are specifically expressed in quantifiable terms.

Project-level evaluations take place at the end of the project. Evaluations make use of the project's original logical framework, drawing on baseline information internal monitoring data, information from interim reviews and other data. Sustainability is a key area to evaluate, since projects may be deemed successful only as long as external funding is available.

It is essential to make the distinction between what the project is formally expected to achieve and the wider perspective applied in the evaluation. The former perspective is important for accountability, and the latter for extracting experience for the future. The project should not be judged on the basis of what has formally been agreed if the original objectives are later seen as unrealistic. The final assessment should be based on what could realistically have been achieved with the available resources.

3.3   Programme evaluations

Purpose	Challenges	Opportunities
Focuses more on questions of institutional performance, processes, changes and interrelationships, as well as the development impact of the programme area	Involvement and ownership of local partners in a donor designed programme	An evaluation of the overall programme rather than individual project components may save time and costs

In programme evaluations, the involvement of country partners in the evaluation and in the ownership of results and recommendations may be difficult where the partner perceives the programme as the donor's creation rather than their own. In the evaluation, the concern of the donor may be to prove their own organisational effectiveness and to be accountable for aid disbursements, whilst the country partner may be more interested in the contribution of the programme to overall improvements in that sector. Overall programme management is likely to figure as significant in the evaluation.

3.4   Sector evaluations

Purpose	Challenges	Opportunities
These evaluations focus on several donor-supported projects within a given sector in one or several partner countries.	Must involve joint evaluation with all the contributing development stakeholders	Cost-effective way to evaluate clusters of smaller projects together with larger projects. Can reduce transaction costs and have considerable impact on developing policies.

Sector evaluations will normally cover completed as well as ongoing activities. A sector may be defined broadly as the education sector, or, more narrowly as a sub-sector such as primary education. A sector evaluation will cover policies and strategies applied, and a sample of projects representing typical interventions.

3.5   Evaluation of developmental aid[4] instruments

Purpose	Challenges	Opportunities
Focus on the efficiency and effectiveness of specific forms of aid such as balance of payment support, direct budget support, project support, technical assistance, etc. Such evaluations are usually cross-sectoral and cross-country and designed to assess the efficiency of a particular aid instrument as a means of promoting development.	Individual cases to be studied often have very different objectives. This makes it difficult to explore the effects and impact of development assistance systematically.	Drawing experience from a larger sample of aid interventions can reveal lessons that could improve development effectiveness by changing policies and practice.

Evaluations of aid instruments will usually focus on the efficiency, effectiveness and impact of the particular aid instrument and the way it has been managed between the partners.

Different aid instruments such as humanitarian assistance, funding for research or budget support, require different evaluation methods and approaches. The evaluation will need to consider the performance of the various parties involved in this particular development intervention and whether best use was made of the human and financial resources available.

---

[4] In international development terminology, the term 'aid' has increasingly been replaced by 'development assistance' or 'development interventions'. In these guidelines, the term 'aid' is used when describing an aspect of financial disbursement, for example aid instrument, aid disbursement, or bilateral aid. In addition, the term official development assistance (ODA) is sometimes used to specify financial assistance, whereas the term development assistance is reserved for the practical products and activities that ODA funds.

3.6   Country development evaluations

Purpose	Challenges	Opportunities
Focuses on the entire developmental support from all donors and funding agencies to the country evaluated <u>from the partner country perspective</u>. It aims to document past activities and performance in order to provide the partner government with a basis for deciding on changes in their co-operation strategy.	These are complex evaluations requiring an analysis of all aid flows and forms of development cooperation into a country.	Important for policy and planning at the highest level, to provide the basis for a country to conduct their bilateral and multilateral negotiations. They also represent an opportunity to focus on specific issues such as the need for donors to be flexible and to harmonise their efforts.

This is a relatively new form of evaluation, allowing the partner country to be in the driving seat, reflecting on the value added by the combined efforts of the international development community. A major challenge is the collation of all the information on aid flows into the country given the different ways in which donors account for their money and the different start and end dates of financial years in different countries. Donors' financial information is not always transparent or accessible to partner countries. Besides calculating the volume of overall aid disbursement, this type of evaluation will also look at the value of different forms of technical assistance and access to global knowledge networks that development agencies bring.

3.7   Country programme evaluations

Purpose	Challenge	Opportunities
Focuses on the entire donor developmental support to a partner country and aims to document past activities and performance in order to provide the donor and the partner government with a basis for deciding on changes in the co-operation strategy.	These are complex evaluations involving both donor and country partner national policies and objectives and cover all sectors and different forms of aid.	Important for policy and planning at the highest level to provide a basis for country strategies and bilateral negotiations. They also represent an opportunity to focus on specific relevant issues such as the country's dependency on aid, institutional capabilities, democracy, human rights, gender, and environment in a wide context.

The primary aim of country level evaluations is to reflect on experience to improve the delivery of future assistance to the country.

These evaluations cover all sectors and modes of co-operation with the country. Preparing Terms of Reference (ToR) acceptable to all stakeholders will take time. Focusing the evaluation and making it relevant for all users and stakeholders requires close consultation between the stakeholders, the evaluation manager and the evaluation team. The evaluation team should be multi-disciplinary, with expertise reflecting the issues to be given particular attention by the evaluation.

The country programme evaluation looks at the relevance of the individual donor's assistance in relation to other funding agencies, and in relation to the policy and MDG strategies of the particular donor agency and the partner country government. It reviews the instruments used in the bilateral co-operation, the way aid money is used, and the relative weight given to assistance for economic and social development. Data collection will involve an overview of the donor's programme and project work, usually in the form of a document

review and interviews in country with government officials, other development partners, and donor project and programme staff.

In order to make country programme evaluations manageable, the trend is to focus on the following issues: policy and strategy coherence, donor co-ordination, development effectiveness and organisational effectiveness. The evaluation could involve all other forms of support reaching the country in question through contributions to multilateral agencies, such as the donors' contribution to EU and World Bank funds or the effects of untied aid or trade policy.

3.8   Evaluation of partnerships and global institutions

Purpose	Challenges	Opportunities
Reviews the organisational structure of partnerships and global alliances and their effectiveness at country and regional level.	As there tends to be a wide range of stakeholders, agreement on the focus of evaluation may prove difficult.	The credibility and utility of findings may be increased due to the wide range of stakeholders who are closely involved.

Evaluations of partnerships and global institutions usually involve a number of joint funders and a wide range of different users of evaluation results. Each may have a different interest and may want to prioritise the focus of the evaluation differently. Since the funders of these global organisations usually pay for evaluations, the focus of the evaluation may be biased towards funders' priorities.

3.9   Thematic evaluation at global and national level

Purpose	Challenges	Opportunities
Extracts and aggregates information on a specific theme. It may involve different forms of aid, sectors or countries, and make use of information from different types of evaluations.	Methodological constraints in comparing different activities in different contexts, and different forms of assistance in the context of different national activities and policies.	Can be highly influential in shaping development policy.

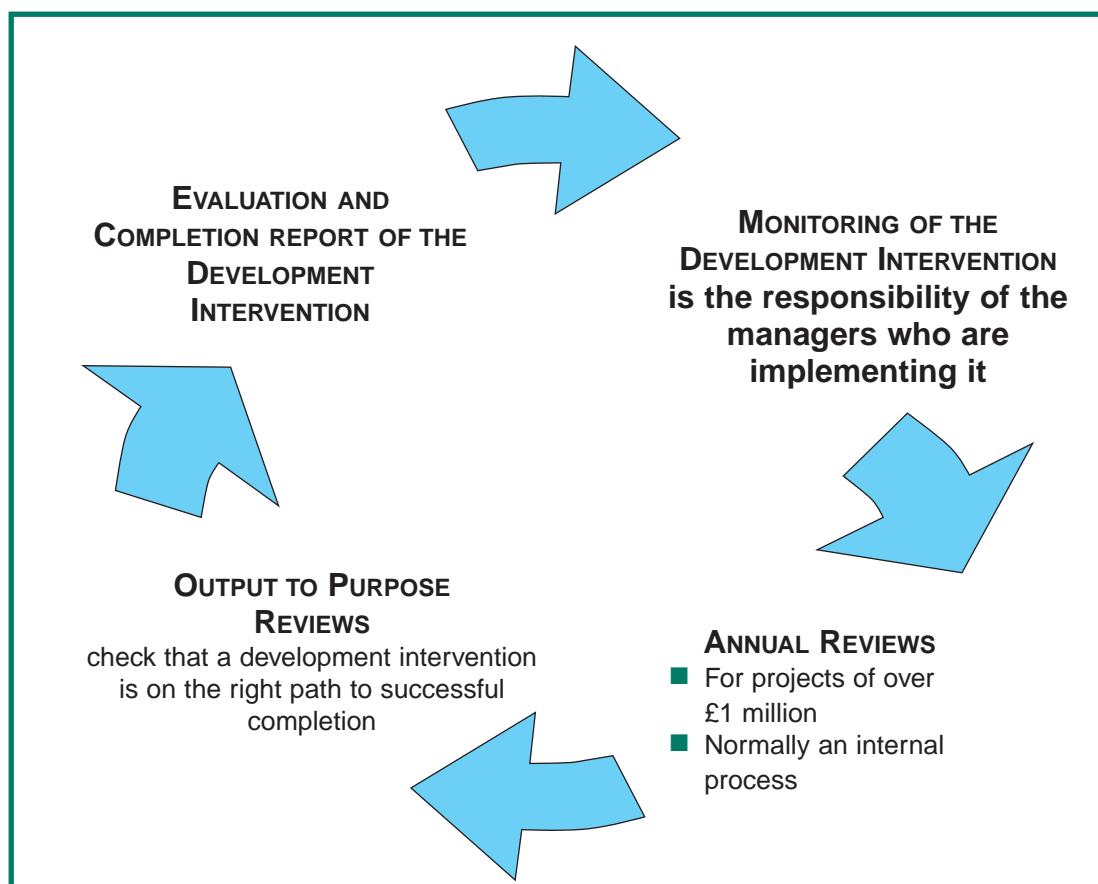T   s-cutting issues' such as gender, institutional development, human rights, democracy, and environment, and cover a number of countries or selected case studies.

In a thematic evaluation the subject may be analysed not only in the context of the donor's development policies, but also in terms of international conventions, and development partners' priorities and strategies. Clearly there will be methodological constraints in distinguishing the impact of any individual contributor in a multi-actor and inter-active context.

Information gathering is usually based on relatively large samples of development activities implemented over a relatively long period of time. Much of the study will therefore have to be based on older documents on file. Older work or work in different countries may reflect different development policy contexts in donor and recipient countries, so it is difficult to compare like with like across time and national context. Fieldwork will essentially concentrate on the more recent development activities and be aimed at verifying information and acquiring a deeper understanding of current issues.

Chapter 4

Conducting Reviews and Evaluations in DFID

EVALUATION AND COMPLETION REPORT OF THE DEVELOPMENT INTERVENTION

MONITORING OF THE DEVELOPMENT INTERVENTION **is the responsibility of the managers who are implementing it**

OUTPUT TO PURPOSE REVIEWS
check that a development intervention is on the right path to successful completion

ANNUAL REVIEWS
- For projects of over £1 million
- Normally an internal process

This guidance concentrates on evaluation processes, but since these are intimately connected with monitoring and review, a brief overview of monitoring and review and evaluation processes, related to development intervention cycle in DFID, is offered here.

Monitoring of the development intervention

Before development interventions are approved and funded in DFID, the Manager of the development intervention will need to complete a memorandum that explains the details of the intervention, its logical framework and how it will be monitored and reviewed.

Monitoring is primarily the responsibility of the managers who are implementing the development intervention. They will keep records of any baseline information collected at the start, together with data on progress as the intervention proceeds.

At agreed intervals (three monthly or six monthly) the Manager of the development intervention will discuss progress with the project team at a monitoring meeting. Records of these meetings will be filed.

Annual Reviews

For development interventions of over £1 million in the value of DFID committed funds, DFID requires an annual review to check on progress and to score the intervention in two ways:

- To check that it is on course for successful completion
- To check for any change in the risk score given at the start of the development intervention

This annual review would normally be an internal process involving the DFID Manager of the development intervention and the implementing team.

The data is recorded on PRISM. This System can collate all project information across DFID so that senior managers can see how well DFID is doing overall and whether the balance of risk is acceptable.

Output to Purpose Reviews (OPR)

An OPR's main purpose is to check that a development intervention is on the right path to successful completion. If this is not the case, the review team should help the implementation team of the development intervention to make the necessary adjustments to get the intervention back on course. For development interventions of over £1 million, DFID expects the Manager of the intervention to organise an OPR. This is sometimes referred to as the Mid Term Review (MTR). Development interventions of under £1 million in value of DFID funds committed are not obliged to have an OPR, but the Manager of the development intervention may choose to have one. A review of this kind is a useful way to help everyone involved to step back to reflect on progress, and to see what wider lessons are being learned.

The OPR team will use monitoring data and their own enquiries and observations to answer five fundamental questions:

- Are the planned activities likely to achieve the outputs?
  - ➢ and if not, should they be done differently, or are different activities required?
- If all the outputs are achieved, will they achieve the purpose of the development intervention?
  - ➢ and if not, are extra outputs or altered outputs required?
- Are the assumptions in the logical framework correct?
  - ➢ and if not, does the logical framework require more fundamental revision?
- Are the risks being managed successfully?
  - ➢ and if not, should extra outputs be designed to manage the risks?
- What lessons are being learnt for wider interest?

Unlike a fully independent and impartial external evaluation, Output to Purpose Reviews occupy the middle ground[5]. They usually involve external evaluators working alongside the DFID Manager of the development intervention to assess overall progress against the intervention's logical framework. As far as possible, OPRs should be participatory exercises, involving the implementation team, beneficiary views and other stakeholders. A participatory approach is more likely to encourage the ownership necessary for sustainability, and encourage a more accurate view of project progress. This approach is also more likely to ensure that any emerging lessons are learnt where they are needed most immediately.

The OPR will result in a report and the completion of a special form on PRISM. This data will be particularly useful for the future evaluation team to see. Fuller details of how to conduct OPRs are available on Insight. However, much of the information in this evaluation guidance will also be of interest to DFID staff who are planning Output to Purpose Reviews.

---

[5] Other bilateral donors refer to similar reviews as 'evaluations' involving different degrees of 'internality' and 'externality', 'independence' and 'impartiality'

Evaluation and Project Completion Reports of the development intervention

In DFID, the term 'evaluation' is usually reserved for independent, impartial, external evaluations conducted on completion of the project or programme. This guidance looks at this evaluation process in detail. Again the evaluation may feed into the preparation of a Project Completion Report (PCR).

PCRs are reviews to assess the extent to which planned outputs have been achieved and to capture lessons learned for the future. A PCR must be completed for eligible projects/programmes of £1 million and over:

- when actual expenditure reaches the 95% threshold; or
- during the last three months of a project; or
- when the stage of the project first changes from operational to completed on PRISM.

Further guidance on PCRs is available from the PRISM link on inSight and from the Blue Book.

4.1   Stages of the evaluation process

We show in sections 4.2 and 4.3 two different kinds of evaluations: a project evaluation and a more complex joint evaluation. In each case the evaluation process involves three stages as follows:

STAGES IN A PROJECT EVALUATION PROCESS		
Stage I	Planning the evaluation	By the DFID Project Manager
Stage II	Carrying out the evaluation	By the evaluation team
Stage III	Following-up the evaluation	By the commissioning partners

STAGES IN A PROJECT EVALUATION

1. PLANNING THE EVALUATION

FOR DFID PROJECT MANAGER

1. Financial resources and authority to commission the evaluation

2. Thinking through the focus of the evaluation

3. Preparing the terms of reference and finding resources

4. Finding a suitable evaluation team and checking their suitability with the stakeholders

5. Organising the logistics

6. Briefing the evaluation team

2. CARRYING OUT THE EVALUATION

FOR THE EVALUATION TEAM

7. The evaluation teams take over and gets on with the job, but you stay in touch

8. Receiving the draft report

3. USING THE EVALUATION

FOR COMMISSIONING PARTNERS

9. Getting the management response

10. Publishing the report

11. Completing the information on PRISM

4.2   Step by step through a sample evaluation process

The guidance here assumes you are relatively new to DFID and that you have been asked to commission and manage a project evaluation for the first time. It takes you step by step

through the three-stages of the process. As mentioned above, this guidance would be similar if you are planning a project review.

Stage I: Planning and commissioning the evaluation

Step one … Do you have the money and the authority to commission the evaluation?

Evaluation exercises involve employing a small team of evaluators and paying for their travel and subsistence. Even quite modest evaluations can be expensive so try to identify the resources available before you start and make a rough estimate of the number of days and the amount of travel needed. The key question here is whether the estimated cost of the evaluation will fall within European procurement limits of about £100,000. If not, then the evaluation will have to go out to competitive tender, following DFID's European procurement procedures. You will need to seek advice from DFID's Procurement Group about the time you should allow for the tendering process.

Step two … Thinking through the focus of the evaluation

As the DFID Evaluation Manager, you would discuss the details of the evaluation with the project implementing team. The questions you would need to resolve together are as follows:

- What is he purpose of the planned review. Is it to be a full evaluation, perhaps, to inform the design of a specific new project? Is it lesson learning? Or is it simply to finalise the Project Completion Report?
- What are the three most important questions you would want the evaluation to answer?
- Who should be involved in the evaluation? - other DFID staff, partner country representatives, other donor representatives?
- The timing of this evaluation? - should it look at immediate impact? Should it be scheduled some time after the end of the project to look at whether the impact has been sustained? When will key people be around to be interviewed or to take part in the evaluation?
- The scope and scale of the evaluation? - a desk study of the data produced by the project only? A desk study and interviews with key stakeholders? Visits to the field consulting a sample of stakeholders? or a full participatory evaluation involving meetings and interviews with primary beneficiaries at a number of different sites?
- What lessons would it be useful to focus on? - technical lessons, lessons about project management, lessons about the country context?
- What resources are available? - this may determine the scope of the evaluation
- Who should be on the evaluation team? - special technical background required, a health or education specialist, a management specialist, someone with a governance, economist or social development background, someone who knows the country context, someone who knows about DFID's policies and procedures, a gender specialist?
- Who should be the team leader? - what kind of qualities would the team leader need to have in this situation, do they need to be a specialist or a generalist to command the respect of the team, should they be an experienced evaluator, do they need to have an understanding of DFID, or other donor or country policy or context knowledge?
- Who are the target audiences? How can the different audiences be best reached? What are the main messages? Who needs to know what? Guidance on designing a dissemination strategy is given in Chapter 4.3, Stage III.

This checklist of questions will help you to think through what the evaluation is about and how it needs to be conducted. Working with the project implementation team, and other key stakeholders who will be interested in the results of the evaluation, will help you to write the first draft of the terms of reference.

Step three … Preparing the terms of reference and identifying resources

The terms of reference give a brief description of the following:
- The project and what it has set out to achieve
- The reasons for the evaluation and the key questions that the evaluators should answer

- Who the main client and interested stakeholders are
- The context in which the evaluation is taking place, e.g. 'decisions will be made about a new project on the basis of this report'
- A rough guide to the tasks to be completed and the schedule
- e.g. 'desk study of 5 days, interviews with the following people … visits to project sites during a 5 day visit to the country and 5 days to write the report equals 15 days in total, during May and June'
- What the outputs from the evaluation may be and when they are due, with specific dates if the report is required for a particular meeting
- e.g. 'a workshop presentation and brief report to the representatives of the Ministry of Education in Malawi, before leaving the country, a draft report by May 10th 2005 and a final report by June 6th 2005 with a completed project completion report, to inform a meeting on July 10th 2005 designing the next phase of support'
- Details of the evaluation team
- e.g. 'a management specialist and experienced evaluator with a knowledge of DFID to act as the team leader; an education specialist with experience in Malawi, an economist and a local consultant who is a gender specialist'
- Any special information about the size and format of the report
- e.g. 20 pages, clearly written in English, using Arial 12 point, with an executive summary of no more than two pages and additional information in the annexes. (A report template and a guide to the style in which Evaluation Department's reports are produced are available from Evaluation Departments insight pages.)
- The name and contact details of the Evaluation Manager and reporting requirements during the evaluation and whom the report should be sent to.

More detail on preparing terms of reference can be found in Chapter 5.

Once you have prepared a draft terms of reference, you need to share them with all the key stakeholders whose interest and commitment is needed to support the evaluation team or to follow up the recommendations from the report. If stakeholders are consulted at this stage, they are likely to be more interested in the results of the evaluation later, and prepared to provide relevant information to the evaluation team. Once the terms of reference are agreed you will need to obtain agreement to the release of resources for the evaluation.

Step four … finding a suitable evaluation team and checking their suitability with the stakeholders

Once the terms or reference and the budget are agreed, you will need to find suitable consultants for the evaluation team. As mentioned in step one, if the costs are above the

European tendering limit then you will have to go out to open tender and follow the guidance of Procurement Group, inviting companies to tender for the work.

If the costs of the evaluation fall within the tendering limit, then you may be able to find suitable consultants by asking colleagues, or approaching one of DFID's Resource Centres, to assist you through their Enabling Contract with DFID. The Health Systems Resource Centre, the Resource Centre for Social Dimensions of Business practices or the Economists Resource Centre are other specialised DFID resource centres that can help you with more specific thematic requests. For a full list of DFID resource centres, please consult the PARC website www.parcinfo.org.

It is most helpful when seeking consultants to make sure the terms of reference state the technical and language skills you require, and whether you need consultants to have specific country experience and evaluation experience. It is important in terms of capacity development to encourage teams to involve local consultants, and to ensure a gender balance across the team. It may also be crucial to have the perspectives of local and national representatives and women evaluators represented within the evaluation team.

You will need to check the CVs of the team with the key stakeholders so that they are happy with your selection. Depending on whether you are organising the team visits or whether this task falls within the terms of the evaluation team's contract, the next stage is to plan the logistics of the evaluation.

Step five … Organising the logistics

On larger evaluations the evaluation team will be responsible for organising the actual evaluation logistics, and making the travel arrangements. Please see the next section for **planning more complex evaluations**. You will have checked their plans as part of the tendering and contracting process. For smaller evaluations you may be more involved in these arrangements. You will certainly be involved in getting formal permissions where visas require letters of invitation or where a letter from DFID is required to ensure a meeting with a senior official. You should try to assemble the key documents at the start.

Step six … Briefing the evaluation team

You need to plan carefully for this initial briefing session. Make sure everyone is properly introduced to key people. At this session you should provide the background information to the project and access to all the key documents. Check the evaluation plan and agree how you and the evaluation team will communicate with each other during the course of the evaluation. As the client and funder, the evaluation team will be keen to understand what you expect to get out of the evaluation. You may wish for other key stakeholders to be present at this briefing too, so that they can explain their expectations of the evaluation.

Depending on how the evaluation is organised, you may have arranged for the initial briefing to include a presentation, and question and answer session with the project implementation team. The more comprehensive the briefing at this point, the more this will help the evaluators to get hold of the big picture before they start to grapple with the data collection.

Good evaluation processes result from clear and open communication between the evaluation manager and the evaluation team. At this first briefing session you have an opportunity to set the tone for this. Since the briefing session may be the first occasion when the members of the evaluation team come together, it is important that your session helps them to gel together as a team.

Stage II: Carrying out the evaluation

Step seven … the evaluation team takes over and gets on with the job, but you stay in touch

If you have prepared good terms of reference, selected a good team and briefed them well, you may be able to step back at this point and await the evaluation team's report. You should be in communication in case they need your help to find a document or facilitate a

meeting. Most importantly you should stay in touch with the team leader informally to make sure the evaluation stays on course.

During this stage, the evaluation team may prepare interim reports at the end of particular phases of the evaluation exercise. These may include presentations to stakeholder workshops to present the findings and receive feedback, or debriefing reports (aide-memoire) to country representatives on completion of individual country visits. Interim reporting is particularly valuable in preparing stakeholders for the presentation of any unwelcome findings and for encouraging interest in the final reports.

You may need to attend one of these sessions to make sure the evaluators are still addressing the main questions. They will have a tendency to be diverted into other interesting areas as they engage with the detail. You may have to pull them back on track.

If sensitive issues are emerging you need to know about these as soon as possible so that you can deal with the implications. Evaluators may be particularly reluctant to reveal any interim findings because they will not have reached their final conclusions. You will need to be sensitive about the fact that this is an independent evaluation and you should not be seen as trying to influence the evaluators' conclusions and recommendations.

The final reason for remaining in communication is in case the members of the evaluation team are not working in harmony. Evaluation assignments can be very stressful since the people whose work is being evaluated are usually apprehensive and often emotional about feeling judged. Team members are often struggling with huge amounts of information and heavy travel schedules. Working intensely together under stress can lead to arguments. In any case the team members may come from different disciplines and see different aspects of the Project in a very different light. It is up to the team leader to manage and resolve these differences, but if the team is not working well, the evaluation manager may have to assist with mediating the situation.

You may also have to guide the evaluators to make sure their methodology is sound and their data is accurate. Your help in these areas will strengthen the credibility of the conclusions and recommendations of the evaluation. However it is the evaluators who have the professional and ethical responsibility for the quality of the evaluation. You can remind the evaluators that these responsibilities are part of their contract (see Annex 4 on evaluator standards).

Step eight … receiving the draft report

You will have made the deadline for the receipt of the draft report clear in your terms of reference and you will have tried to resist all requests from the evaluation team to extend the deadline! The team may use a workshop to present the draft report in the form of a PowerPoint presentation in order to get feedback from the key stakeholders including you as the client. Clarifying ideas in PowerPoint helps the evaluators to present their work effectively and can help them in structuring their report. It is normal practice for the deadline

for the receipt of the draft report to be scheduled to give you time to circulate the report to all the people who should be attending the feedback workshop, with time to read it first. After getting feedback from you and others the evaluation team will prepare a final report and submit to you. Once you are satisfied with its quality and that it conforms to the terms of reference you can authorise payment. Responsibility now reverts from the evaluation team back to you for the third stage.

Stage III: Using the evaluation

Step nine … getting the management response

As the evaluation manager you will need to make arrangements for following-up the recommendations of the report and for the wider publication of the report and the management response to it. Sometimes a summary of management responses and other follow-up action is published with the report. In EvD these summaries are published as EVSUM's. Examples of them are available from EvD's insight pages.

Step ten … publishing the report

Reports are increasingly published on the internet. The moment of publishing may be accompanied by a special announcement on the DFID intranet. The Freedom of Information Act means that all such reports will be available to the public. Evaluation managers may actively promote published reports in the Press or encourage wider links through notices sent to relevant bulletin boards, such as the Development Gateway. A special seminar may be organised to target the dissemination of lessons more effectively to specific audiences rather than by simply publishing the report. Dissemination can be encouraged by using a variety of media such as CD ROMs where potential audiences may not have access to the internet.

Step eleven … completing the information on PRISM

Forms for summarising the evaluation lessons and recommendations should be made available to the evaluators, so that information can be completed and entered onto PRISM.

4.3 The evaluation process for more complex multi-stakeholder joint evaluation

Stage I: Planning and commissioning the evaluation

For smaller evaluations the funding agency that has designed and contracted the delivery of the project usually commissions and manages its own evaluation processes through an external consultancy team. A complex multi-stakeholder evaluation is likely to start with one interested party writing the concept note or approach paper, outlining the main reasons for conducting an evaluation. This concept note will be the trigger to bring all the stakeholders together to agree to conduct the evaluation together.

For larger scale development programmes, such as health or education sector reform programmes, country governments are key stakeholders who want to know the results of evaluation but may not have the resources to commission sector evaluations without donor assistance. In such cases groups of donors contribute to a joint evaluation with their host country partners. They would normally come together to form a Steering Group. The Steering Group will meet together at specific times to agree the terms of reference, approve the evaluation team, ensure oversight of the evaluation, introduce balance in the final evaluation judgements, and to take responsibility for the dissemination and use of results.

To save the time of its members, the Steering Group will normally depute a Management Group with perhaps one agency acting as the Evaluation Manager to handle the task of recruiting and managing the evaluation team.

The Evaluation Manager will prepare terms of reference for consultation and agreement by the Steering Group. Although it may take time to resolve different stakeholder requirements, it is important that consultation at this stage results in a tight focus and coherent set of questions for the evaluation. The overall scope of the evaluation needs to be feasible and realistic.

Once the terms of reference are agreed, the Evaluation Manager initiates the recruitment of the evaluation team. Where international tendering rules apply, the recruitment and selection process may take some time. Consideration of this should be built into the overall scheduling of the evaluation deadlines. The evaluation manager puts out a request for expressions of interest in doing the work. Steering Group partners may be involved in agreeing the shortlist of potential evaluators who will then be asked to submit a full tender.

The lead agency will normally evaluate the bids submitted, although all the partners should be consulted to ensure that the team as a whole is credible, impartial, without vested interests, and not engaged in the implementation process.

Once the contract is agreed, the Evaluation Manager will be responsible for providing the initial briefing to the consultancy team. The briefing will indicate the logistical support, documents and access to data to be provided by the Evaluation Manager to support the evaluation team, and clarify arrangements for communicating with the team.

Stage II: Carrying out the evaluation

The evaluation team, under the direction of the team leader, will conduct an initial information gathering exercise to complete the inception report (with a full description of the methodology and workplan to be adopted). The inception period provides the evaluators with an opportunity to appreciate the difficulties that they may have in completing the evaluation and to make a realistic workplan. At this stage a desk study of the key documents may also be completed. The workplan, and any logistical or conceptual difficulties in meeting the contract as laid out in the terms of reference, will be included in the inception report for presentation to the Steering Group. The meeting to discuss the inception report should clarify any misunderstandings with the contract now that the evaluation team understands the context more clearly. This meeting is an opportunity for all sides to agree any necessary amendments to the contract. The meeting should also reaffirm the scope of the evaluation and any sampling, such as the selection of site visits, case studies or countries to be visited, and resolve logistical difficulties such as responsibilities for security clearance, access to important interviewees, and the organisation of focus groups. At this stage the Steering Group may wish to check the methodology in terms of coverage, scheduling, and whether the data sources will be triangulated sufficiently to produce reliable evidence. Triangulation means verifying the same information from a number of different perspectives to ensure accuracy e.g. matching qualitative information from interviewees with quantitative data and site visits to check perceptions.

Once the workplan is accepted, the main information collection and analysis phase begins. This may involve a more extensive document review, collection of statistical data, individual interviews and telephone or email questionnaires, focus group meetings, and site visits.

During this phase the evaluation team may prepare interim reports at the end of particular phases of the evaluation exercise. These may include written reports to the Management Group, interim workshops to present the findings and receive feedback from specific case studies, and debriefing reports (aide-memoire) to country representatives on completion of individual country visits.

Evaluation team meetings should also be held during the evaluation to help the team leader to plan the contributions to the final report and to focus the team's efforts as the evaluators become more immersed in potentially confusing detail. Interim evaluation team meetings enable the team members to share and review interim findings, re-orientating themselves as necessary to answer the central questions of the evaluation. These team meetings may also reveal real difficulties, such as problems of gaining access to data, which need to be resolved at the earliest stage between the team leader and the Evaluation Manager.

As the information and data collection period comes to an end, the evaluation team members meet to collate findings and complete the analysis. It is helpful if the team actually meet together in person rather than by email to agree the conclusions, lessons and recommendations. At this point, the team leader may wish to lay out the structure of the draft report and then clarify how individual team members will contribute their part to the overall product.

The team leader will plan the feedback arrangements on the draft report with the Management Group or the Evaluation Manager. The Evaluation Manager may take the prime responsibility for calling all the stakeholders together and organising any feedback workshop. The team leader is responsible for collating the draft report and editing it for

clarity of structure and consistency of style. This draft report is circulated to the Steering Group and other stakeholders for their consideration, before any workshop is held to present and discuss it. The team leader may prepare a PowerPoint presentation for the feedback workshop to explain the methodology, findings, conclusions, lessons and recommendations and will facilitate the session in such a way that differences of opinion about the feedback can be resolved.

Feedback may contain corrections of factual inaccuracies and challenges to the conclusions or disagreement with the recommendations. The Evaluation Manager may help to channel feedback and comments between the Steering Group and the evaluation team. All forms of feedback should be addressed before finalising the report. Factual inaccuracies should be corrected and misunderstandings clarified. Where necessary the evaluators would be expected to justify their conclusions by presenting additional evidence to convince any critics. However evaluations often present stakeholders with hard messages and whilst the team leader should be sensitive to the needs of stakeholders and may wish to amend the report to avoid unnecessary misunderstandings, the report should not shy away from presenting the results of the team's analysis honestly.

The Evaluation Manager will check the overall quality of the final report before paying the evaluation team.

The final task of the evaluation team is **to prepare a well-presented copy of the final** report. The use of a copy editor at this stage can avoid wasting the time of the team leader and Evaluation Manager in making minor typographical amendments. Translation may also be required.

Stage III: Using the evaluation

Once the final copy of the report is received from the evaluation team, the Evaluation Manager makes arrangements for wider publication and for following-up the recommendations of the report. The Steering Group members may also be involved in making formal responses to the recommendations. Sometimes a summary of management responses and other follow-up action is published with the report.

The Evaluation Manager will publicise lessons learned by drawing these to the attention of relevant bodies, and to the public through press releases. Reports are usually published on the internet. There are increasingly useful web portals for locating the reports of different agencies.

Given the wide range of possibilities for publishing and disseminating lessons from the evaluation, the Evaluation Manager, programme staff and stakeholders should consider which media would be most appropriate. Ideally this should be done before commissioning the evaluation, as the dissemination efforts need to be fully budgeted for.

Since evaluation exercises are expensive, important findings should not remain locked in unread reports. Too often the same mistakes highlighted in evaluations are repeated in the design and implementation of new development activities because the lessons are not communicated well or fail to be understood by development organisations and their stakeholders. Those commissioning major evaluations should spend some time considering how to ensure that an evaluation's findings are influential. Here are a set of common mistakes to avoid:
- Stakeholders do not feel sufficiently involved in the outcomes of evaluation exercises because the questions they want answered are not built into the evaluation design
- Key stakeholders are not sufficiently consulted or do not feel involved in the evaluation process

- The findings from the evaluations are not fed back in time to inform the next set of decisions
- The evaluation methodology was ill-conceived and corrective actions were not taken
- Reports are too long and too difficult to digest
- The lessons are not easy to understand or are not communicated well enough
- Lessons from past development evaluations may not be perceived as relevant to new contexts, especially as development priorities shift
- Implementers only want to hear about success and do not want to give time to learning from mistakes
- Evaluation findings are rarely fed back to beneficiaries so that they can feel more powerful in designing future interventions implemented in their name

To avoid these mistakes, evaluation needs to be an integral part of the knowledge management and performance culture of development organisations.

Designing the dissemination strategy:

Who are the target audiences?
- Discuss with relevant stakeholders to whom and when the results of the evaluation should be disseminated

How can the different audiences best be reached?
- Take different dissemination options into consideration, including dialogue with partners, seminars or workshops at key moments
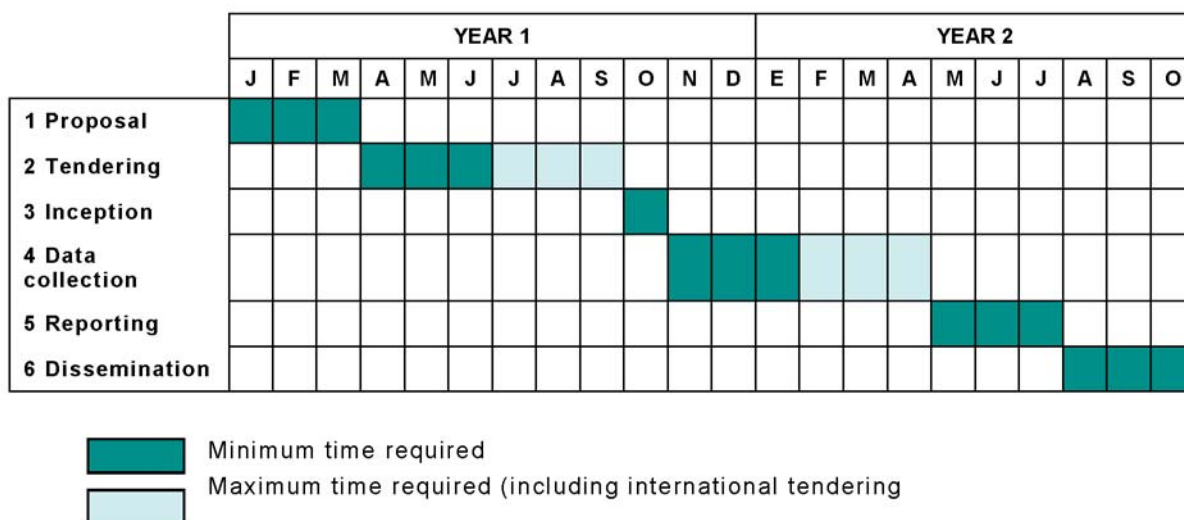
What are the main messages? Who needs to know that?
- Focus the dissemination efforts and tailor them to the specific needs of different audiences

---

4.4   The planning timescale for more complex evaluations

Evaluations are often conducted to inform particular decisions and therefore the timetable for reporting is crucial. Planning an evaluation exercise to reach a particular deadline needs to take account of the time-scale for the whole evaluation exercise.

A typical example of the time-scale for a large-scale evaluation

	YEAR 1												YEAR 2									
	J	F	M	A	M	J	J	A	S	O	N	D	E	F	M	A	M	J	J	A	S	O
1 Proposal	■	■	■																			
2 Tendering				■	■	■	░	░	░													
3 Inception										■												
4 Data collection											■	■	░	░	░							
5 Reporting																	■	■	■			
6 Dissemination																				■	■	■

■ Minimum time required

░ Maximum time required (including international tendering

1. Proposal for the evaluation and agreement on the focus and terms of reference by commissioning partners

2. Completion of the tendering and contracting process

3. Inception period for the contracted evaluation team to plan the detailed methodology and logistics for conducting the evaluation

4. Data collection, document reviews, questionnaires, interviews and visits

5. Report preparation, drafting, consultations on the draft and report finalisation

6. Final agreement on the report, follow-up and official responses to recommendations, translation, publication and dissemination

Chapter 5

Terms of Reference

The Terms of Reference (ToR) lay down the Evaluation Manager's requirements and expectations from the evaluation. The ToR represent the basis of the contract with the evaluators. Well-written ToR can prevent misunderstandings between the range of commissioning stakeholders, the Evaluation Manager, and the evaluation team members as the evaluation progresses. The ToR should be as tightly specified as possible at the start to attract relevant responses from prospective evaluators. In rare cases, where those who are commissioning evaluations prefer the evaluators to design the evaluation with the stakeholders, only an outline is specified in the ToR.

THE MAIN OBJECTIVES OF TERMS OF REFERENCE	**Identify the reasons** for the evaluation, and what is expected to be achieved **Describe the investment** being evaluated and the context for performance to be assessed - Establish the **focus and scope** **Guide the execution** of the evaluation - Set scheduling and **time** frames - Provide the basis for a **cost** projection.

All the stakeholders involved in commissioning the evaluation should have ownership of the contents of the ToR and agree them from the start. Commissioning stakeholders often have conflicting priorities and are exceedingly ambitious about all the areas that they wish the evaluation to cover. When drafting the ToR, Evaluation Managers should focus on what it is feasible to expect the evaluators to do. This is especially important in relation to identifying the impact on poverty reduction and attributing shifts in such indicators to a specific donor's activities.

Individual responsibilities and lines of accountability are an important framework for the evaluation. Therefore it is important to specify the following in the ToR: who the client for the evaluation is and who will be managing the exercise; the role of any stakeholder representatives who may join the team; and the role of reference groups, management groups or steering groups regarding advice or feedback during the evaluation process. An outline of the supporting role that the Evaluation Manager intends to play in terms of logistical support and any administrative arrangements is helpful.

A specification of the deliverables should include the number of copies of the evaluation report required, the length, style and font size of the report and any translation requirements.

Your comprehensive check list for drafting Terms of Reference

Title of the evaluation, that summarises the broad intention
Background and context of the **investment**
✔ **Reasons** for the evaluation
Objectives — indicating the timing of the report if it is to inform specific time-bound decisions

✔ **Focus and scope**

Principle level of focus (policy/strategic, operational, impact)
Specific issues to address — best expressed as key questions
Coverage - estimated number of case studies or site visits

✔ **Stakeholder involvement**
People involved in commissioning the evaluation
People who should be consulted

✔ **Documents available**
✔ **Accountabilities and Responsibilities**
Including the roles and support offered by the evaluation manager and others

✔ **Evaluation process**
Phases, desk study, field synthesis, other methodological issues, contributors

✔ **Deliverables**
Types of reports and/or workshops

✔ **Evaluation team qualifications**
Technical skills, knowledge and experience of the consultants, specific roles of the team leader, local consultants, proof copy editor and translators and interpreters

✔ **Timing**
When the report is due and any key dates for special meetings where report findings are needed for decision-making

✔ **Person days** or overall **cost estimates**
An indication of the total estimated person days or budget limits

Speed check list: Have you ...?

✔ Clarified the use and purpose of the evaluation?
  ➔ How does it fit with DFID's internal policymaking process?
✔ Described clearly the background of the intervention and its logic?
✔ Outlined how stakeholders will be involved during the evaluation?
✔ Consulted colleagues and stakeholders to list evaluation questions based on a set of criteria?
✔ Specified what kind of recommendations and lessons the evaluation intends to provide?
✔ Specified gender issues to be addressed?
✔ Clarified the timing of the individual evaluation steps within the work plan?
✔ Specified the format of the report including number of pages and any other dissemination activities?
✔ Described clearly the skills, knowledge and qualification the evaluators should encompass and when the work needs to be done?

Chapter 6

Evaluation teams – roles and skills

Consultancy companies bidding to work on a particular evaluation will assemble the evaluation team. Individual companies or a consortium of companies may join together to bid for the work. The consultancy company's project manager for the evaluation team is likely to be responsible for assembling the initial Expression of Interest (EOI).

Once the EOIs are evaluated, a smaller range of the best applicants is invited to submit a proposal. If invited, the company's project manager and the evaluation team leader, with inputs from other team members, will submit the final tender documents. The company's project manager will manage the commercial aspects of the evaluation assignment and oversee the overall quality control of the evaluation.

The team leader will manage the team and the evaluation process, and act as the prime interface with the Evaluation Manager who represents the commissioning partners. This is shown in the following diagram:
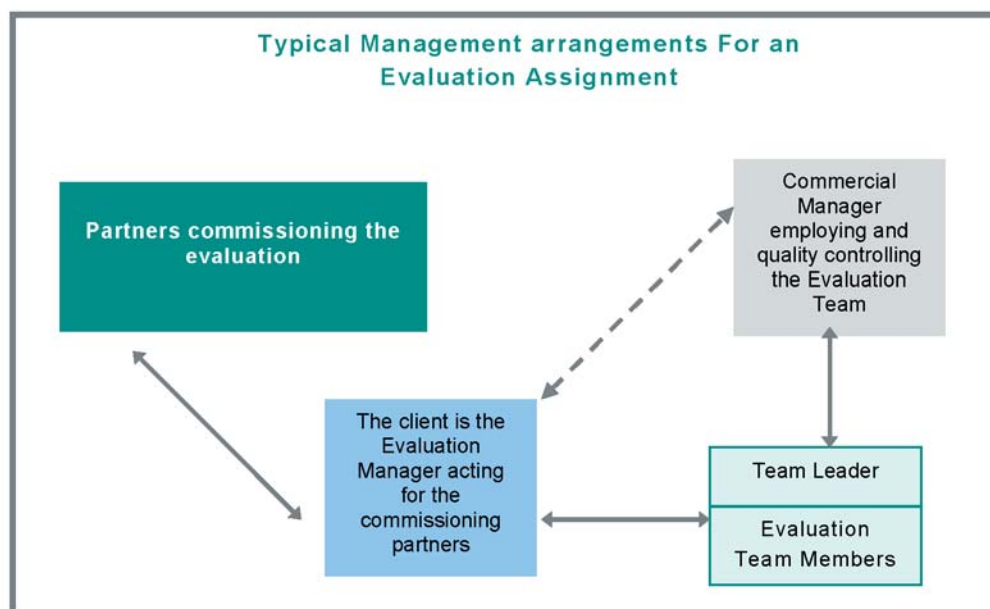
Once the evaluation begins, the Team Leader will:

- work closely with the Evaluation Manager throughout the process to ensure the expectations of the client are met
- manage the team to ensure all aspects of the ToR are met
- oversee the preparation of the evaluation workplan
- oversee the data collection phase, in compliance with international and ethical standards
- facilitate agreement amongst the team on the findings, conclusions and recommendations
- draw together the draft report, present the draft report
- present the report at any workshop required and facilitate feedback
- produce the final report, amended in response to comments
- deliver all the required reports and other outputs on time to the client.

These duties require good interpersonal, management, writing and presentational skills and relevant technical skill in the area of evaluation.

Team members are selected for their technical skills, knowledge and experience and their local understanding. Ideally they would have experience of the evaluation techniques required and know how to work well as part of an international team. There are strong arguments for involving the insights and language skills of local consultants and young consultants who may have limited experience of evaluation. This practice will encourage wider understanding of international evaluation and help build local capacity.

Attention to the overall gender balance of the team is important in any case, but is especially important where access to women respondents may not be possible without women interviewers. The needs of women and girls are easily overlooked in contexts where women's voices are rarely heard and women's rights are not a priority. The presence of a female evaluator often encourages women to participate and express their views. It is important to have at least one evaluator in the team who will take responsibility for ensuring a proper understanding of gender dynamics within the evaluation process.

Evaluation reports need to be accessible and the involvement of copy editors and translators may be necessary to ensure the quality and accessibility of any written material.

To ensure the independence and impartiality of the evaluation process, evaluation team members would normally have no connection with the implementation of the project or programme being evaluated. Representatives of the funders, or the implementing agency may work with the team for part of the time as resource persons to provide information, but would not be expected to be involved in the preparation of the conclusion and recommendations or any project scoring decisions. These roles are explored further in Chapter 9.

Speed check list: Have you …?

✔ Identified a team leader with good interpersonal skills?
  ➔ Are you satisfied with the Team leader's management, writing, presentation and relevant technical skills?
✔ Chosen members of the evaluation team with the relevant evaluation skills, technical knowledge and understanding of the context?
✔ Included local consultants? Remember that each evaluation exercise is a chance to build capacity and to share insights and skills. An additional way of capacity building is to allow new young local consultants to shadow the evaluation team.
✔ Taken gender issues into account?
  ➔ the gender balance of the evaluation team? Does the team possess gender analytical skills? Are gender responsibilities clearly assigned within the team? Guidance on the policy and practice in support of gender equality and women's empowerment are available from EvD's insight pages.
Clarified that a copy-editor and potentially translators need to be included in the team? Don't forget that the evaluation report needs to be easily accessible for all stakeholders!
✔ Double checked that no member of the evaluation team has had any previous involvement in the implementation of the intervention?

Chapter 7

Reporting

Reporting is more than written reports. The aim of reporting is to ensure the stakeholders are briefed about progress and about the outcomes at the end of the evaluation exercise. Written reports are normally supplemented with verbal presentations and discussions at specially convened meetings of the relevant stakeholders or wider public.
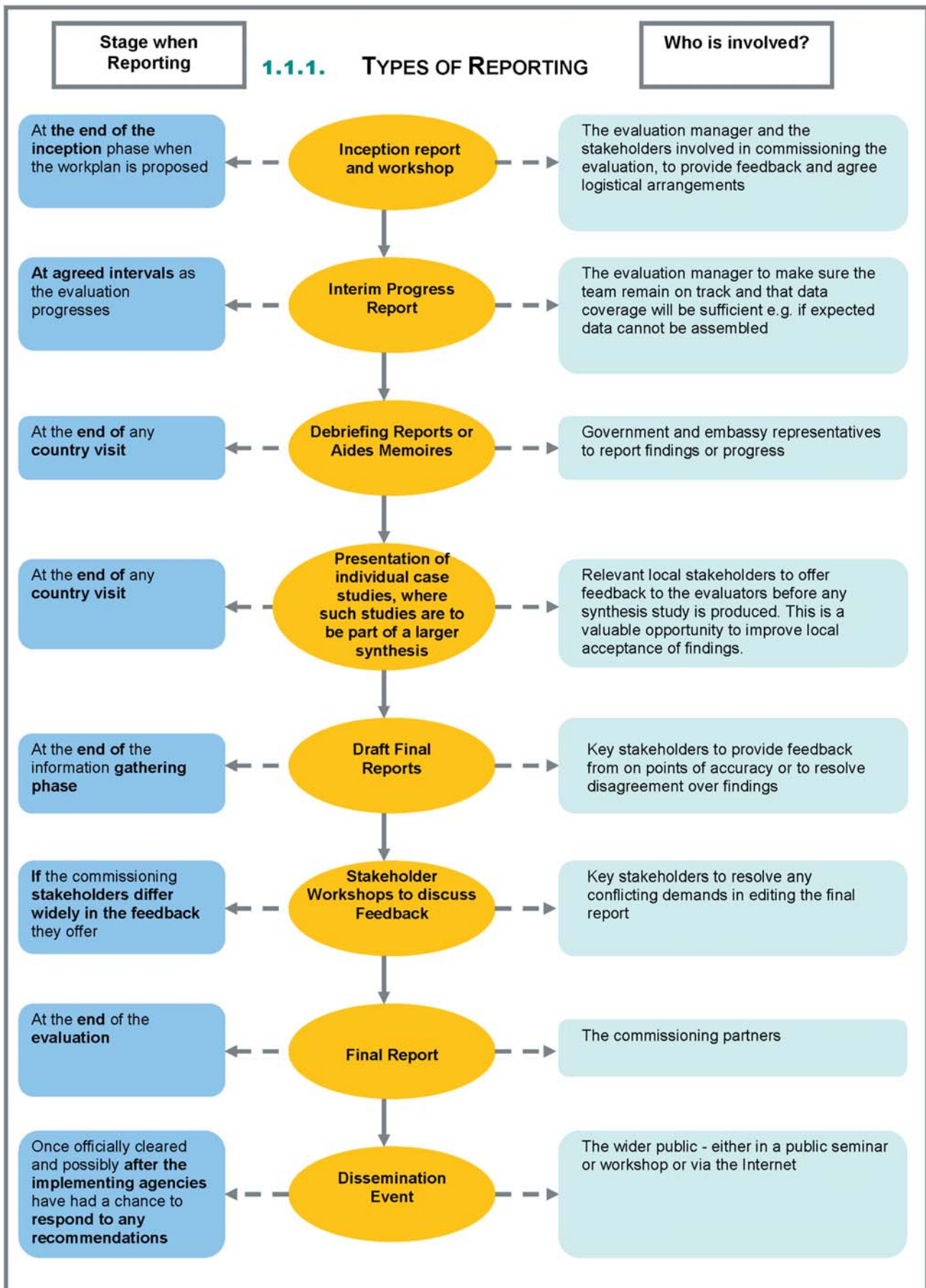
Regular interaction between the evaluation team and the commissioning partners helps to maintain the focus of the evaluation and teases out any problems with data collection or team dynamics as early as possible. Opportunities to reflect on the findings as soon as possible helps the stakeholders to prepare for the conclusions and recommendations and makes hard messages easier to respond to before the final report becomes public. Good feedback opportunities improve accuracy and raise the quality of the final report. The table on the next page summarises the different stages of reporting and the different types of reporting at each stage. And, as an example of best practice, an EvD report template and style guide are available from EvD's insight pages.

The accessibility of the findings, conclusions and recommendations and the constructive reaction of the commissioning partners are crucial tests of the value added by the evaluation. Written reports need to be well-structured, clear and concisely written in relevant languages. Evaluators should not express their own opinions in the section on findings but report the evidence presented to them and then draw their conclusions. A typical report structure is shown in the following table.

Many readers will have limited time to read so the structure of the report and sub-headings should be used to guide the reader to the key areas. Since readers may not be used to reading evaluation reports or may be reading in a language other than their own, the language should be simple, free from jargon and with specialist terms explained.

All reporting should be timely. Good evaluation reports will be wasted if they arrive too late to inform decisions.

After opportunities for feedback, the conclusions and recommendations in the final written report are those of the evaluation team only. Although interested parties may attempt to influence the outcomes of an evaluation report, the professionalism of the evaluation team requires the team to hold their ground where they do not find the evidence for changing their opinions. Robust discussions are to be expected where findings are controversial, but any attempts to intimidate the team should be reported to the Evaluation Manager.

1.1.1. TYPES OF REPORTING

Stage when Reporting	Types of Reporting	Who is involved?
At **the end of the inception** phase when the workplan is proposed	**Inception report and workshop**	The evaluation manager and the stakeholders involved in commissioning the evaluation, to provide feedback and agree logistical arrangements
At agreed intervals as the evaluation progresses	**Interim Progress Report**	The evaluation manager to make sure the team remain on track and that data coverage will be sufficient e.g. if expected data cannot be assembled
At the **end of** any **country visit**	**Debriefing Reports or Aides Memoires**	Government and embassy representatives to report findings or progress
At the **end of** any **country visit**	**Presentation of individual case studies, where such studies are to be part of a larger synthesis**	Relevant local stakeholders to offer feedback to the evaluators before any synthesis study is produced. This is a valuable opportunity to improve local acceptance of findings.
At the **end of** the information **gathering phase**	**Draft Final Reports**	Key stakeholders to provide feedback from on points of accuracy or to resolve disagreement over findings
If the commissioning **stakeholders differ widely in the feedback** they offer	**Stakeholder Workshops to discuss Feedback**	Key stakeholders to resolve any conflicting demands in editing the final report
At the **end of** the **evaluation**	**Final Report**	The commissioning partners
Once officially cleared and possibly **after the implementing agencies** have had a chance to **respond to any recommendations**	**Dissemination Event**	The wider public - either in a public seminar or workshop or via the Internet

A Sample Final Evaluation Report Structure	
Title page	✔ with the title, date of publication and names of the authors responsible for the report, and with clear reference to commissioning partners and where the document can be found electronically
Preface	✔ by the evaluation manager acting for the commissioning partners
Map	✔ if helpful to the reader
Contents page	✔ indicating the sections and annexes for easy reference
Acronyms and Abbreviations	✔ these are usually explained in full on the first occasion that they are used in the text to assist the reader
Acknowledgements	✔ thanking those who have contributed to the study
Executive summary	✔ as brief as possible to ensure it is read by busy officials e.g. the purpose, context and coverage of the evaluation (1 paragraph), the methodology (1 paragraph), the main findings, lessons and recommendations may be written in bold or italicised with brief explanatory paragraphs to follow
The main report	✔ should be illustrated where possible by diagrams to minimise the text. It is usually laid out as follows:
1. Introduction	➔ the purpose, scope and focus of the evaluation ➔ any limitations of the evaluation design gained in retrospect ➔ the policy context of the programme ➔ the programme's relative size in terms of the budget in ➔ relation to the overall aid programme ➔ a brief description of the project its logic and assumptions ➔ an explanation of the structure of the report ➔ introduction of the team
2. Methodology (how the evaluation has been undertaken):	✔ phases in the data collection (desk study, field visits) ✔ reasons for selection (of the time-slice in the life of the programme, or of the countries or sites or case studies chosen for detailed examination) ✔ how information is collected (use of questionnaires, official data, interviews, focus group discussions and workshops) ✔ problems encountered, such as key people not available for interview or documents not available
3. Findings: (evidence)	✔ reporting of the data, (what happened and why, what actual outcomes were achieved in relation to those intended, what positive or negative, intended or unintended impact, what effects on target groups and others — organised by DAC criteria, issues, or key questions in the terms of reference)

4. Conclusions (justified and arising logically from the findings)	✔ summary of achievements against the initial project design or logical framework ✔ summary of problems encountered and the reasons for this ✔ overall effect on poverty and cross-cutting issues ✔ why things happened as they did, questioning the assumptions, design, implementation, management and partnership arrangements
Lessons:	✔ lessons at the level of policy, strategy, design, implementation and management that may have implications for future work
Recommendations	✔ a short number, succinctly expressed, and addressed to those with the means and responsibility for implementing them
Appendices/ Annexes	✔ terms of reference ✔ schedule ✔ list of people consulted ✔ list of documents consulted ✔ sub-reports of country visits and case studies which formed part of the whole evaluation and have been drawn upon to produce the main report ✔ details of the members of the evaluation team

Speed check list:

When you are holding the first draft of the evaluation report in your hands, look at the questions below: Have you…?

✔ Checked that the report is well-structured, clear and concisely written?

✔ Been helped to understand the report through the use of:

➔ Graphics, tables and illustrations?

➔ Use of simple non-specialist language, free of jargon?

➔ Appropriate headings and subheadings?

✔ Checked that findings, conclusions and recommendations are not mixed up putting the credibility of the evaluation at risk?

Findings are based on evidence and not the opinion of the evaluator?

➔ Conclusions and recommendations arise logically from findings?

➔ Findings, conclusions and recommendations for specific gender issues are clearly stated?

✔ Made sure that the limit of page numbers is not heavily exceeded?

Chapter 8

Using evaluations – and sharing lessons

When planning an evaluation exercise, the evaluation manager should think about the purpose of the evaluation and who the main users of findings and lessons may be. Involving the potential users of the findings, in the evaluation design, is the best way to ensure interest in the results. If those who want to use the evaluation results are involved in commissioning and designing the evaluation, they can specify key questions that they want the evaluation to answer.

Users may want to decide whether to extend the programme into another phase, or to know how to replicate the project methodology in new contexts. They may wish to use the conclusions and lessons to inform alternative policy formation. The first use of a completed evaluation exercise is therefore to share results with those commissioning the study. This is usually done by circulating the report at the draft stage and requesting feedback individually, or as part of consultation workshops.

The process of conducting the evaluation can also be useful in sharing the learning from the evaluation. Local officials who wish to learn more about the programme and how it has been implemented may accompany the evaluation team. Shadowing or being part of the evaluation team can be a valuable opportunity to learn about the evaluation process itself. Lesson learning can start during the evaluation process as funders and implementers reflect on what they have been doing as they answer the questions the evaluators raise with them.

Lessons can be shared beyond those immediately involved in the evaluation through planned dissemination activities. These are likely to include circulating the report or tailoring information derived from the report for particular interested parties, arranging dissemination events, and organising press releases. The planning and management of dissemination processes may be specified in the terms of reference and built into the workplan of the evaluation team.

In some cases it may be sensible to ensure that a management response to the report is prepared before the report is published, so that this can be released with the report to show changes being implemented as a result of the recommendations.

Primary beneficiaries rarely receive feedback on the conclusions of an evaluation. This empowering additional step should be built into evaluation exercises as those affected by the project or programme are in a good position to see how work done in their name can be influenced to yield better results in future.

The executive summary is the essential part of the report for most stakeholders and interested members of the public in the North and the South. It should be short and easily digestible. It is useful to publish the executive summary separately. A translation of the executive summary into the language of the host country is also useful and cheaper than translating the whole report. This is common practice for a number of bilateral donor agencies.

Parts of the evaluation report might serve as case studies for training purposes. UNICEF is currently using evaluation reports for the training of its staff. Alternative media for publishing evaluation results are videos, web-broadcasts, PowerPoint presentations, radio and television.

Speed check list: Have you ...?

✔ Ensured that the findings and lessons from the evaluation are targeted to potential users?

✔ Taken into account the needs of key stakeholders: Funders and implementers, local officials and primary beneficiaries?

→ Checked that the evaluation is well-conducted and the report is written in a constructive manner in order to be used by funders and implementers?

→ Offered local officials the chance to be part of team?

→ Assured that primary beneficiaries are consulted during the evaluation and receive feedback on the conclusions of the evaluation?
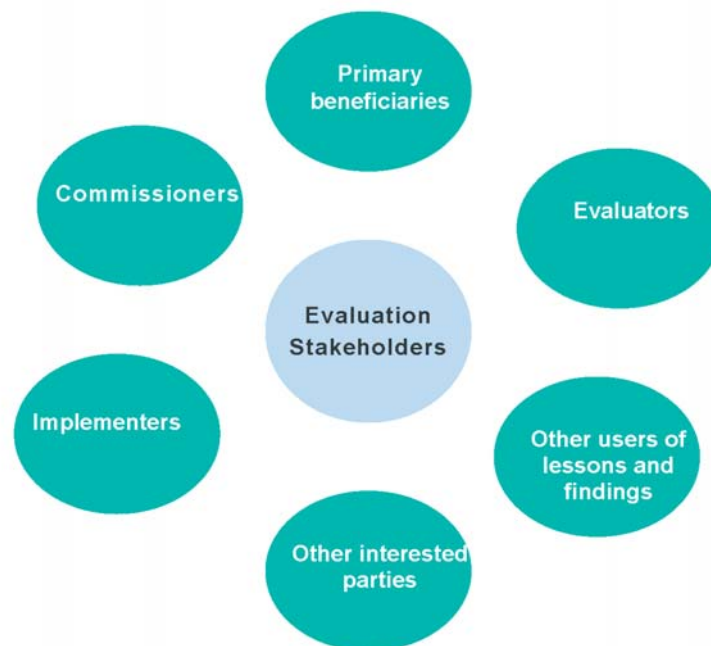
Chapter 9

Different stakeholders and their roles

9.1   Who are the stakeholders in an evaluation exercise?

A large number of people are involved in, or affected by, an evaluation. They are referred to as stakeholders.

Evaluation Stakeholders

Commissioners

Those commissioning evaluations usually involve funders and recipients, and other development partners who have a prime interest in the results of the development intervention. This commissioning group decides the questions they want the evaluation team to answer and work together to plan and conduct the evaluation. The commissioners are normally responsible for responding to any recommendations made in the final report. One or more representatives of the commissioning group will also act as the Evaluation Manager, responsible for liaising between the commissioning group and the evaluation team.

Primary beneficiaries

The main intended beneficiaries of the development intervention will want the opportunity to express their views and to hear the results of the evaluation. Beneficiaries often have the least powerful voice in evaluations. Their views should be specifically sought and listened to. It may also be necessary to ensure the voice of women and other less powerful groups are distinguished when consulting beneficiaries.

Implementers

Partner governments and institutions, donor agencies, advisors, and consultants as well as the project and programme implementation team may all be involved in the implementation of the particular development intervention being evaluated. In an independent evaluation, all these staff members may contribute their perspectives but they should not be responsible for the final conclusions and recommendations, which would be made independently by the impartial evaluation team members.

Evaluators

Inter-disciplinary teams of local and international consultants who are responsible for conducting the evaluation, producing the report and feeding back their findings, conclusions and recommendations. Partner agency staff may also join the evaluation team to participate in meetings and visits to absorb the learning derived from the evaluation process, and to act as resource persons, to contribute reports and other data. As with implementers, resource persons may contribute perspectives, but the independent and impartial evaluators alone should be responsible for the conclusions and recommendations.
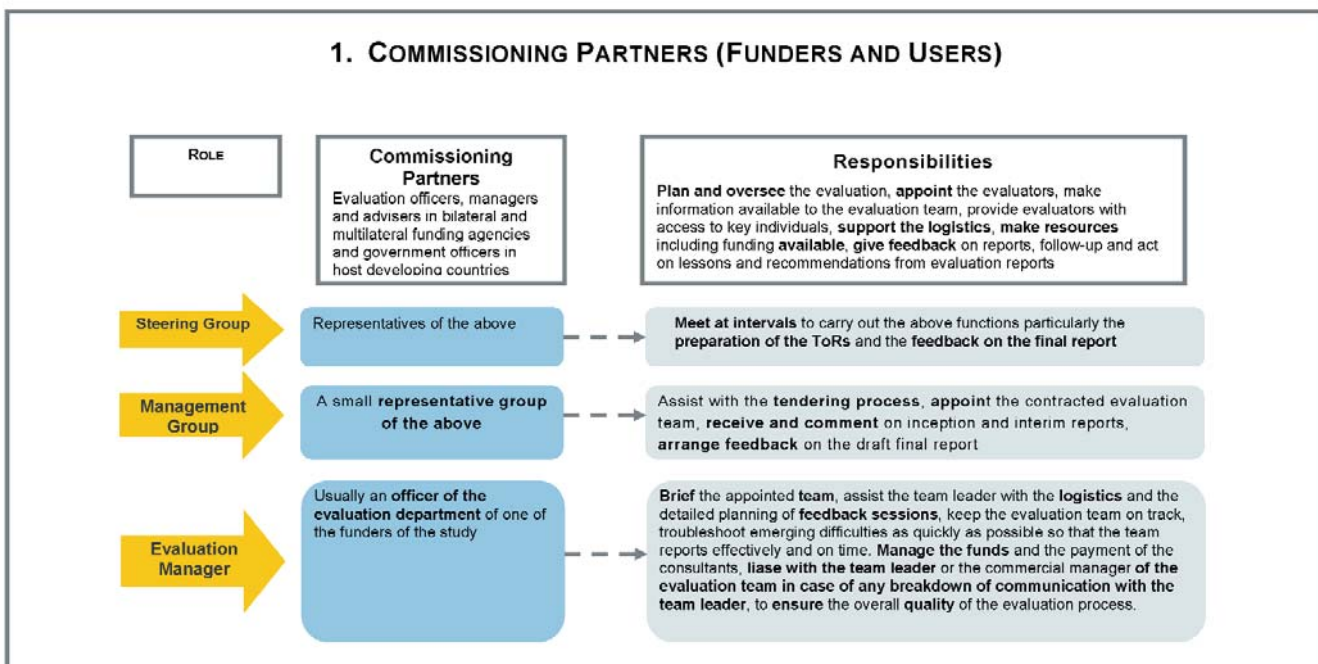
Other users and interested parties

Representatives of affected institutions, the wider international development community, policy-makers, planners, agency officials, consultants, academics, local development workers, may all have an interest in the outcomes of the evaluation. The Evaluation Manager should consider disseminating the final evaluation report to members of this group or at least bringing it to their attention.
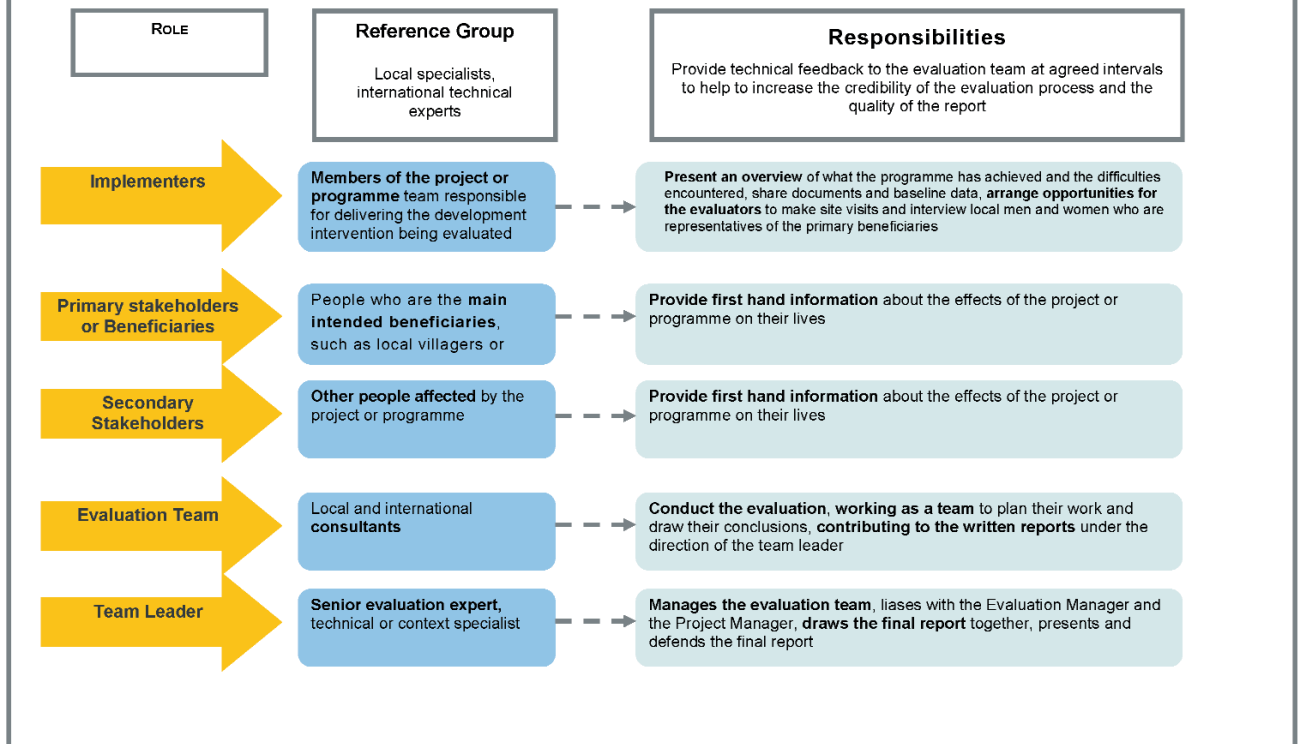
All stakeholders to the evaluation need to be treated with sensitivity, especially where the findings are controversial or affect any one group adversely. The proper involvement of all stakeholders improves the quality and use of the evaluation, but maximising participation can also make evaluation exercises time-consuming and expensive. The commissioners must balance participation and cost. The evaluation manager and evaluation team manager must ensure that stakeholders' time and resources are not over-taxed during the process of the evaluation.

9.2 Clarifying roles and responsibilities in complex evaluation exercises

Multi-stakeholder evaluations are complex exercises to manage. Many of the difficulties can be avoided if the roles and responsibilities of the different individuals and organisations involved are clear from the start, as shown in the following diagram:

1. COMMISSIONING PARTNERS (FUNDERS AND USERS)

ROLE	Commissioning Partners Evaluation officers, managers and advisers in bilateral and multilateral funding agencies and government officers in host developing countries	Responsibilities **Plan and oversee** the evaluation, **appoint** the evaluators, make information available to the evaluation team, provide evaluators with access to key individuals, **support the logistics**, **make resources** including funding **available**, **give feedback** on reports, follow-up and act on lessons and recommendations from evaluation reports
Steering Group	Representatives of the above	**Meet at intervals** to carry out the above functions particularly the preparation of the ToRs and the **feedback on the final report**
Management Group	A small **representative group of the above**	Assist with the **tendering process**, **appoint** the contracted evaluation team, **receive and comment** on inception and interim reports, **arrange feedback** on the draft final report
Evaluation Manager	Usually an **officer of the evaluation department** of one of the funders of the study	**Brief** the appointed **team**, assist the team leader with the **logistics** and the detailed planning of **feedback sessions**, keep the evaluation team on track, troubleshoot emerging difficulties as quickly as possible so that the team reports effectively and on time. **Manage the funds** and the payment of the consultants, **liase with the team leader** or the commercial manager **of the evaluation team in case of any breakdown of communication with the team leader**, to **ensure** the overall **quality** of the evaluation process.

2. REFERENCE GROUPS

ROLE	Reference Group Local specialists, international technical experts	Responsibilities Provide technical feedback to the evaluation team at agreed intervals to help to increase the credibility of the evaluation process and the quality of the report
Implementers	**Members of the project or programme** team responsible for delivering the development intervention being evaluated	**Present an overview** of what the programme has achieved and the difficulties encountered, share documents and baseline data, **arrange opportunities for the evaluators** to make site visits and interview local men and women who are representatives of the primary beneficiaries
Primary stakeholders or Beneficiaries	People who are the **main intended beneficiaries**, such as local villagers or	**Provide first hand information** about the effects of the project or programme on their lives
Secondary Stakeholders	**Other people affected** by the project or programme	**Provide first hand information** about the effects of the project or programme on their lives
Evaluation Team	Local and international **consultants**	**Conduct the evaluation**, **working as a team** to plan their work and draw their conclusions, **contributing to the written reports** under the direction of the team leader
Team Leader	**Senior evaluation expert,** technical or context specialist	**Manages the evaluation team**, liases with the Evaluation Manager and the Project Manager, **draws the final report** together, presents and defends the final report

---

Speed check list: Have you…?

✔ Clarified bilaterally all the roles and responsibilities of the different stakeholders in the evaluation, particularly between yourself as the Evaluation Manager in relation to the Steering Group, Management Group and the Evaluation Team?

✔ Ensured clear lines of communication with the evaluation team through the team leader and the commercial manager of the consultancy team?

✔ Clarified who shares responsibility for the overall quality control of the evaluation?

Annex 1: International development evaluation standards

The evaluation process should be impartial and independent from both the policy-making process and the delivery and management of development assistance.

Independence is best achieved where evaluation activities are independent from managers who have an interest in showing accomplishment and good performance, or any other decision-makers for whom evaluation raises a conflict of interest.

In considering what should be covered in evaluating development assistance, OECD's Development Assistance Committee has agreed a set of standard international criteria to guide all evaluations of development assistance. The DAC criteria are shown below:

DAC Criteria	
Relevance	The extent to which the objectives of a development intervention are consistent with beneficiaries' requirements, country needs, global priorities and partners' and donors' policies.
Effectiveness	The extent to which the development intervention's objectives were achieved, or are expected to be achieved, taking into account their relative importance.
Efficiency	A measure of how economically resources/inputs (funds, expertise, time, etc.) are converted to results.
Impact	Positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended.
Sustainability	The continuation of benefits from a development intervention after major development assistance has been completed. The probability of continued long-term benefits. The resilience to risk of the net benefit flows over time.

Taken together, each of these five analytical measures can be applied to every development intervention. They are applicable at all administrative or aggregation levels — for example at the project, programme or global level — and to all forms of development assistance such as technical or humanitarian assistance. However, in order to be realistic, it is not necessary for all five criteria to be applied at every level in every evaluation. An evaluation of a vaccine research programme for example may examine the relevance, effectiveness and efficiency of the programme rather than the question of impact that might not be answerable until many years after the development phase is ended.

The combination of these criteria should provide decision-makers with the essential information to assess whether an intervention has been successful.

Covering all criteria in all evaluations may be an ideal but is not always practical. The evaluation may be conducted at too early a stage to assess impact or sustainability in the longer term. It should always be possible to assess relevance, effectiveness and efficiency as minimum criteria and to look for early signs of impact and sustainability, or at least establish the benchmarks or baselines against which longer term impact could be assessed.

Relevance

Relevance means the extent to which a development intervention conforms to the needs and priorities of the beneficiaries, and the policies of recipient countries and donors. Relevance implies a consistent fit with the priorities of the principle stakeholders and beneficiaries.

At the policy level, the intervention needs to be in line with the partner governments' policies and priorities. From a bilateral donor's perspective, taxpayers contribute funds for development assistance on the understanding that it will be spent in accordance with the policies for which they have voted. At another level the objectives of the intervention must be in line with the expectations of the beneficiaries in whose name the work is being done.

Reviewing relevance is a useful way of checking that the design of the programme was appropriate, since the activities and outputs of the interventions should be relevant to the overall outcomes and purpose of the intervention. The means of achieving the outcomes should be relevant to the overall purpose.

Thinking about relevance is important throughout the life cycle of the intervention from its design through to implementation and afterwards when looking back during the final evaluation. One key question when assessing relevance is whether the intervention has a potential for replication.

> *For example: in a road project, relevance could be assessed in terms of the rationale for constructing the road: was it to serve a political agenda of the few or to exploit real economic potential. In a sector programme to support agriculture, relevance could be assessed in terms of domestic market responses to the various programme initiatives etc .*

Standard questions regarding relevance are:
- Are objectives in line with needs, priorities and partner government policies?
- Is the intervention in line with the livelihoods strategies and cultural conditions of the beneficiaries?
- Is the design of the intervention relevant to the context?
- Is the timing of the intervention relevant from the point of view of the beneficiaries?
- Do proposed interventions have a potential for replication?

Effectiveness

Effectiveness is a measure of the extent to which a development intervention has achieved its objectives, taking their relative importance into account. While measuring efficiency is looking more from the perspective of implementers, measuring effectiveness means taking the perspectives of the beneficiaries of a development intervention into account.

Effectiveness can be measured at the level of the outputs and outcomes attributable to the intervention. Questions relating to effectiveness need to be carefully tailored to the specific level of objectives. In the case of the output level, effectiveness refers to the achievement of inputs in line with the objectively verifiable indicators agreed at the outset. This level is much easier to assess than the achievement of outcomes or impact that assess the successful achievement of the overall purpose of the intervention.

There are two steps to assessing effectiveness at the outcome level:

Step 1: Measurement of the achievement of the objectives
- ■ Has the development intervention achieved what it set out to do?
- ■ How have the conditions of the intended beneficiaries group changed since the beginning of the development intervention?
- ■ How do these identified changes compare with the intended changes?

Step 2: Clarifying causal attribution
- ■ To what extent have the identified changes been caused by the development intervention rather than external factors?

For example: In a road project, effectiveness could be measured in terms of traffic flow, or transport of different commodities, better market or school access for uses taking the opinions of the local community users of the road into account. In a sector programme to support agriculture, effectiveness could be measured in terms of the scope of productive investments made by the farmers involved, and the sustainable development of human resources through training and institutional development in the sector.

Standard questions regarding effectiveness that could be applied during mid term evaluations are:

- ■ To what extent have the agreed objectives been achieved?
- ■ Are the successfully achieved activities sufficient to realise the agreed outputs?
- ■ To what extent is the identified outcome the result of the intervention rather than external factors?
- ■ What are the reasons for the achievement or non-achievement of outputs or outcomes?
- ■ What could be done to make the intervention more effective?

Efficiency

Efficiency is comparing the outputs against the inputs. It illustrates the relation between means and ends. The underlying question to ask when looking at efficiency is this: to what extent can the costs of a development intervention be justified by its results, taking into account alternatives? Does the intervention represent the quickest way to transform aid money into development gains, whilst minimising unnecessary transaction costs? In multi-lingual international meetings it is sometimes difficult to explain the difference between the English terms 'efficiency' and 'effectiveness'. The term 'cost-effectiveness' is sometimes used instead of 'efficiency'.

Optimal efficiency is reached when the value of a development outcome outweighs the cost of the inputs and effort required, and is greater than it would have been if compared with using these same resources in different ways.

This raises the question of what standards to adopt as a reference. The standard will rarely be predetermined and will depend either on the terms of reference given to the evaluation team or the evaluator's own professional judgement.

The following criteria can serve as benchmarks of efficiency:

- ■ Comparable development activities
- ■ Best practice standards or
- ■ Standards which are considered to be reasonable

For example: in a road-building project efficiency could be measured in terms of the quality, cost and the timeliness of construction of the road (the physical output) in comparison with other local road building costs and practices

Standard questions regarding efficiency are:

- ● What measures have been taken during the planning and implementation phase to ensure that resources are efficiently used?
- ● To what extent have the development components been delivered as agreed?
- ● Could the intervention have been done better, more cheaply or more quickly?
- ● Could an altogether different type of intervention have solved the same problem at a lower cost?

Impact

The term impact describes the totality of the effects of a development intervention, positive and negative, intended and unintended. In this respect, impact is wider and distinct from effectiveness where the focus is on intended and positive effects. The assessment of impact starts from the perspective of the intended beneficiaries of a development intervention, and moves on to see how others are affected by the intervention.

When studying impact, it is important to clarify what effects the development intervention is responsible for in relation to other external factors (causal attribution). Process-oriented methods such as participant observation may serve this purpose.

Different impacts may appear at different times and it may be necessary to distinguish between short-term and longer-term effects.

As with effectiveness, changes can be measured against baseline data that describes the situation before a development intervention took place. Timelines, critical incident reviews and documented audit trails are methods of recording change processes.

Missing baseline data is a common problem. It hampers the assessment of impact. In some cases memory recall is applied to reconstruct a baseline, although this is a less reliable method than a baseline assembled before the intervention started. When assessing impact, it is also important to consider that a baseline assembled before the intervention is usually part of the design process and is likely to focus on areas where positive changes and signs of impact were expected. At this point it might be useful to specify different types of negative, unintended impacts:

Mistakes in targeting
- Has the intervention resulted in benefits going to unintended beneficiaries? Is there a coverage bias in terms of the intended beneficiaries? For example, have children, women, elderly or disabled people been excluded?
- Have outsiders seized benefits intended for the primary beneficiaries?

Substitution and displacement
- Do subsidised workers, for example, replace unsubsidised workers who would otherwise have been employed?
- Has the creation of subsidised jobs in one firm or sector led to the reduction of workers in other firms or sectors?

Recoil effects
- Are recipient country organisations or systems overburdened by an externally financed development intervention?

Fungibility
- Is it possible that by lending support for a particular activity in a partner country, this has enabled the partner country's resources to be shifted to other activities that are inconsistent with donor priorities or policies, like military expenditure?

Perverse effects
- Is donor support increasing the dependency of a partner country on external aid flows, weakening local systems or increasing unsustainable debt?

For example: in a road project, the impact may be measured in terms of economic benefits and losses that accrue to changes in traffic flow and transport capacity, through new settlement patterns, opening new areas to agricultural production, increased pollution, etc. In a sector programme to increase farm incomes and better family health, positive an and negative employment effects, environmental effects, economic dependency etc.

Standard questions regarding impact are:
- How has the intervention affected the wellbeing of different groups of stakeholders?
- What would have happened without the intervention?
- What are the positive and negative effects?
- Do the positive effects outweigh the negative ones?
- What do the beneficiaries and other stakeholders perceive to be the effects of the intervention on themselves?
- To what extent does the intervention contribute to capacity development and the strengthening of institutions?

Sustainability

The term sustainability describes the extent to which the positive effects of the development intervention continue after the external assistance has come to an end.

The four evaluation criteria described above primarily concern the specific development intervention, whereas an assessment of sustainability is looking at the long-term effects of the development process. Sustainability might be regarded as a tougher test of whether a development intervention has been a success, since the sustainability of a development intervention depends on local management and local resource allocation continuing once the intervention ends. This partly depends on the degree of local ownership the intervention has achieved.

> *For example, in a road construction project, sustainability can be measured in terms of whether the road is likely to be maintained, the extent to which it will be needed in the future, etc. In a sector programme to support agriculture, it could be measured in terms of financial and economic viability of the agricultural production and the supporting institutions, the extent to which economic surplus is reinvested productively by farmers, etc.*

Standard questions regarding sustainability are:

- To what extent does the positive impact justify continued investments?
- Are the stakeholders willing and able to keep facilities operational and to continue activities on their own?
- Is there local ownership?
- Did partner country stakeholders participate in the planning and implementation of the intervention to ensure local engagement from the start?
- Do relevant partner country institutions possess sufficiently strong governance structures and professional capacity to sustain the activity?
- Is the technology utilised in the intervention appropriate to the economic, social and cultural conditions in the partner country?
- Is the intervention harmful to the environment?

In development co-operation sustainability does not have the same meaning for short-term emergency and humanitarian assistance as in interventions with long-term development objectives. While the criterion may be the same, the standards will differ. The following annex provides further detailed information for this specific area of evaluation:

Annex 2: International development evaluation standards applied to humanitarian assistance

Criteria for evaluation of humanitarian assistance
(Source: DAC: Guidance for evaluating humanitarian assistance in complex emergencies (Paris, 1999)

Efficiency

Measures the outputs — qualitative and quantitative — in relation to the inputs. This generally requires comparing approaches to achieving the same outputs, to see whether the most efficient process has been used. Cost-effectiveness is a broader concept than efficiency in that it looks beyond how inputs were converted into outputs, to whether different outputs could have been produced that would have had a greater impact in achieving the project purpose.

Effectiveness

Measures the extent to which the activity achieves its purpose, or whether this can be expected to happen on the basis of the outputs. Implicit within the criteria of effectiveness is timeliness (for if the delivery of food assistance is scientifically delayed the nutritional status of the population will decline). There is value in using it more explicitly as one of the standard criteria because of its importance in the assessment of emergency programmes. Similarly, issues of resourcing and preparedness should be addressed.

Impact

Looks at the wider effects of the project — social, economic, technical, environmental — on individuals, gender and age-groups, communities, and institutions. Impact can be immediate and long-range, intended and unintended, positive and negative, macro (sector) and micro (household). Impact studies address the question: what real difference has the activity made to the beneficiaries? How many have been affected?…

Relevance

is concerned with addressing whether the project is in line with local needs and priorities (as well as with the donor policy). An evaluation of humanitarian assistance replaced the criteria of relevance with the criteria of appropriateness — the need "to tailor humanitarian activities to local needs, increasing ownership, accountability, and cost-effectiveness accordingly (Minear, 1994). However, the two criteria complement rather than substitute each other. 'Relevance' refers to the overall goal and purpose of a programme, whereas 'appropriateness' is more focused on the activities and inputs. The expansion of the criteria draws attention to the fact that even where the overall programme goal is relevant — for example to improve nutritional status — there are still questions to be asked about the programme purpose. Distributing large quantities of food aid may not be the best way of improving nutritional status. Alternatives could include food for work, cash for work, or measures to improve the functioning of local markets. Furthermore, even if distribution of food aid is deemed appropriate, it is still necessary to examine the appropriateness of the food that is distributed …

Sustainability

is of particular importance for development aid — is concerned with measuring whether an activity or an impact is likely to continue after donor funding has been withdrawn. Projects need to be environmentally as well as financially sustainable. However, many humanitarian interventions, in contrast to development projects, are not designed to be sustainable. They still need assessing, however, in regard of whether, in responding to acute and immediate needs, they take the long-term into account. Larry Minear has referred to this as Connectedness, the need 'to assure that activities of a short-term emergency nature are carried out in a context which takes longer-term and interconnected problems into account' (Minear, 1994) …

(Examples of the problems to be dealt with are the environmental effects of refugee camps, the damage of roads through food transports, the damage to local institutions as a result of international NGOs taking over central government functions and recruiting skilled staff from local government institutions.)

Coverage

includes the need 'to reach major population groups facing life-threatening suffering wherever they are, providing them with assistance and protection proportionate to their need and devoid of extraneous agendas'. Minear alerts evaluators that complex emergencies and associated humanitarian programmes can have significantly differing impacts on different population sub-groups … Programmes need to be assessed both in terms of which groups are included in a programme, and the differential impact on those included.

Coherence

refers to policy coherence, and the need to assess security, developmental, trade and military policies as well as humanitarian policies, to ensure that there is consistency and, in particular, that all policies take into account humanitarian and human right considerations. … Coherence can also be analysed solely within the humanitarian sphere — to see whether all the actors are working towards the same basic goals …

Co-ordination

could be considered under the criteria of effectiveness, for a poorly co-ordinated response is unlikely to maximize effectiveness or impact. However, given the multiplicity of actors involved in an emergency response, it is important that co-ordination is explicitly considered — the intervention of a single agency cannot be evaluated in isolation from what others are doing, particularly as what may seem appropriate from the point of view of a single actor, may not be appropriate from the point of view of the system as a whole.

Protection

Given the context of conflict and insecurity, protection issues are also critical to the effectiveness of humanitarian action. Where levels of protection are poor it is feasible that the target population of an otherwise effective project distributing relief assistance are being killed by armed elements operating within the project area or even within the displaced persons/refugee camp. Assessment of the levels of security and protection in the area of the project or programme and, where relevant, the steps taken to improve them should be part of all humanitarian assistance evaluations.

Annex 3: Donor quality standards in international development

Different donors have developed further quality standards to act as guiding principles for evaluators and managers of evaluation studies. These are summarised in the following table:

Adds value	*Process and outcomes add to overall knowledge and transfer that knowledge to partners effectively*
Ethical and professional	*Evaluators treat all stakeholders with respondents with sensitivity and respect, observing confidentiality and maintaining the principles of independence, neutrality, transparency and fairness throughout.*
Credible	*Conclusions and recommendation derive from reliable and valid findings and analysis based on adequate and triangulated data collection derived from a rigorous and transparent process.*
Useful	*Utilisation–focused throughout, ensuring learning informs decision-making*
Accessible	*Reports and presentations are clear and succinct and well adapted to their audiences*
Cost-beneficial	*The costs of evaluation are proportional to the budget committed to the development intervention being evaluated, and remain within budgetary limits. Resources used with care.*
Meets client expectations	*Meets contractual obligations laid out in the terms of reference and agreed in the workplan.*
The evaluation partnership works well	*Sustains the spirit of partnership with the evaluation managers.*
Timely	*Deliver outputs on time.*

Annex 4: Ethical standards for international development evaluators

SUMMARY OF INTERNATIONAL EVALUATOR ETHICAL STANDARDS	
Cultural Intrusion	Local customs regarding dress, personal interaction, religious beliefs and practices should be respected.
Anonymity / confidentiality	Evaluators must respect people's right to provide information in confidence, and must ensure that sensitive information cannot be traced to its source.
Individual responsibilities	All team members should have an opportunity to disassociate themselves from particular findings and recommendations. The report should acknowledge any unresolved differences of opinion within the team. .
Participation / privacy	Evaluators should provide maximum notice, minimise demands on time, and respect people's rights to privacy or refusal to be involved
Fundamental values	While evaluators should respect other cultures, they must also be aware of international values regarding minorities, women, children etc
Omissions and wrong-doing	The evaluation team should consult with the evaluation manager when there is any doubt about if and how issues, such as evidence of wrongdoing, should be reported.
Evaluation of individuals	Evaluators are not expected to evaluate individuals, and must balance an evaluation of management functions with this general principle.
Disclosure	Briefings and unofficial summaries may be shared, with the permission of the evaluation manager, as part of the transparent evaluation process.
Integrity	Principles of independence, impartiality and accuracy are paramount.

Annex 5: Glossary of standard DAC evaluation terms

(This list is derived from OECD DAC definitions supplemented with additional terms and definitions from Danida and DFID)

Accountability

Obligation to demonstrate what has been achieved, conducted in compliance with agreed rules and standards, fairly and accurately on performance results vis a vis mandated roles and/ plans. This may require a careful, legally defensible, demonstration that work is consistent with the contract .

Note: Accountability in development may refer to the obligations of partners to act according to clearly defined responsibilities, roles and performance expectations, often with respect to the prudent use of resources. For evaluators, it connotes responsibility to provide accurate, fair, credible monitoring reports and performance assessments. For public sector managers and policy-makers, accountability to taxpayers/citizens.

Activity

Actions taken or work performed through which inputs, such as funds, technical assistance and other types of resources are mobilised to produce specific outputs.

Related term: development intervention.

Analytical tools

Methods used to process and interpret information during an evaluation.

Appraisal

An overall assessment of the relevance, feasibility and potential sustainability of a development intervention prior to a decision of funding.

Note: In development agencies, banks, etc., the purpose of appraisal is to enable decision-makers to decide whether the activity represents an appropriate use of corporate resources.

Related term: ex-ante evaluation

Assumptions

Hypotheses about factors or risks which could affect the progress or success of a development intervention.

Note: Assumptions can also be understood as hypothesised conditions that bear on the validity of the evaluation itself, e.g. about the characteristics of the population when designing a sampling procedure for a survey. Assumptions are made explicit in theory based evaluations where evaluation tracks systematically the anticipated results chain.

Attribution

Ascribing a causal link between observed (or expected to be observed) changes and a specific intervention.

Note: Attribution refers to that which is to be credited for the observed changes or results achieved. It represents the extent to which observed development effects can be attributed to a specific intervention or to the performance of one or more partners, taking account of other interventions, (anticipated or unanticipated) confounding factors, or external shocks.

Baseline study

An analysis describing the situation prior to a development intervention, against which progress can be assessed or comparisons made.

Benchmark

Reference point or standard against which performance or achievements can be assessed.

Note: A benchmark refers to the performance that has been achieved in the recent past by other comparable organisations, or what can be reasonably inferred to have been achieved in the circumstances.

Beneficiaries

The individuals, groups, or organisations, whether targeted or not, that benefit, directly or indirectly, from the development intervention.

Related terms: reach, target group.

Cluster evaluation

An evaluation of a set of related activities, projects and/or programmes.

Conclusions

Conclusions point out the factors of success and failure of the evaluated intervention, with special attention paid to the intended and unintended results and impacts, and more generally to any other strength or weakness. A conclusion draws on data collection and analyses undertaken, through a transparent chain of arguments.

Counterfactual

The situation or condition, which hypothetically may prevail for individuals, organisations, or groups, were there no development intervention.

Country Programme Evaluation/ Country Assistance Evaluation

Evaluation of one or more donor's or agency's portfolio of development interventions, and the assistance strategy behind them, in a partner country.

Data Collection Tools

Methodologies used to identify information sources and collect information during an evaluation.

Note: Examples are informal and formal surveys, direct and participatory observation, community interviews, focus groups, expert opinion, case studies, literature search.

Development Intervention

An instrument for partner (donor and non-donor) support aimed to promote development.

Note: Examples are policy advice, projects, and programmes.

Development objective

Intended impact contributing to physical, financial, institutional, social, environmental, or other benefits to a society, community, or group of people via one or more development interventions.

Economy

Absence of waste for a given output.
Note: An activity is economical when the costs of the scarce resources used approximate the minimum needed to achieve planned objectives.

Effect

Intended or unintended change due directly or indirectly to an intervention.

Related terms: results, outcome.

Effectiveness

The extent to which the development intervention's objectives were achieved, or are expected to be achieved, taking into account their relative importance.
Note: Also used as an aggregate measure of (or judgement about) the merit or worth of an activity, i.e. the extent to which an intervention has attained, or is expected to attain, its major relevant objectives efficiently in a sustainable fashion and with a positive institutional development impact.

Related term: efficacy.

Efficiency

A measure of how economically resources/inputs (funds, expertise, time etc.) are converted to results.

Evaluability

Extent to which an activity or a programme can be evaluated in a reliable and credible fashion.
Note: Evaluability assessment calls for the early review of a proposed activity in order to ascertain whether its objectives are adequately defined and its results verifiable.

Evaluation

The systematic and objective assessment of an on-going or completed project, programme or policy, its design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, development efficiency, effectiveness, impact and sustainability. An evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision-making process of both recipients and donors. Evaluation also refers to the process of determining the worth or significance of an activity, policy or programme. An assessment, as systematic and objective as possible, of a planned, on-going, or completed development intervention.

Note: Evaluation in some instances involves the definition of appropriate standards, the examination of performance against those standards, an assessment of actual and expected results and the identification of relevant lessons.

Related term: review.

Ex-ante evaluation

An evaluation that is performed before implementation of a development intervention.
Related terms: appraisal, quality at entry.

Ex-post evaluation

Evaluation of a development intervention after it has been completed.

Note: It may be undertaken directly after or long after completion. The intention is to identify the factors of success or failure, to assess the sustainability of results and impacts, and to draw conclusions that may inform other interventions.

External evaluation

The evaluation of a development intervention conducted by entities and/or individuals outside the donor and implementing organisations

Note: For some agencies the term 'evaluation' automatically assumes externality and independence. Other agencies reserve the term evaluation for evaluation studies which take place after the development intervention is completed

Feedback

The transmission of findings generated through the evaluation process to parties for whom it is relevant and useful so as to facilitate learning. This may involve the collection and dissemination of findings, conclusions, recommendations and lessons from experience.

Finding

A finding uses evidence from one or more evaluations to allow for a factual statement.

Formative evaluation

Evaluation intended to improve performance, most often conducted during the implementation phase of projects or programs.
Note: Formative evaluations may also be conducted for other reasons such as compliance, legal requirements or as part of a larger evaluation initiative.
Related term: process evaluation.

Goal

The higher-order objective to which a development intervention is intended to contribute.

Impacts

Positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended.

Independent evaluation

An evaluation carried out by entities and persons free of the control of those responsible for the design and implementation of the development intervention.

Note: The credibility of an evaluation depends in part on how independently it has been carried out. Independence implies freedom from political influence and organisational pressure. It is characterised by full access to information and by full autonomy in carrying out investigations and reporting findings.

Indicator

Quantitative or qualitative factor or variable that provides a simple and reliable means to measure achievement, to reflect the changes connected to an intervention, or to help assess the performance of a development actor.

Inputs

The financial, human, and material resources used for the development intervention.

Institutional Development Impact

The extent to which an intervention improves or weakens the ability of a country or region to make more efficient, equitable, and sustainable use of its human, financial, and natural resources, for example through: (a) better definition, stability, transparency, enforceability and predictability of institutional arrangements and/or (b) better alignment of the mission and capacity of an organisation with its mandate, which derives from these institutional arrangements. Such impacts can include intended and unintended effects of an action.

Internal evaluation

Evaluation of a development intervention conducted by a unit and/or individuals reporting to the management of the donor, partner, or implementing organisation.

Related term: self-evaluation.

Joint evaluation

An evaluation to which different donor agencies and/or partners participate.
Note: There are various degrees of 'depending on the extent to which individual partners cooperate in the evaluation process, merge their evaluation resources and combine their evaluation reporting. Joint evaluations can help overcome attribution problems in assessing the effectiveness of programmes and strategies, the complementarity of efforts supported by different partners, the quality of aid co-ordination etc.

Lessons learned

Generalisations based on evaluation experiences with projects, programs, or policies that abstract from the specific circumstances to broader situations. Frequently, lessons highlight strengths or weaknesses in preparation, design, and implementation that affect performance, outcome and impact.

Logical framework (Logframe)

Management tool used to improve the design of interventions, most often at the project level. It involves identifying strategic elements (inputs, outputs, outcomes, impact) and their causal relationships, indicators, and the assumptions or risks that may influence success and failure. It thus facilitates planning, execution and evaluation of a development intervention.

Related term: results based management.

Meta-evaluation

The term is used for evaluations designed to aggregate findings from a series of evaluations. It can also be used to denote the evaluation of an evaluation to judge its quality and/or assess the performance of the evaluators.

Mid-term evaluation or review

Evaluation performed towards the middle of the period of implementation of the intervention.

Related term: formative evaluation.

Monitoring

A continuing function that uses the systematic collection of data on specified indicators to provide the management team and the main stakeholders of an ongoing development intervention with indications of the extent of progress and achievement of objectives and progress in the use of allocated funds.

Related term: performance monitoring, indicator.

Outcome

The likely or achieved short-term and medium-term effects of an intervention's outputs.

Related terms: result, outputs, impacts, effect.

Outputs

The products, capital goods and services which result from a development intervention; may also include changes resulting from the intervention which are relevant to the achievement of outcomes.

Participatory evaluation

Evaluation method in which representatives of agencies and stakeholders, including beneficiaries, work together in designing, carrying out and interpreting an evaluation.

Partners/partnership

The individuals and/or organisations that collaborate to achieve mutually agreed objectives.
Note: The concept of partnership implies shared goals, common responsibility for outcomes, distinct accountabilities and reciprocal obligations. Partners may include governments, civil society, non-governmental organisations, universities, professional and business associations, multilateral organisations, private companies etc.

Performance

The degree to which a development intervention or a development partner operates according to specific criteria/standards/ guidelines or achieves results in accordance with stated goals or plans.

Performance indicator

A variable that allows the verification of changes in the development intervention or shows results relative to what was planned.
Related terms: performance monitoring, performance measurement.

Performance measurement

A system for assessing performance of development interventions against stated goals.
Related terms: performance monitoring, indicator.

Performance monitoring

A continuous process of collecting and analysing data to compare how well a project, programme, or policy is being implemented against expected results.

Process evaluation

An evaluation of the internal dynamics of implementing organisations, their policy instruments, their service delivery mechanisms, their management practices, and the linkages among these.

Programme evaluation

Evaluation of a set of interventions, marshalled to attain specific global, regional, country, or sector development objectives.
Note: a development programme is a time-bound intervention involving multiple activities that may cut across sectors, themes and/or geographic areas.
Related term: Country programme/strategy evaluation.

Project evaluation

Evaluation of an individual development intervention designed to achieve specific objectives within specified resources and implementation schedules, often within the framework of a broader programme.
Note: Cost benefit analysis is a major instrument of project evaluation for projects with measurable benefits. When benefits cannot be quantified, cost effectiveness is a suitable approach.

Purpose

The publicly stated objective of the development programme or project.

Quality Assurance

Quality assurance encompasses any activity that is concerned with assessing and improving the merit or the worth of a development intervention or its compliance with given standards.

Note: examples of quality assurance activities include appraisal, Results Based Management (RBM), reviews during implementation, evaluations etc. Quality assurance may also refer to the assessment of the quality of a portfolio and its development effectiveness.

Reach

The beneficiaries and other stakeholders of a development intervention.
Related term: beneficiaries.

Recommendations

Proposals aimed at enhancing the effectiveness, quality, or efficiency of a development intervention; at redesigning the objectives; and/or at the reallocation of resources. Recommendations should be linked to conclusions.

Relevance

The extent to which the objectives of a development intervention are consistent with beneficiaries' requirements, country needs, global priorities and partners' and donors' policies.

Note: Retrospectively, the question of relevance often becomes a question as to whether the objectives of an intervention or its design are still appropriate given changed circumstances.

Reliability

Consistency or dependability of data and evaluation judgements, with reference to the quality of the instruments, procedures and analyses used to collect and interpret evaluation data.

Note: evaluation information is reliable when repeated observations using similar instruments under similar conditions produce similar results.

Results

The output, outcome or impact intended or unintended, positive and/or negative of a development.
Related terms: outcome, effect, impacts.

Results Chain

The causal sequence for a development intervention that stipulates the necessary sequence to achieve desired objectives beginning with inputs, moving through activities and outputs, and culminating in outcomes, impacts, and feedback. In some agencies, reach is part of the results' chain.
Related terms: assumptions, results framework.

Review

An assessment of the performance of an intervention, periodically or on an ad hoc basis.

Note: Frequently 'evaluation' is used for a more comprehensive and/or more in depth assessment than 'review'. Reviews tend to emphasise operational aspects. Sometimes the terms 'review' and 'evaluation' are used as synonyms.

Related term: evaluation.

Results framework

The program logic that explains how the development objective is to be achieved, including causal relationships and underlying assumptions.

Related terms: results chain, logical framework.

Results-Based Management – RBM

A management strategy focusing on performance and achievement of outputs, outcomes and impacts.

Related term: logical framework.

Risk analysis

An analysis or an assessment of factors (called assumptions in the logframe) affect or are likely to affect the successful achievement of an intervention's objectives. A detailed examination of the potential unwanted and negative consequences to human life, health, property, or the environment posed by development interventions; a systematic process to provide information regarding such undesirable consequences; the process of quantification of the probabilities and expected impacts for identified risks.

Sector program evaluation

Evaluation of a cluster of development interventions in a sector within one country or across countries, all of which contribute to the achievement of a specific development goal.

Note: a sector includes development activities commonly grouped together for the purpose of public action such as health, education, agriculture, transport etc.

Self-evaluation

An evaluation by those who are entrusted with the design and delivery of a development intervention.

Stakeholders

Agencies, organisations, groups or individuals who have a direct or indirect interest in the development intervention or its evaluation.

Summative evaluation

A study conducted at the end of an intervention (or a phase of that intervention) to determine the extent to which anticipated outcomes were produced. Summative evaluation is intended to provide information about the worth of the program.

Related term: impact evaluation.

Sustainability

The continuation of benefits from a development intervention after major development assistance has been completed. The probability of continued long-term benefits. The resilience to risk of the net benefit flows over time.

Target group

The specific individuals or organisations for whose benefit the development intervention is undertaken.

Terms of reference

Written document presenting the purpose and scope of the evaluation, the methods to be used, the standard against which performance is to be assessed or how analyses are to be conducted, the resources and time allocated, and reporting requirements. Two other expressions sometimes used with the same meaning are 'scope of work' and 'evaluation mandate'.

Thematic evaluation

Evaluation of a selection of development interventions, all of which address a specific development priority that cuts across countries, regions, and sectors.

Triangulation

The use of three or more theories, sources or types of information, or types of analysis to verify and substantiate an assessment.

Note: by combining multiple data sources, methods, analyses or theories, evaluators seek to overcome the bias that comes from single informants, single methods, single observer or single theory studies.

Validity

The extent to which the data collection strategies and instruments measure what they purport to measure.

Annex 6: Additional International development terms

Aid Modalities

Different types of support to a country, a national programme and/or an element. Can range from non-earmarked balance-of-payment support to training, technical assistance, investments and commodity support, and sector budget support to operation and maintenance.

Evaluation Design

The logical model or conceptual framework and the methods used to collect information, analyse data and arrive at conclusions.

Immediate Objective

The effect which an activity is expected to achieve if completed successfully and on time.

Impact Assessment — an assessment of impact using before/after and/or with/without comparisons.

Impact Monitoring — the systematic, ongoing assessment of impact (intended and unintended) during project implementation.

Implementation — the process of realising a project 'on the ground' in line with the agreed work plan. It involves project management and monitoring - both financial and non-financial.

Institutional Appraisal — examines the organisation that will receive funds and support to look at the organisational structure, management, staffing, policies and procedures, in order to ensure that these are sound. It also looks at the way government policies impact on the organisation's operating environment. Institutional appraisal looks at whether the organisation's structure is adequate to carry out the project and whether policy or institutional changes outside the organisation are required if the project objectives are to be achieved. It also looks at whether local capabilities and expertise are being employed to maximum advantage. Institutional appraisal is only possible with an extensive understanding of the cultural, political and social operating environments of the given locality.

Mission

A particular task or goal assigned to a person or a group; a statement issued by an organisation setting out their key tasks and aims (see also: Vision).

Objectively Verifiable Indicators – OVIs

OVIs are the measures to verify accomplishment of all stages of the project through the Logical Framework. OVIs should be targeted in terms of quantity, quality, time.
Indicators at the Purpose level measure end of project Impact.

Ongoing Evaluation
An evaluation designed to help the development of a project as it proceeds. It is more wide-ranging and thorough and objective than an internal review. Sometimes referred to as 'concurrent' or 'formative' evaluation. Contrast ex post or terminal evaluations.

Output to Purpose Review —**OPRs** — the periodic review of ongoing projects, focusing on progress at the Purpose level, and whether the performance of the project implementation in terms of the achievements under each Output are sufficient to reach the Purpose at the end of the project. Additional outputs may be added as a result of the OPR. Other Impacts (e.g. at the Goal level and unplanned impacts) should also be considered. OPRs may be carried out by DFID staff and/or consultants. They usually involve the participation of stakeholders either directly or indirectly by assessing their views. OPRs require information on which to base judgements. The collection of that information requires a synthesis of monitoring data and related studies, as well as the commissioning of specific evaluation and impact studies. Building in the necessary information systems to deliver monitoring data will be a part of project design. The likelihood of success is scored at the Output and Purpose level on a 5-point scale.

Ownership

Appropriation or taking of responsibility for a certain endeavour. Ownership implies formal or real authority as well as effective self-authorisation to assume management responsibility. (Danida)
Claiming as one's own; possession; the idea that an organisation is more likely to be committed to a project if it has been involved in its development. (DFID)

Process Indicator
A measure or sign of whether planned activities are being carried out, and how they are being carried out.

Project Appraisal — a comprehensive and systematic review of all aspects of the project — technical, financial, economic, social, institutional, environmental — to determine whether an investment should go ahead.

Project Cycle
A useful tool for understanding the various stages that any project will probably go through. The same logic applies to simple project ideas within one organisation and complex projects supported by a number of external funders.

The basic model includes the following stages:
(a) Project Identification — ideas for potential projects are identified and explored.
(b) Project Preparation — the project idea is carefully developed.
(c) Project Appraisal — the project is rigorously assessed.
(d) Project Implementation — the project is progressed in the agreed manner.

The more detailed version of the cycle, used by many funding agencies, emphasises decision-making with feedback loops and includes aspects of negotiation, effectiveness, supervision and completion:

(e) Negotiations and Presentation — at the stage where a project proposal has been developed, the parties involved (usually the recipient/borrower and the donor/funder) negotiate the detailed requirements for implementation.

f) Monitoring — the project will be monitored throughout the implementation period in order to make sure that everything is going according to plan and that corrective action can be taken where necessary.

(g) Evaluation — the project is assessed against its objectives in terms of performance, efficiency and impact.

Project Memorandum

The paper which records all details of a project.

Project Monitoring

Measuring how the project is being implemented and how it is operating — essential for evaluation and measuring performance. Reporting on costs, savings, other benefits, reporting to decision-makers, may support the generation of new projects. Continuous tracking of a project's development through the collection of specified variables. Progress is monitored according to the planned methodology and time-scale. Monitoring covers both financial and non-financial aspects of the project — the use of resources, measurement of performance and outputs (see also: Project Cycle).

Programme

A group of related projects or services directed toward the attainment of specific (usually similar or related) objectives.

Sector

A coherent set of activities which can be relevantly distinguished in terms of policies, institutions and finances, and which needs to be looked at together to make a meaningful assessment.

Sector Programme Support (SPS)

A longer-term framework for a broader assistance to a national endeavour (called a National Sector Framework), including various and flexible modalities of support at different levels within agreed objectives and management procedures.

Strategic Framework — the overall aims and objectives of a country's approach to development based on analysis of problems, and including a statement of priorities. This sets the context for programme and project development (see also: Strategy).

Sustainability Factors

Six areas of particular importance to ensure that development interventions are sustainable, i.e. institutional, financial and economic, technological, environmental, socio-cultural and political.

Annex 7: Standard research terms used in monitoring and evaluation

Baseline — information collected before or at the start of a project which provides a basis for planning and/or assessing subsequent progress and impact.

Information will be required on a control site or comparable group outside the project if with/without comparisons, not just before/after comparisons, are to be possible subsequently.

Bias

Repeated, systematic error. Bias in sampling, for example, means ignoring or under-representing parts of the target population.

Case Study

Study based on a small number of 'typical' examples. Results are not statistically reliable.

Correlation

The strength of relationship between two (or more) variables. Positive correlation means that one variable tends to increase together with another variable. Negative correlation means that one variable decreases as the other one increases.

Data

Information. The term "data" often describes information stored in numerical form. Hard data is precise, numerical information. Soft data is less precise, verbal information. Raw data is the name given to survey information before it has been processed and analysed. Missing data are values or responses which fieldworkers were unable to collect (or which were lost before analyses). To promote gender balanced analyses, gender segregated data are generally preferred.

Distribution

The full set of sample or population values. A frequency distribution is a table that shows how many times each particular value or item occurs in the sample data.

Error

Variation between sample data and the actual situation. Error can be systematic (see bias), or random. Random error is a result of chance differences from sample to sample. Systematic error (bias) occurs where there is a fault in the way the sample is chosen.

Experiment

Carefully designed trial or test, controlling the effect of outside factors. Used in agriculture, for example, to test alternative seed or crop types in controlled conditions.

Formal Methods

Structured research methods, usually using structured interview forms with large samples (see also 'Informal methods').

Informal Methods

Less structured, exploratory research methods, such as case studies or participant-observation (see also 'Formal methods').

Key Informant

Person carefully chosen for interview because of their special knowledge of some aspect of the target population.

Mean

A type of average calculation for interval variables. Total of the sample values divided by the number of values in the sample.

Median

A type of average calculation for interval variables. The middle value when the data are arranged in order of size. Where the set of data has an even number of values the median is the mean of the two middle values.

Objective Data

Data based on observable facts that do not involve personal feelings. Objective data can be quantitatively or qualitatively measured.

Participatory Research

Research undertaken in partnership with end-users who, in general, are involved in the definition of the research agenda, the conduct of the research, the evaluation of results and the dissemination of findings.

Primary Data

Data collected by an evaluation team specifically for the evaluation study

Qualitative Data

Observations that are categorical rather than numerical and often involve attitudes, perceptions and intentions.

Qualitative Indicator

An indicator drawing on perceptions and judgements.

Response Rate

Proportion (or fraction) of the sample who are successfully interviewed.

Secondary Data

Data collected and recorded by another (usually earlier) person or organisation, usually for different purposes than the current evaluation.

Subjective Data

Observations that involve personal feelings, attitudes and perceptions. Subjective data can be quantitatively or qualitatively measured.

Surveys

A data collection method that involves a planned effort to collect needed data from a sample (or a complete census) of the relevant population. The relevant population consists of people or entities affected by the programme (or of similar people or entities).

Variable

A single characteristic of the target population, such as household size, household income or spending priorities.

Annex 8: Approaches and tools for commissioners of complex evaluations

The most common approach in the evaluation of development assistance is the deductive approach. Those commissioning the study, or the evaluators themselves normally formulate hypotheses at an early stage in the evaluation design. These hypotheses are then tested against reality when the evaluation is carried out.

The advantage of this approach is that it takes development experience as its point of departure. This makes it easier to choose indicators and interpret data. The main criticism of this approach is that it may limit the themes and problem areas examined so that important aspects or conditions may be overlooked.

One solution to this problem is to use an inductive approach where the evaluation team starts with an open mind. Through observation and investigation, themes arise that in turn demand new knowledge, gradually leading to new insights. The advantage of this approach is that it is often the only means available for achieving an in-depth understanding. The criticism is that it is time consuming. To avoid subjective biases, it may be necessary to place special emphasis on the selection of particular data, perspectives and methods and then to explain these in detail.

APPROACH	PROCESS	ADVANTAGE	DISADVANTAGE
DEDUCTIVE	Hypothesis formulated and tested against reality	Starting point based on existing experience, which eases the selection and analysis of data	Limits the themes and problem areas examined
INDUCTIVE	Evaluators start with an open/questioning mind	In-depth understanding as themes arise leading to new insights	Time consuming; risk of subjective bias

Clearly, most evaluations will rely on a combination of these two approaches. The part of the evaluation that focuses on efficiency and effectiveness may benefit strongly from a deductive approach. Here the perspective is narrow. It is the implementers' and users' perspective. In measuring impact, relevance and sustainability, a more inductive approach may be more useful because the field of study is more open-ended and the perspective is the widest since it is the society's perspective.

Evaluating the strategy behind a development intervention

The development intervention strategy or theory of change is the starting point for the evaluator, since this identifies the approach and anticipated achievements against which assessments will be made. Ex-post evaluations are primarily concerned about the broad overall assessment of development interventions, and less with the implementation of activities as such. In other words, the significance of the process, rather than simply the process itself.

The intervention strategy often changes during implementation. The level of ambition may be increased or decreased in the course of events, as well as the direction of activities. The evaluator must review changes in the strategy in order to gain a full picture of what is to be evaluated.

It is a key concern of the evaluation to assess the extent to which the initial objectives are still relevant, whether a possible shift in the objectives was justified and whether other changes in the objectives are warranted.

Complex programmes may have many different objectives. In such cases, it may be useful to organise the objectives in a logical cause/effect sequence. The development assistance of most donor agencies is planned according to the Logical Framework (Logframe) method, where the formally agreed objectives are established and sequenced together with

accompanying indicators. If changes have been made in these objectives, the reasons for the changes should be analysed specifically.

This organisation of all the objectives helps to provide an initial overview of the complexity of the strategy. In many cases assembling this overview will reveal a picture of confusing, even conflicting, objectives. It will often demonstrate that the objectives have gradually changed from what was initially intended, and that this has been done without being discussed explicitly or approved between the parties involved.

Is the strategy realistic?

The analysis should consider whether the objectives are realistic. Often objectives are too ambitious when seen in relation to the resources available and the agreed time-frame of the aid intervention. Unrealistic objectives are a problem for two reasons. Firstly, if they are used as the formal reference for the evaluation the final assessment will be that the aid activities have failed. Secondly, the analysis necessary to assess the fulfilment of ambitious objectives would be unduly expensive. The evaluation is unlikely to be able to demonstrate that the changes can be attributed to the aid intervention, separating them from other factors that might have caused major changes.

Evaluation design

In academic research, the design to study the effects of an intervention involves establishing the situation in both a target group and control group, before and after the intervention. This is a precise procedure, but costly and complicated.

Evaluation cannot replicate pure research in this sense. There is an ethical dilemma involved in making use of a control group if this means that its members will not benefit from the development intervention. Time and resources for most evaluation assignments are usually severely restricted, but the evaluation must satisfy a minimum standard of quality of information in terms of validity and reliability.

	TYPE OF DESIGN	COMPLEXITY	COSTS
STUDY DESIGN	"Before and after" Study with control group	Complex	Costly
	"Before and after" Study without control group	Simple	Less costly
	"After" Study with control group	Simple	Less costly
	"After" Study without control group	Very simple	Low cost

The most commonly used design in evaluation is the simple 'after'-study without a control group. The evaluator must make judgements on the basis of general experience and comparing effects against data from any baseline study conducted prior to the intervention. To strengthen the reliability of these judgements, other supporting information must be made available to verify the findings. These methods are brought together in the evaluation design.

The evaluation cannot answer every question that the various stakeholders want answered without becoming too time-consuming or expensive to commission. The process of selecting key questions is therefore a very important and often delicate process. Limiting

the evaluation to about three key questions is a good discipline. Similarly in a large scale or high level evaluation, the evaluation team cannot look in detail at every level or visit every country or location where the development intervention is operating. The stakeholders need to be involved in the selection of case study areas or countries and the Evaluation Manager will need to help to defend the Evaluation Team from unrealistic demands for visits.

Thinking about the use of a website and communication strategy is an important part of the evaluation design process. The website can encourage widespread information sharing and consultation processes as the evaluation study progresses. The communications strategy can also build in special conferences and feedback sessions as part of the process, or encourage the use of a newsletter to report interim findings and maintain the interest of potential users in the final recommendations.

The analytical process

The process that takes place during an evaluation design is depicted below. In the preparatory phase, the first step is to translate very general evaluation criteria into specific evaluation questions that relate to the development intervention under study. There will always be several evaluation questions under each evaluation criterion, in order to substantiate conclusions and provide enough information for decision-making.

The next step is to make the evaluation questions operational by turning them into evaluation instruments. In this context the evaluation instrument means the evaluation question plus the information collection method applied. The evaluator must ensure that the evaluation instruments produce valid and reliable information.

Prior to, and during, fieldwork the instruments are used to collect data and generate findings. Much of the information collected is in the form of raw data, time-series, individual statements etc. Most of this information is at a level that is too detailed to be useful for decision-makers. Each piece of information provides only a small part of the combined picture. Many pieces of information may be conflicting.

The information therefore has to be analysed and aggregated to a higher, less detailed, level. This should result in a number of single statements providing answers to the various evaluation questions. These statements are the conclusions. Taken together, the conclusions should be comprehensive enough to accurately depict the situation.

The Analytical Process of Evaluation

Criteria → Questions → Instruments (Questions & the information collation method) → Conclusions (Answers to evaluation questions) → Recommendations
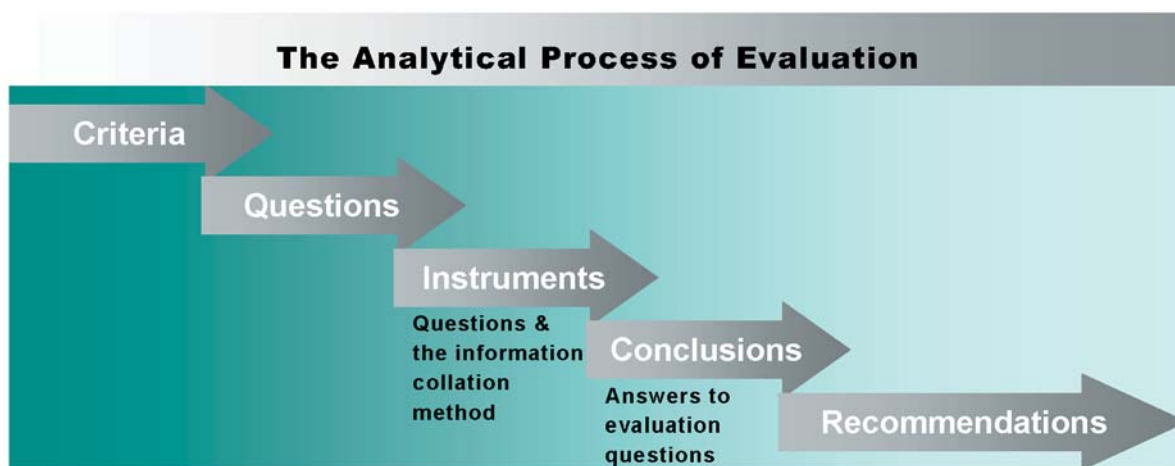
The strength of the conclusions will rely on the amount and quality of underlying information. In some cases the analysis of information will reveal conclusions that are linked to other evaluation questions, or other evaluation criteria, than the one in focus. Such conclusions will add to the picture and can provide opportunities for verifying or rejecting other conclusions.

The analytical part of the evaluation is the one that transforms the detailed pieces of information into conclusions at a more aggregated level, and finally makes use of these to draw the overall conclusions at the highest level, linking to the evaluation criteria.

The purpose of analysis is to transform data into credible evidence about the development intervention and its performance. Individual pieces of information are of little use unless they can be extracted or aggregated to address the questions in focus. This assumes that the evaluation design and instruments are appropriate, that the data are valid and reliable, and are analysed properly.

The evaluator is then expected to draw the conclusions and translate the conclusions into main recommendations, which can be considered by management and subsequently translated into more specific and detailed decisions. The evaluator is not usually expected to provide recommendations at a detailed level. This should be the result of more comprehensive decision processes, which will involve the organisation that has commissioned the evaluation and the different stakeholders. Audiences for the findings and recommendations of an evaluation should be identified early on in the design process

Collecting information

Evaluations often produce controversial results. Therefore, the evaluation is often criticised, and the design and data collection methods may be scrutinised closely. The basic distinction is between the formal methods typical of academic research and the simple, less formal methods of evaluation which asks primarily what happened and why.

Characteristic of formal methods is that the procedure is clearly defined from the outset. The result will normally be qualitative or quantitative information with a high degree of both validity and reliability. Examples are the formal survey, participatory observation and direct measurement. The main problem is that they are expensive. Preferably they will be used in connection with pre-studies, more seldom as a consistent part of the evaluation itself.

Informal methods, on the other hand, have less precise procedures and rely to a larger extent on experience, intuition and subjective judgement. Examples are field visits and unstructured interviews.

Such methods are not suitable for generating systematic information. Personal opinion and biases often affect validity and reliability. Still, most decisions are based on information gathered by means of quite informal approaches. This is also the case in evaluations.

In evaluation of development assistance, the most frequently used data collection method is interviews with key informants. This will inevitably produce subjectively biased information that will have to be verified. The evaluator therefore needs to use supplementary methods that maintain a reasonable level of precision without being too time-consuming and expensive. The concept of triangulation is used to bring together data on the same topic from several different sources.

> ### Triangulation
>
> This means compensating the use of simple data collection methods and a simple study design with the use of several information sources and different methods simultaneously, to generate information about the same topics.
>
> For instance, information from a simple survey may be supplemented with general experience data from similar interventions, and interviews with a variety of key informants to provide contextual information. In a situation that affects several parties with different interests, representatives of all parties, as well as some neutral respondents, should be interviewed. This provides a triangulation effect that largely helps to verify information, cuts through conflicting evidence, and reveals insights, in a cost-effective way.

The different methods of collecting data that may be used to triangulate information are shown in the following table:

DATA COLLECTION METHODS

Advantage	Type	Disadvantage
Economic and efficient way of obtaining information.	Literature Search	Difficult to assess validity and reliability from secondary data
Registration of quantifiable or classifiable data by means of an analytical instrument. Precise, reliable and often requiring few resources.	Direct measurement	Difficult to assess validity and reliability from secondary data
Involves inspection, field visits, observation to understand processes, infrastructure/services and their utilisation.	Direct observation	Dependent on the observer's understanding and interpretation.
Involves quantitative surveys of small samples. Reasonable and rapid.	Informal survey	Risk of sampling errors/biases. Less suited for generalisation.
In-depth review of one or a small number of selected cases. Well-suited for understanding processes and for formulating hypotheses to be tested later.	Case studies	Choice of cases may affect the whole synthesis
In-depth observations over an extended period of time, participatory or non-participatory. Well-suited for understanding processes.	Observation	Limited potential for generalisation
Oral interviews or written questionnaires in a representative sample of respondents. Data collection is demanding but often produces reliable information.	Formal survey	Open questions hard to collate but offer interesting insights. Can canvas widespread views by email questionnaires

Participatory appraisal techniques

Advantage	Type	Disadvantage
Flexible, in-depth approach. Easy to implement.	Key informant interviews	Risk of biased presentation/interpretation from informants/interviewer
For analysis of specific, complex problems, to identify attitudes and priorities in smaller groups. Reasonable in terms of cost, and efficient. Stimulates the generation of new ideas.	Focus group interviews	Risk of one-sidedness on the part of participants and the moderator. Individuals may dominate the group
Low-cost, efficient. Direct contact with those affected, especially useful if separated by gender or special interest.	Group interviews	Susceptible to manipulation and less suitable for sensitive issues, unless groups are separated by gender or special interests
Same questions make comparing results easier	Structured interviews	Inflexible
More questions can arise according to initial responses	Semi-structured interviews	Highly flexible - richer and more varied insights

Quantitative and qualitative analysis

The level of sophistication of analysis is a matter of concern in evaluation. Tables, percentages and averages often give a clear picture of the sample data particularly for non-

specialists, and many users will only be interested in this level of analysis. In addition, measures of spread, including percentiles and standard deviations, may add valuable information on how a variable is distributed throughout a sample population. There are a wealth of more sophisticated research methods that can be applied, according to the subject of the evaluation. However, much of the evaluation work can be done using very basic methods. After all, evaluation is not scientific research and the quality of data and the size of samples are often such that sophisticated methods are not appropriate.

Conclusions on the qualitative versus quantitative debate

A common distinction is the one between quantitative and qualitative analysis. Different parties will often have different opinions of what type of analysis should be preferred. Most parties, however, will agree that both types of analysis should be employed simultaneously in varying proportions.

Quantitative analysis involves systematic manipulation of data, either 'hard' objective data or 'softer' subjective, attitudinal data. It has three main purposes:

- To describe phenomena in a concise format using statistical tabulation
- To test relationships among variables of interest
- To generalise findings to an overall population

One of the main advantages of statistical analysis is that it can be used to summarise the findings of an evaluation in a clear, precise and reliable way. However, not all information can be analysed quantitatively. For example, responses to an open-ended interview survey may provide lengthy descriptions that may be very difficult to categorise, let alone quantify, without losing subtle differences in the responses.

Another problem with quantitative methods is that their validity depends on initial assumptions about the data being used. Sophisticated analysis requires that the data have a certain quality. With available statistical software packages, the danger is that the analysis becomes more sophisticated than the data. This, of course, weakens the credibility of the conclusions. It is essential to be aware of the assumptions as well as the limitations of the statistical technique employed.

The evaluator therefore needs to make sure that the choice of methods for quantitative analysis matches the quality of data. In many evaluations this implies that simple methods should be preferred. This has the added advantage that the results will be easier to communicate to non-specialists.

- Qualitative data includes detailed descriptions, direct quotations in response to open-ended questions, the transcript of opinion of groups, and observations of different types.
- Qualitative analysis typically includes content analysis, analysis of case studies, inductive analysis and logical and sequential analysis. All methods may produce descriptions (patterns, themes, tendencies, trends, etc.), and interpretations and explanations of these patterns.

The analysis of qualitative data helps broaden the view of the phenomena of interest in an evaluation, but can also increase depth and detail, where needed. The process of analysing qualitative information is often inductive and without any particular guiding theory.
Qualitative analysis relies on the evaluator's professional judgement concerning the relevance and significance of available data. The evaluator must therefore be both

knowledgeable concerning the evaluation issues and attuned to potential biases.

Qualitative analysis is best done in conjunction with the statistical analysis of related (quantitative or qualitative) data.

The evaluation should be designed so that the two sorts of analysis, using different but related data, will be mutually reinforcing. It is often the case that quantitative analysis raises questions that trigger further qualitative analysis and vice versa.
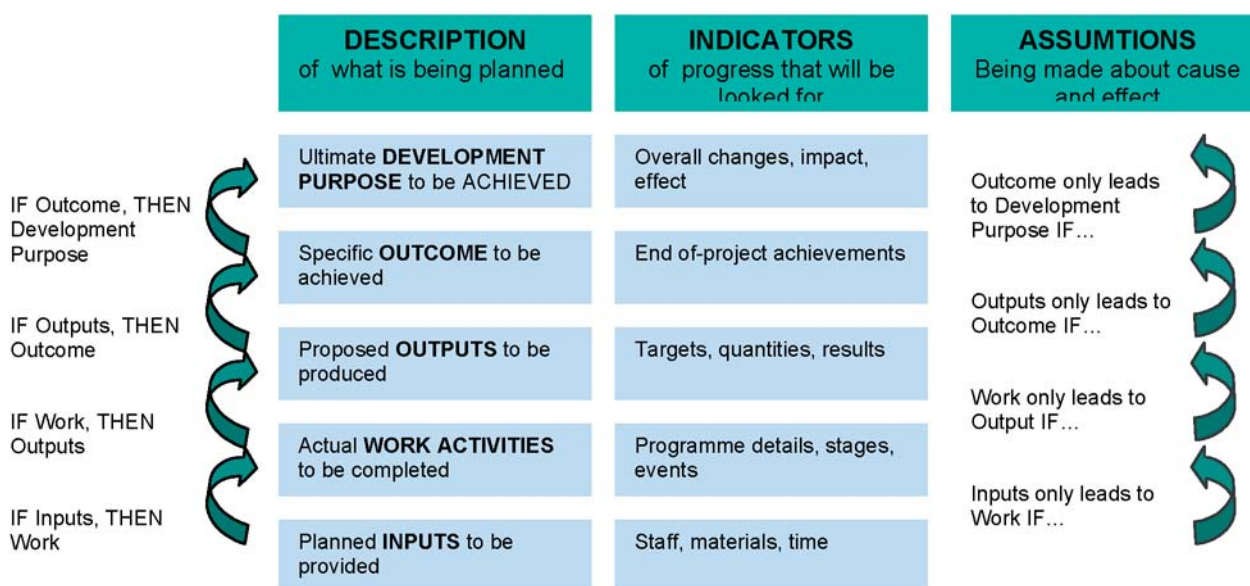
Other useful tools

Logical frameworks: the logical framework or logframe is the most common planning, monitoring and evaluation tool used to describe the anticipated chain of cause and effect in development interventions.

The logframe is developed by development partners at the design phase, and used by project implementers to guide their work.

Evaluators use the framework to understand the initial logic and assumptions behind the project, and to assess the success of planned achievements against the success indicators. The outputs and indicators may be refined over the course of implementation in agreement with project partners. All such changes should be documented.

The Logical Framework
(Describing the logic and assumptions behind
the development intervention)

DESCRIPTION of what is being planned	INDICATORS of progress that will be looked for	ASSUMPTIONS Being made about cause and effect
Ultimate **DEVELOPMENT PURPOSE** to be ACHIEVED	Overall changes, impact, effect	Outcome only leads to Development Purpose IF…
Specific **OUTCOME** to be achieved	End of-project achievements	Outputs only leads to Outcome IF…
Proposed **OUTPUTS** to be produced	Targets, quantities, results	Work only leads to Output IF…
Actual **WORK ACTIVITIES** to be completed	Programme details, stages, events	Inputs only leads to Work IF…
Planned **INPUTS** to be provided	Staff, materials, time	

IF Outcome, THEN Development Purpose

IF Outputs, THEN Outcome

IF Work, THEN Outputs

IF Inputs, THEN Work

Timelines and critical incident reviews: visual tools such as timelines and critical incident reviews have been developed by the Performance Assessment Resource Centre as overarching analytical tools, to enable evaluators and stakeholders to grasp the multi-level and multi-dimensional changes that take place in large programmes over time.

The timeline and critical incident overview
The critical incident overview places the logical framework in a much more vivid context. It makes the often invisible actors, whose actions affect the outcomes of the logframe, much more accessible. It connects the logframe with a time-scale. Often logframe inputs and outputs are not related to time, other than by the assumption that they will be achieved 'by the end of the project'.

The timeline and critical incident overview show changes affecting the performance of the programme at national, state, international levels, and in the funding agency and programme itself.

The logframe alone extracts the project or programme into an abstract world where predicted results, based on linear logic, risks and assumptions represent the products of the design team's imagination during the design phase. The critical incident overview grounds the project in real time and real experience. In this way it brings the historical context, in which the project design is operating, into a much sharper focus.

The combination of reviewing multi-level critical incidents affecting the programme along a timeline is a simple way of showing inter-related events over time, in a graphical way. The critical incident timeline is particularly useful for showing key change moments such as national elections, changes of leadership or organisational policy that may have affected the direction of long-term large-scale programmes and which may have impinged on the development effectiveness and organisational effectiveness of the overall programme.

The critical incident overview should be produced by the project or programme managers in preparation for a review or evaluation. It will help the review or evaluation team to focus on key areas of difficulty and enable them to elicit further relevant information from the programme team or other stakeholders.

As a preliminary tool, this overview provides a graphical introduction to the project history that may be easier for the evaluators to grasp than a voluminous series of project documents. References in the overview to significant changes in objectives may need to be verified through an audit trail of related project documents.

Source:
http://www.parcinfo.org/documents/e-learningDownloads/PARC_Web_Document_02.doc

The Document Audit trail

In project evaluations, it is often easier to track the disbursement of money than the development of ideas over the project cycle. Project managers should be encouraged to maintain a documentary audit trail as part of the internal monitoring system. The documentary audit trail is basically a filing system of key project documents including the original project memorandum, the records of monitoring meetings and reviews, key changes of policy and agreed decisions affecting the direction of the project.

AUDIT TRAIL: DOCUMENTS REQUIRED	■ Project Memoranda including any initial logframe ■ Project tendering documents and any subsequently agreed changes to the design ■ Project monitoring reports ■ Review reports ■ Documents which mark critical incidents that have affected the project.

The document audit trail should provide accompanying documents to explain the critical incidents recorded in the timeline or critical incident review shown above. The documents may provide evidence of policy changes made or new agreements negotiated by new partner representatives, e.g. after elections. Documents may include key emails as well as more formal letters, reports or press cuttings etc.

It is not uncommon for the evaluators to find no one able to answer questions about what happened and why during the earlier stages of the programme being evaluated. Evaluators are then left with the dilemma of whether they should be evaluating effectiveness against the original intentions laid out in at the design stage, or the changed intentions that appear to have emerged over time after exposure to the real context in which the development intervention has been operating. Where the changes and reasons for change have been properly documented the evaluator's task is easier. What appear to be major changes of direction are sometimes the result of shifting national or donor policies. At other times they are simply the aggregated result of many small incremental changes none of which seemed significant individually at the time.

The tools described here can help evaluators to reveal the critical interplay of both the drivers and inhibitors of change, and therefore to explain better what has happened and why.

Source:
http://www.parcinfo.org/documents/e-learningDownloads/PARC_Web_Document_02.doc

Humanitarian assistance

The DAC provides 'Guidance for evaluating humanitarian assistance in complex emergencies'. This document describes the differences between the evaluation of humanitarian assistance and the evaluation of conventional development assistance. Evaluation techniques like discourse analysis are explained in detail. Apart from the standard DAC criteria, coverage, coherence, co-ordination and protection are listed as sub-criteria (see Annex 2).

Source: DAC, Guidance for evaluating humanitarian assistance in complex emergencies (Paris, 1999)

Evaluation of humanitarian assistance demands specific methodological approaches, because of the speed and turbulence of these interventions. Evaluating emergency operations in "real-time has recently gained ground at UNHCR, and in the wider humanitarian world. A real-time evaluation (RTE) is a timely, rapid and interactive peer review of a fast evolving humanitarian operation. With knowledge of lessons from past emergency evaluations, a real-time evaluator is able to approach an emergency from a number of different angles and to view the situation from a number of vantage points — headquarters, country of asylum (capital and field) and possibly regional centres and the country of origin. UNHCR provides a short paper to answer some frequently asked questions regarding these kinds of evaluations, based on UNHCR's experience in this area to date.

Source:
http://www.unhcr.ch/cgi-bin/texis/vtx/home/
+OwwBmHejDdpwwwnwwwwwwwmFqo20I0E2glhFqo20I0E2gltFqoGn5nwGqrAFqo20I0E
2glcFqmGBnDzmxwwwwwww/opendoc.pdf

Gender

DFID is looking to strengthen the mainstreaming of gender considerations within evaluation processes. The following box provides a list of issues that DFID staff should take into account when commissioning evaluations:

Gender in evaluation: Checklist

- Does the composition and skills of the evaluation team ensure adequate capacity to address gender issues in the evaluation?
- Are gender issues clearly assigned within the evaluation team?
- Do indicators and data collection methods reflect the potential differences between men and women?
- Have gender issues been addressed at the design and implementation phase of an intervention?
- What are the intended and unintended effects of an activity on men and women?
- Are there clear findings, conclusions and recommendations on gender issues in the evaluation report?

Although gender issues are supposed to be mainstreamed in international development, experience shows that this is still not the case. For evaluations focusing specifically on gender issues, a useful resource is the International Labour Organisation's "Guidelines for the Integration of Gender Issues into the Design, Monitoring and Evaluation of ILO Programmes and Projects".

Source: http://www.ilo.org/public/english/bureau/program/eval/guides/gender/index.htm

Participatory evaluation

An evaluation is characterised not only by its focus, but also by its process. Who were the findings and recommendations for and how were they arrived at?

Participatory evaluation has the potential to:

- Build partnerships and sense of local ownership over donor funded interventions
- Build consensus among project or programme staff and partners about goals/objectives of the intervention
- Enhance local learning, management capacity and skills
- Provide timely, reliable, and valid information for management decision-making
- Increase cost-effectiveness of M&E information
- Empower local people to make their own decisions about the future, to initiate, control and take corrective action

Source: Tools for Development DFID (2002 amended)

Participation can be seen as an end in itself, as the right of people to have a voice in matters that significantly affect their interests. It can also be justified in more instrumental terms. Participation in this case helps mobilise local knowledge. While in a conventional evaluation the evaluators are hired to assess the merits of the intervention, their role is different in a participatory evaluation. In that situation they become mainly facilitators and instructors, helping others to make the assessment.

Sources: CIDA evaluation guide, appendices B
http://www.acdi-cida.gc.ca/INET/IMAGES.NSF/vLUImages/
Performancereview5/$file/Evaluation%20Guide.pdf

Participatory rural appraisal
http://www.gtz.de/participation/download/schoenhuth-kievelitz1995.pdf

Tools for Development DFID
http://www.dfid.gov.uk

Annex 9: List of useful web based resources

DAC, DAC member states and international organisations with online evaluation guidance

- DFID Evaluation Department
  http://insight/evd
- DAC
  http://www.oecd.org/dac/evaluationnetwork
- CIDA
  http://www.acdi-cida.gc.ca/cida_ind.nsf/
  49d9f10330ed2bb48525677e00735812/061a4e025fc6fe2785256c6b00155e2d
  ?OpenDocument
- Danida
  http://www.um.dk/da/menu/Udviklingspolitik/MaalOgResultatstyring/Evaluering/
  Guidelines/EvaluationGuidelines
- EC
  http://europa.eu.int/comm/europeaid/evaluation/methods/guidelines_en.pdf
- DIDC Finland
  http://global.finland.fi/julkaisut/yleis/pdme/index.html
- IFAD
  http://www.ifad.org/gbdocs/eb/ec/e/34/EC-2003-34-WP-3.pdf
- JICA
  http://www.mofa.go.jp/policy/oda/evaluation/guideline.pdf
- SIDA
  http://www.parcinfo.org/e-learningDownloads.asp

Resource centres

- **ALNAP** (The Active Learning Network on Accountability and Performance in Humanitarian Assistance)

The ALNAP Evaluative Reports Database offers access to the evaluative reports of humanitarian action in order to ensure experiences and learning are shared, and to foster discussion on good practices.

- **Centre for Aid and Public Expenditure**

The Centre for Aid and Public Expenditure was established in 1999 and has since then carried out high quality policy-focused research on aid instruments, budget processes and the reform of public expenditure systems.

- **Eldis**

ELDIS is a gateway to information on development issues, providing free and easy access to wide range of high quality online resources.

- **IDRC (The International Development Research Centre)**

The Evaluation Unit promotes the use of state-of-the-art evaluation and management methodologies and practices throughout IDRC and with its Southern initiatives.

- **MandE**

A news service focusing on developments in monitoring and evaluation methods relevant to development projects and programmes with social development objectives.

● **PARC** (Performance Assessment Resource Centre)

The PARC supports international efforts to improve performance assessment in international development practice. The PARC has been set up to support the evaluation work of development partners in both the South and the North.

● **Poverty Action Lab**

The Poverty Action Lab serves as a focal point for development and poverty research based on randomised trials. The objective is to improve the effectiveness of poverty programmes by providing policy makers with clear scientific results that help shape successful policies to combat poverty

For further information on the topics below, consult the following web sites:

Terms of Reference

● How to write ToR: CIDA Checklist: Chapter 5, p. 25
➔ http://www.acdi-cida.gc.ca/INET/IMAGES.NSF/vLUImages/
    Performanceereview5$fil/Evaluation%20Guide.pdf

● Sida format for Terms of Reference
➔ http://www.sida.se/content/1/c6/02/56/24/SIDA3753en_mini.pdf — p.94: Annex A:

● Example of good Terms of Reference from CIDA: Annex D
➔ http://www.acdi-cida.gc.ca/INET/IMAGES.NSF/vLUImages/
    Performancereview5/$file/Evaluation%20Guide.pdf

Evaluation Team Roles and Skills

● Sida evaluation manual. Chapter 2.7; Recruiting consultants, p. 76
➔ http://www.sida.se/content/1/c6/02/56/24/SIDA3753en_mini.pdf

Reporting

● DIDC Finland
➔ http://global.finland.fi/julkaisut/yleis/pdme'
● Sida reporting and Dissemination, Chapter 4,p.83
➔ http://www.sida.se/content/1/c6/02/56/24/SIDA3753en_mini.pdf

● CIDA, Preparing evaluation reports; Chapter 9, p. 63
➔ http://www.acdi-cida.gc.ca/INET/IMAGES.NSF/vLUImages/
    Performancereview5/$file/Evaluation%20Guide.pdf

● EC, Format and presentation of evaluation reports, Part 3, p. 16
➔ http://europa.eu.int/comm/europeaid/evaluation/methods/guidelines_en.pdf

Using Evaluations and Lesson Learning

● CIDA, Optimising benefits, Chapter 10, p. 77
➔ http://www.acdi-cida.gc.ca/INET/IMAGES.NSF/
    vLUImages/Performancereview5/$file/Evaluation%20Guide.pdf

● Finland EC, Format and presentation of evaluation reports, Part 3, p. 16
➔ http://global.finland.fi/julkaisut/yleis/pdme/evaluation.htm#_Toc393865649

● Sida, Dissemination of evaluation results, Chapter 4.3, p. 88
➔ http://www.sida.se/content/1/c6/02/56/24/SIDA3753en_mini.pdf

Ten Principles to Ensure Successful Use of Evaluations

● Swiss Agency for Development and Co-operation
➔ http://www.deza.ch/index.php?navID=3523&userhash=27983680&l=e

Please note that these sources reflect their respective organisational settings as a basis for reporting but do include useful elements to apply more widely.