Ministry of Foreign Affairs of Denmark
Danida

evaluationguidelines

EVALUATION GUIDELINES

The Guidelines can be obtained free of charge by ordering on www.evaluation.dk
or from:

DBK Logistik Service
Unionsvej 4
4600 Koege
Denmark
Phone: +45 32 69 77 88

The Guidelines can also be downloaded through the homepage of the Ministry of
Foreign Affairs www.um.dk or directly from the Evaluation Department:
www.evaluation.dk.

Contact: eval@um.dk

PREFACE

Evaluation has been used systematically since the early days of Danish development assistance. In 1982 Danida established a special unit responsible for evaluation. The use of evaluations has developed successively in three main stages:

Prior to 1982 evaluations focussed essentially on individual projects and programmes. Most of these were mid-term or phase evaluations conducted as the project moved from one phase to the next. Only a few end-evaluations were conducted, and only occasionally were ex-post evaluations carried out to study the long-term effect of projects.

In the period 1982-87, after Danida's Evaluation Unit was established, it was agreed to use evaluations more systematically to improve the quality of Danida financed development activities. Also in this period most evaluations were mid-term or phase evaluations of individual projects. The tendency was to replace mid-term evaluations with internal reviews and increase the number of end-evaluations. The use of evaluations was more systematic in the sense that it was guided by an annual evaluation programme to ensure that the sample of evaluated projects and programmes were representative for Danish bilateral assistance.

In the period 1987-97 the number of individual project evaluations was reduced and the number of thematic and sector evaluation increased. As a policy, all evaluation reports were made public. In 1992, informing the public was included as a distinct objective for evaluation in accordance with DAC principles. In this period evaluations became more experimental and included a small number of impact evaluations as well as use of participatory methods. In general, all evaluations were – and still are – conducted by independent consultants.

In 1997 a formal evaluation policy was promulgated. At the same time the Evaluation Secretariat was established as a separate, independent entity within the Ministry of Foreign Affairs. Since then the development

evaluation scene has changed with joint evaluations becoming much more prominent. The Millennium Development Goals, the Rome Declaration on Harmonisation and Alignment and the Paris Declaration on Aid Effectiveness have wide implications for the ways evaluations are carried out. Consequently the evaluation policy was revised in 2006 (see Annex 2).

The first evaluation guidelines were adopted in 1986. They were substantially revised in 1994 and again in 1999. These guidelines build on the 1999 guidelines and reflect the new policy as well as the DAC Guidance for Conducting Effective Joint Evaluations and the DAC Evaluation Quality Standards, both issued in 2006.

The primary purpose of the Evaluation Guidelines is to communicate to partners and external consultants Danida's expectations of the quality of evaluations. They constitute a framework built on principles, criteria, standards, good practices and information about Danida's evaluation process. Particular emphasis has been put on defining the roles of various stakeholders in evaluations and on developing codes of conduct for these stakeholders.

The guidelines are not intended as an evaluation manual. Danida requires that its external consultants have the competencies expected of evaluation professionals: a sound grasp of evaluation methodologies, the skills and abilities to carry out evaluations as well as up to date knowledge of developments in the field of evaluation, in particular development evaluation.

Many people have contributed to these guidelines. First and foremost the staff of the Evaluation Department with support from Ian Davies. Feedback from consultants, colleagues and partners on how we do business has been invaluable as has been comments on drafts from Ray Rist and Linda Morra of IPDET.

Niels Dabelstein
Head of Evaluation Department

TABLE OF CONTENTS

INTRODUCTION

Purpose
The primary purpose of the Evaluation Guidelines is to communicate to partners and external consultants Danida's expectations of the quality of evaluations carried out on its behalf.

Quality
Because quality touches on different aspects of the evaluative enterprise such as use, methodology and management, the guidelines explain Danida's approach to evaluation of development interventions and identify those attributes it considers important to quality.

The Evaluation Guidelines apply to evaluation activities financed fully or partly by Danida: they constitute a framework for achieving quality that is built on principles, criteria, standards, good practices and information about Danida's evaluation process.

External consultants
The guidelines are not intended as an evaluation manual. Danida requires that its external consultants have the competencies expected of evaluation professionals: a sound grasp of evaluation methodologies, the skills and abilities to carry out evaluations as well as up to date knowledge of developments in the field of evaluation, in particular development evaluation.

As such, the guidelines incorporate concepts, methods, tools and issues basic to professional evaluation practice without discussing them at length; in some cases they provide a reference to more information.

Interested parties
The guidelines are also intended as a source of information on Danida's evaluation practice for partner countries, colleagues outside of the Evaluation Department (EVAL), staff of the Ministry of Foreign Affairs and of the government of Denmark in general, Danida's technical assistance

personnel in the field, bilateral and multilateral assistance organisations, and the broader evaluation community.

Contents
The guidelines are presented in five chapters.

Chapter 1 provides key definitions and distinctions used by Danida in its evaluation activities, identifies main parties to the evaluation process and discusses joint evaluation in Danish development assistance.

A discussion of differences and similarities in evaluating different types and aspects of development cooperation is included in Chapter 2.

Chapter 3 presents the steps and procedures taken by the Evaluation Department to prepare the evaluation, oversee its implementation and disseminate its results.

The evaluation criteria that serve as a basis for developing evaluative questions are presented and discussed in Chapter 4.

Chapter 5 addresses the implementation of the evaluation by the external evaluator.

CHAPTER 1
EVALUATION DEFINITION AND PARTIES

What is evaluation?

In general terms evaluation is used in society at large to establish achievements with some accuracy. All major development agencies involved in international development assistance undertake formal evaluations of part of their development activities each year.

Definition
Danida uses the OECD Development Assistance Committee's[1] (DAC) definition of evaluation.

> *An assessment, as systematic and objective as possible 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, developmental 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.*
>
> (DAC 1991)

Evaluation
Danida carries out or participates in evaluations at different stages of development interventions and defines their purpose as either primarily knowledge or accountability uses, or both.

Evaluation of *ongoing* development activities with a primary purpose to generate information to improve the quality of the intervention typically focus on implementation issuesand and operationel activities, but may

---

1) *OECD (1991), Principles for Evaluation of Development Assistance. Paris.*

also take a wider perspective and consider effects. As they are performed usually about midway in the cycle of the intervention, these evaluations may also be referred to as mid-term evaluations by other organisations.

Other evaluations are undertaken after completion of the aid intervention to understand the factors that affected performance, to assess the sustainability of results and impacts, and to draw lessons that may inform other interventions. The terms 'formative' or 'summative' evaluations are also utilized to distinguish the evaluation types.

Management's operational tools

Although there is increasing agreement in the evaluation community on the use of basic terms[2], differences exist and it is useful for clarity to explain the terms used in evaluation and various assessment tools related to the operational preparation and monitoring of the implementation of development interventions:

Appraisal

Some agencies use the term evaluation for prospective studies conducted prior to a project's approval. For example the European Commission conducts ex-ante evaluations of intended programmes[3].

Such studies are termed "appraisals" by Danida: "The aim of the appraisal is to provide quality assurance at an advanced stage of the preparation process of a programme or a component, but early enough for the recommendations provided by the appraisal to influence the final preparation of programme and component documents." (Guidelines for Programme Management)[4]. Danida's Technical Advisory Services is responsible for all major appraisals

---

2) *OECD (2002), Glossary of Key Terms in Evaluation and Results Based Management. Paris.*
3) *European Commission (2001), Guide to Ex Ante Evaluation. Bryssels.*

Review

DAC defines "review" as an assessment of the performance of an intervention, periodically or on an ad hoc basis. Danida refers to the same definition in its Aid Management Guidelines. The main distinction is that a review is regarded as an internal management tool for operational monitoring of the implementation of the targets of the interventions, while the evaluation is an independent, in-depth external assessment of the objectives, implementation and results of the interventions. Major reviews are undertaken by the Technical Advisory Services, frequently as part of joint partner reviews.

Performance monitoring

The trend of public administrations worldwide to reform their input-driven and highly procedural administrative practices in favour of results-based approaches to management and accountability, and the movement in the international donor community towards providing a increasingly greater proportion of funding in the form of sector or general budget support to governments of partner countries, has underlined management's responsibility to measure and report on the achievement of agreed results. Performance measurement of set targets has become another term for the management's monitoring function. Danida's Quality Assurance Department undertakes regular performance reviews of the Danish Embassies.

Performance audit

The distinction between performance audit and evaluation can be thin as typically both deal with the "what" and "why" questions of programme and organisational performance.

A useful way to distinguish performance audit from evaluation generally is to highlight that performance audit is carried out primarily to provide assurance, i.e. accountability, on a direct or attest basis, with knowledge-generation a secondary spin-off. Evaluation on the other hand is carried

4) *Danish Ministry of Foreign Affairs. Guidelines for Programme Management. Copenhagen.*

out usually to inform policy and decision-making as well as to produce knowledge.

Main parties to an evaluation

Successful evaluations are based on collaboration between key partici-pants and due consideration to the interest of the different parties in the evaluation. These include the commissioning organisations, evaluators, users, and stakeholders.

Main parties to an evaluation

Commissioner	The organisation that initiates and funds the evaluation such as Danida, other donor agencies partner countries or these jointly.
Evaluators	Independent expertise from private companies or public institutions.
Stakeholders	Agencies, organisations, groups or individuals with direct or indirect interest in the development intervention or its evaluation.
Users	A sub-set of stakeholders with formal or direct involvement in the development intervention or its evaluation: decision-makers with a formal role in following up the evaluation, e.g. policy-makers, auditors, etc., and managers of development interventions.

Evaluations are commissioned by Danida's Evaluation Department, alone or jointly with other donors, and with the partner country.

Increasingly, joint evaluations are used to rationalise the evaluation process, reduce transaction costs for partner countries, improve quality of the work undertaken, and increase weight and legitimacy of the evaluation. Usually one of the collaborating agencies is responsible for commissioning the joint evaluation. (See next section on joint evaluation).

Evaluators are the organisations or individuals that carry out the evaluation, i.e. collect and analyse data, judge the value of the assistance intervention and produce the evaluation report. They are selected usually among international and Danish consulting firms and research institutions.

Stakeholders are those agencies, organisations, groups or individuals that have an interest in the development intervention or its evaluation but not necessarily a formal role in it.

Stakeholder and user participation in the evaluation process promotes consensus and ownership in relation to development activities. Participatory approaches are capacity building processes that develop evaluative thinking and contribute to indigenous knowledge and sustainability.

Users of evaluations are stakeholders with a specific relationship to the intervention and its evaluation. They include policy makers, Danida policy and programme managers and staff, managers of assistance interventions, developing-country partners and other parties with a formal or direct role in relation to the development activities under evaluation.

To impact decision-making, evaluations address questions that are relevant to users and responsive to their information needs, i.e. evaluative information is meaningful, reliable and timely.

Inclusion of users and stakeholders in the evaluation process contributes to useful evaluations. Dialogue with stakeholders and users improves the evaluator's understanding of, and the evaluation's responsiveness to, their needs and priorities. The views and expertise of groups affected by the intervention and its evaluation are considered and integrated to the evaluation process whenever appropriate through mechanisms such as steering or advisory groups.

Joint evaluation

Joint evaluations of development assistance have been conducted by various groups of donors and countries since the late 1980's and a significant number have been undertaken since 2000. Danida has participated in joint evaluations since 1989 (see box).

What is joint evaluation?

According to the DAC Network on Development Evaluation[5]: "Joint evaluations are development evaluations conducted collaboratively by more than one agency. The focus here is not on participatory evaluation with its techniques for bringing stakeholder communities into the process, but on evaluations undertaken jointly by more than one development cooperation agency. Joint evaluations vary considerably in the number of participating agencies and in their focus, purpose and approach. Methodologies used can also differ widely, ranging from desk reviews of existing information to fieldwork in developing countries."

Danida and joint evaluation

The changing modes of Danida's, and the global donor community's, development assistance increase the significance of joint evaluations. Sector programme support and national development programmes rely on collaborative, multi-donor assistance efforts with shared common objectives, joint strategies and on partner country leadership roles.

Effective evaluations of such new types of assistance call for more collaboration. Clearly, it is a waste of resources on both the donor and partner side when several donors perform individual evaluations of their support to the same sector.

Key objectives of a joint evaluation process are to ensure that the evaluation becomes an efficient learning tool, helps promote good governance, enables the partners to be fully accountable and is cost-effective. Joint evaluations can focus on vital areas and help consolidate international responses and development policy.

5)   *OECD (2006), DAC Guidance for Managing Joint Evaluations. Paris.*

Joint evaluations with Danida leadership or participation

- The Nordic Agricultural Programme in Mozambique (MONAP) was evaluated jointly by Denmark, Norway and Sweden in 1991 which resulted in a shift of priorities of the three countries.
- UNICEF's technical program was evaluated jointly by Australia, Canada, Denmark and Switzerland in 1993. It contributed to major changes in UNICEF's mode of operation.
- The international response to conflict and genocide in Rwanda was evaluated jointly by 38 bilateral donors under Danish leadership, UN agencies and NGOs in 1996. It identified key issues and problems of international humanitarian assistance and has induced changes of policies and strategies of humanitarian agencies.
- The road sub-sector programme in Ghana was evaluated jointly by eight development agencies and the Ministry of Road & Transport in 2000. A follow-up study was conducted in 2005.
- The Basic and Primary Education Programme in Nepal was evaluated jointly by five development agencies and the Ministry of Education & Sports in 2004.
- The humanitarian and reconstruction assistance to Afghanistan was evaluated jointly by five development agencies in 2004-05.
- The Ugandan Plan of Modernisation of Agriculture (PMA) was evaluated jointly under the overall coordination and guidance of the PMA Steering Committee in 2005.
- The Tsunami Evaluation Coalition with over 50 member agencies from across the humanitarian sector evaluated jointly the response to the Asian earthquake and tsunamis in 2005-06.
- The joint evaluation of general budget support was carried out in 2003-06. More than 20 development agencies were represented in the steering committee together with seven partner countries.

Types of joint evaluations

Joint evaluations, like single agency evaluations, come in different sizes and shapes: their focus spans a continuum ranging from project to policy

including topics such as assistance modalities, themes, delivery mechanisms, etc. They have different purposes, cover single countries, regions or more, use a variety of approaches and methodologies, and involve developing country partners to varying degrees.

In joint evaluations including larger groups of contributors, the involvement of the majority is typically limited to a few key stages such as review and approval of the design and final product. Management and coordination is typically delegated to one or a few agencies taking a lead responsibility or to an external organisation designated to manage the study.

The following table describes the main characteristics of three broad categories of joint evaluations, based on their degree of "jointness", i.e. the extent to which individual partners cooperate in the evaluation process.

Acknowledging the existence and typology of many complex forms of joint evaluations contributes to reducing confusion and misunderstanding when partners work together.

Benefits and challenges of joint evaluations
Joint evaluations provide a number of significant benefits to donors and recipient countries alike particularly in terms of reduced transaction costs for developing country partners and improved harmonisation of development interventions among donors; however, they also present their own set of particular challenges due primarily to the increased complexity of managing and conducting this type of evaluation.

A key advantage of joint evaluations is that they have greater credibility and broader ownership of findings among the larger development community than would be the case with single agency evaluations.

Their value lies in the comprehensiveness of findings and in the validity of lessons learned. They generally yield useful best practises, focus policy discussions among development agencies, and achieve an impetus for joint action that is out of reach of smaller evaluations.

Type of joint evaluation	Mode of work / examples
1. **Classic joint evaluation**	Participation is open to all stakeholder agencies. All partners participate and contribute actively and on equal terms. Examples of classic joint evaluations include: the Rwanda Evaluation, the tripartite evaluation of WFP, the CDF Evaluation and the GBS Evaluation.
2. **Qualified joint evaluation**	Participation is open only to those who qualify – through membership of a certain grouping (e.g. DAC, EU, Nordics, UNEG, ECG, Utstein) or through active participation in the activity (e.g. SWAp, basket funding mechanism, jointly-implemented programme) that is being evaluated. Examples include EU aid evaluations, the evaluation of the Road Sub-sector in Ghana, the Basic Education Evaluation, the ITC Evaluation, and the evaluation of Uganda's Plan on Modernisation of Agriculture (PMA).
3. **Hybrid joint evaluation**	Includes a wide range of more complex ways of joint working: (a) responsibilities are delegated to one or more agencies while others take a 'silent partnership' role; (b) some components of the evaluation are undertaken jointly while other parts are delivered separately; (c) various levels of linkage are established between separate but parallel and interrelated evaluations; (d) the joint activity is agreeing a common evaluation framework and responsibility for implementation of individual evaluations is devolved to different partners; (e) research, interviews and team visits are undertaken jointly but each partner prepares a separate report. Examples include Support to Internally Displaced Persons and the Tsunami Evaluation.

On the other hand, just as co-ordination of development interventions is notoriously difficult, the same applies to evaluations. Donors prepare their evaluation programmes in response to agency needs for lesson learning and accountability, and these are geared to the planning and programming cycle of the agency. For evaluation to develop into a joint management and accountability tool, the donors need to be flexible, to adjust to the planning cycle of the partner countries and to enable them to take the lead in evaluations.

The following is a summary[6] of the key benefits and particular challenges of joint evaluations seen from the perspective of commissioners and organisers:

Benefits

- *Diversity, Legitimacy and Influence.* Joint evaluations enable a diverse range of perspectives and talents on the evaluation. This can lead to high quality, transparent and credible evaluation reports, with broad ownership of findings and greater legitimacy and influence on decision-makers.
- *Broader Scope.* Joint evaluations enable a broader number of evaluation questions to be addressed, given extra (jointly-shared) resources.
- *Harmonisation and Alignment.* Donors are increasingly prioritising partnership-based assistance modalities. Sector-wide approaches, budget support and country-led poverty reduction strategies all stress greater participation and leadership roles for partner countries. Effective evaluations of such assistance modes call for – even require – greater collaboration. Joint evaluations therefore have distinct advantages as development agencies shift their strategies away from isolated projects towards programmatic approaches. Joint evaluations also foster greater consensus among the partner agencies on upcoming development priorities and needs, thus stimulating improved coordination of future programming.

6) *Breier, H. (2005), Joint Evaluations: Recent Experiences, Lessons Learnt and Options for the Future, DAC Evaluation Network Working Paper, OECD, Paris.*

- *Rationalisation and reduced transaction costs for developing countries:* Joint evaluations reduce the burden and associated costs of multiple donor evaluation efforts on partner country institutions, including overlapping team visits and duplicative data collection. Joint evaluations also help to avoid conveying to partner countries too many different and sometimes conflicting evaluation messages.
- *Mutual capacity development:* Joint evaluations provide opportunities for agencies to learn from each other and to share their evaluation processes and techniques. They can also lead to enhanced use of local consultants, with consequent local capacity development.
- *Cost-sharing:* Joint evaluations are normally undertaken with shared or pooled financing, and the cost-burden is therefore divided among the various partners.
- *Applicability:* Joint evaluations have additional advantages and applicability in certain specific circumstances: if there are issues in the evaluation that are too sensitive or controversial for one agency alone to tackle; if the activities being evaluated have been jointly financed or implemented; when a meta-evaluation is being undertaken; and when an evaluation of the work of a multilateral organisation is being undertaken.

Challenges are often related to difficulties in harmonising differences among agencies' objectives, operational processes and organisational cultures.

- *Planning joint evaluations:* A particular problem is that by the time agencies share or make public their evaluation work programmes, it may be too late in the planning cycle of any given agency to agree to undertake it jointly. Also, all agencies have limited time and resources, and getting one's 'own' work completed can take priority over 'joint' efforts.
- *Consensus on purpose:* Agreeing the overall and political objectives for the evaluation, and ensuring that no partners are driving hidden agendas.
- *Consensus on Terms of References:* Agreeing comprehensive, yet manageable, Terms of Reference that accommodate the particular issues and interests of all participating agencies.

- *Management:* Developing joint management structures and communi-cation processes that work effectively.
- *Suitability of evaluators:* Selecting evaluation teams that are mutually acceptable to all participants.
- *Logistics:* Coordinating schedules and travel logistics amongst the various partners.
- *Evaluative process:* Reaching agreement on methodologies, recommen-dations and reporting formats.

As well, joint evaluations create particular conditions for the external evaluators charged with carrying out a joint evaluation and some princi-ples have been developed for organizing and managing the external evaluation team during joint evaluations.[7]

7) *Freeman, T. (2004). Managing the External Evaluation Team During Joint Evaluations: Challenges and Responses. Ottawa, Canada: Goss Gilroy Inc.*

CHAPTER 2
DANIDA EVALUATIONS

Evaluating development assistance

The emergence of new types of development interventions and the focus on specific priority themes in development assistance poses new challenges to evaluators.

Some of those challenges are methodological: how to evaluate support for good governance, human rights, civil service reform and private sector support.

Some are concerned with policy: how to move from project assistance to more varied and flexible modes of assistance within sector frameworks designed to address policy and institutional problems; how to improve development co-operation and partnership; and how to evaluate the increasing volume of humanitarian relief operations.

Danida's evaluation programme includes single intervention evaluations (although these are increasingly rare), impact evaluations, sector evaluations, country programme evaluations, evaluations of forms of development cooperation, multilateral evaluations and thematic evaluations.

As the breadth, depth and diversity of Danida's evaluation topics and collaborations expand, including joint evaluations with other donors or partner governments, their growing complexity poses increasing methodological and logistical challenges to the oversight and conduct of evaluations.

Over time Danida has shifted the focus of its evaluations from projects to more complex modes of development assistance, e.g. sector programme, country programme or thematic evaluations, and these have proved to be increasingly cost effective and with more impact on policy decisions.

Focus on Danida Evaluations

Single intervention	Evaluation of an individual development intervention designed to achieve specific objectives within specified resources and implementation schedule, often within a broader program.
Sector	Evaluation of a cluster of development interventions in a sector within a country or across countries, all of which contribute to the achievement of a specific development goal.
Country Programme	Evaluation of one or more donor's or agency's portfolio of development interventions, and the assistance strategy behind them, in a partner country.
Modes of development co-operation	Evaluation of a specific instrument or channel for development assistance (research, NGOs, humanitarian assistance, balance of payment support, general budget support, technical assistance, etc.).
Thematic	Evaluation of a selection of development interventions, all of which address a specific development priority that cuts across countries, regions and sectors.
Impact	Evaluation of a project, programme or policy to assess the intended or unintended effects, positive or negative, of a specific intervention on people's welfare.
Humanitarian	Evaluation of humanitarian action to draw lessons to improve policy and practise and enhance accountability.

Single intervention evaluations

A single intervention, traditionally called a development project, is defined as a time-bound intervention designed to achieve specific objectives with specified resources, usually within a short to medium term implementation schedule, and often within a broader program.

The limited scope of the single intervention makes it relatively easy to focus the evaluation.

Evaluation of single interventions

Achievements of objectives	Impact

Outputs

Inputs

A single intervention can be described as a single cause-effect chain of identifiable elements which makes it easier to focus the evaluation.
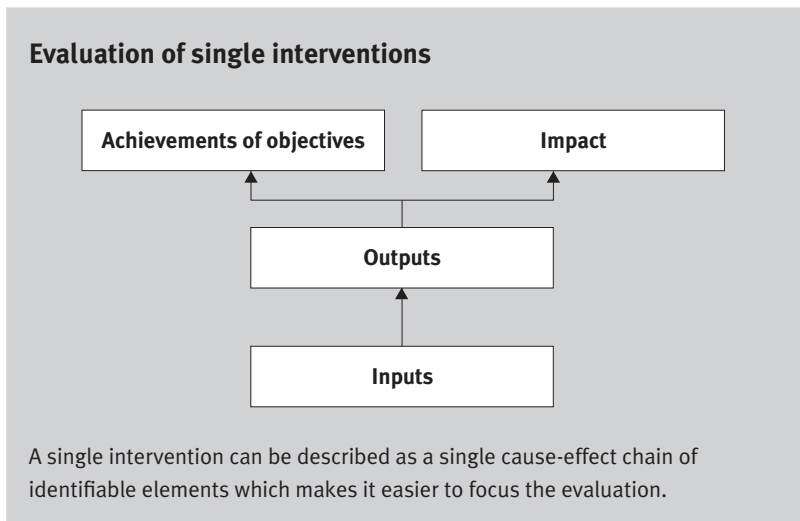
Although overall project success is commonly determined by whether outputs are achieved on time and on budget, single interventions in development contexts are evaluated usually from the broader perspective of attainment of objectives and impact, in order to learn from the experience and to improve the design of similar future projects.

To the extent that objectives are specifically expressed and in measurable terms, single intervention evaluations present few methodological challenges.

However, a typical complicating factor is that the objectives are often complex, confounding, unrealistic, ill defined, or change as the project develops. It follows that a single intervention is not only judged on the basis of what has formally been agreed to, rather the assessment takes into account what can be achieved realistically with the resources available.

A single intervention evaluation takes into consideration, yet distinguishes between, the assessment of the project's outputs, i.e. the quality and quantity of what it is expected to produce, and its related, broader and more complex, effectiveness and impact questions.

Sector programme evaluations

Sector programmes are a well established mode of support for Danida in which the organisational performance of relevant institutions is a key factor for successful development.

Sector programme evaluations

Sector programme evaluations are more complex than single intervention evaluations: there are often several donors as well as government involved making it difficult to identify and assess each donor's contribution.
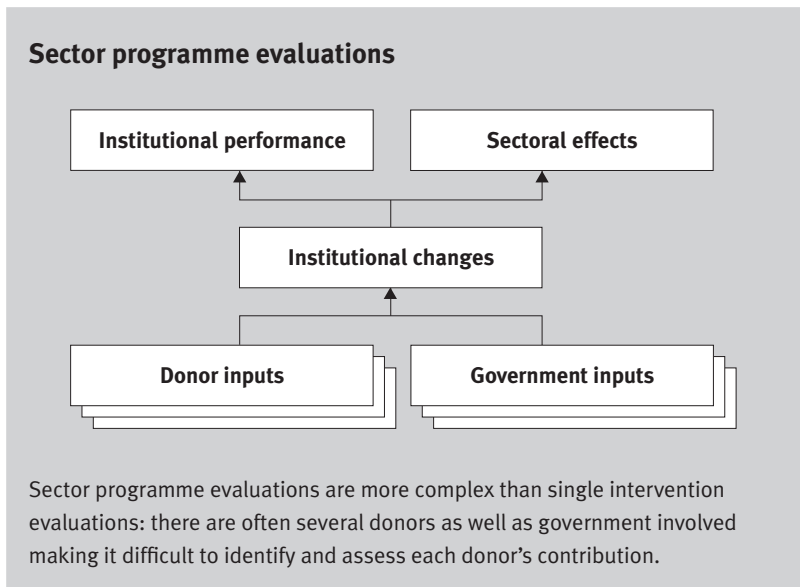
As such, sector programme evaluations focus on questions of institutional performance, processes, changes and interrelationships, and on their effects on the sector.

Questions of organisational capacity of institutions, partner country ownership and responsibility are central to evaluating the success of sector programme support interventions and their sustainability.

A key success factor in sector programme evaluations is the involvement of the partner country and its ownership of the evaluation results.

Consistent with the emphasis on development partnerships, local ownership and good governance, donors increasingly program sector assistance jointly and assist in developing an evaluation culture by incorporating joint evaluations into sector programmes.

Ultimately the goal is for partner countries to coordinate joint evaluations in their sectors.

Country programme evaluations

Country programme evaluations focus on the entire Danish assistance to one of the programme countries.

They provide an assessment of past interventions to Danida and partner countries to improve cooperation strategies, country programmes and sector interventions. As well, they generate knowledge to improve future assistance to the country and other national country programmes.

A country programme evaluation looks at the relevance of Danish assistance against Danida's overall policy and country strategy, as well as the development policy of the partner country. It reviews the instruments used in the bilateral cooperation, the modalities, and the relative weight given to assistance for economic and social development.

Country programme evaluations

| Policy performance | Development effects |

Policy/strategy changes

Donor inputs

Country programme evaluations are complex in the sense that they will relate to policy rather than objectives and usually involve three to four sectors and different forms of assistance.
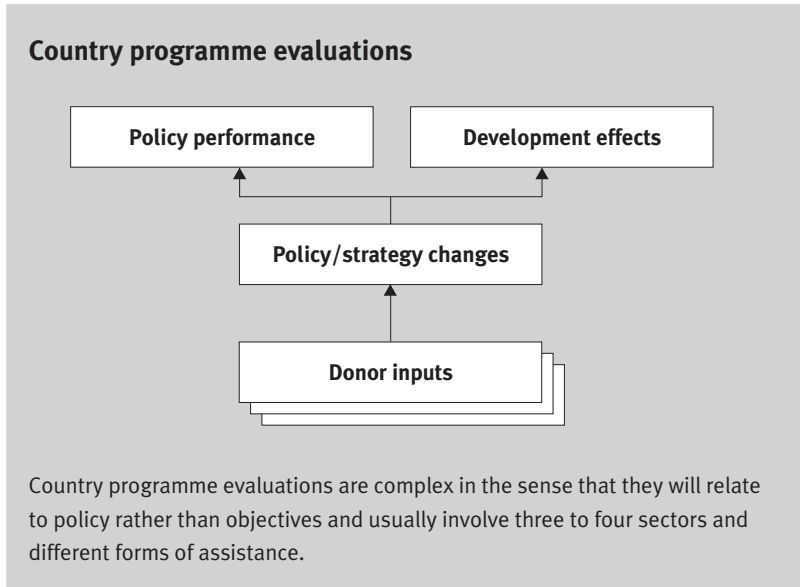
Country programme evaluations are important for policy planning at the highest level and provide a basis for bilateral negotiations. They represent an opportunity to focus on specific relevant issues such as the country's dependency on development assistance, institutional capabilities, democracy, human rights, gender, and environment in a wide context.

Because of the importance of these evaluations, involvement of the partner country in the evaluation process and its ensuing acceptance, ownership and use of the evaluative information is paramount.

Country programme evaluations are truly cross-cutting, i.e. they cover all sectors and modes of cooperation, and, as a result, focussing the evaluation poses difficult challenges. The evaluation team is usually interdisciplinary, with expertise reflecting the key issues the evaluation will focus on.

To focus the country programme evaluations within the resources available, the scope is limited typically to the development issues and the strategic choices made at the national and overall levels; the economic,

political and social context of the country; and the individual development agency's strategy or the joint assistance strategy of the involved agencies.

The evaluation will take its departure from other available sector or programme evaluations and programme documentation. However, the fieldwork will cover observations and interviews with government officials, programme staff, beneficiaries and interested parties as appropriate.

Evaluations of modes of assistance

These evaluations focus on the specific instrument or channel for development assistance, e.g. research, NGOs, balance of payment support, general budget support for poverty reduction, technical assistance. Such evaluations are usually cross-sectoral and cross-country and designed to assess the efficiency of a particular mode of assistance as a means of promoting development.

Evaluations of assistance instruments are initiated to provide accountability information on past activities and to generate knowledge extract to improve future performance. The focus is usually on the efficiency of assistance delivery, but also to some extent on the effect of assistance delivered under a given mode of cooperation.

The methodological approach varies according to the form of development assistance considered. Evaluating humanitarian assistance requires a significantly different approach from evaluating research assistance or balance of payment support.

A typical problem in evaluating assistance instruments is that the individual cases being assessed often have different objectives. This makes it difficult to compare systematically the effects of different modes of development assistance and their impact. As a result, these evaluations tend to focus more on the performance of the various parties involved in development assistance and the use of the resources made available.

Evaluation of modes of assistance

```
┌─────────────────────────┐      ┌─────────────────────────┐
│  Recipient performance  │      │      Effects of aid     │
└─────────────────────────┘      └─────────────────────────┘
              ▲                              ▲
              └──────────────┬───────────────┘
                  ┌──────────────────────┐
                  │    Delivery of aid   │
                  └──────────────────────┘
                             ▲
                  ┌──────────────────────┐
                  │        Inputs        │
                  └──────────────────────┘
```

The focus is on efficiency and effectiveness of specific modes of assistance. The evaluation is complex in the sense that it may involve many sectors and countries, as well as a multitude of different objectives.
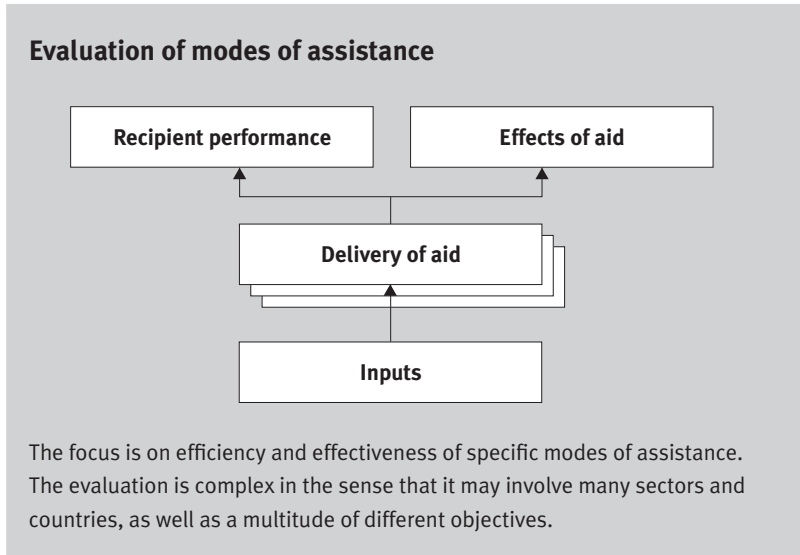
Evaluation of assistance instruments has a narrow perspective yet allows lessons to be drawn from larger samples of development interventions, that in turn can have wide ranging implications for policy and to improve bilateral co-operation.

Thematic evaluations

Thematic evaluations deal with selected aspects or themes in a number of development activities. There are several such themes that have been highlighted in Danish development assistance over the years. These themes are often borne out of policy statements and often termed "cross-cutting issues".

This means that they are systematically integrated in all sectors of co-operation and in all modes of assistance. Some of the factors that have been highlighted in recent years are gender aspects, institutional development, human rights, and environmental aspects.

Over the years, such cross-cutting issues have been explored on a project-by-project basis by including specific questions in the Terms of Reference. This provides a wealth of useful information when thematic evaluations are undertaken.

Thematic evaluations

Evaluation of Sector A	Evaluation of Sector B	Evaluation of Sector C	→	Thematic evaluations

Projects/Components	Projects/Components	Projects/Components	→	
Community Support	Community Support	Community Support	→	Evaluation of modes of assistance
Technical assistance	Technical assistance	Technical assistance	→	
Institutional support	Institutional support	Institutional support	→	
Research	Research	Research	→	
Etc.	Etc.	Etc.	→	
Sector A	Sector B	Sector C		

The thematic evaluation extracts aggregates of information on a specific theme. It may involve different assistance instruments, sectors or countries, and make use of information from different types of evaluations.

Thematic evaluations cut across countries, regions and sectors and have usually the greatest complexity of the evaluations. The topic is analysed in the relation to Danida's development policy, but also in the context of international conventions and the partner's priorities and strategies. As in all complex evaluations, a key methodological challenge lies in distinguishing between the impact of Danish assistance and the impact of national activities and policies and of other donor initiatives.

Another common challenge in thematic evaluations is, in the preparatory phase, to draw a sample of development activities that reflects Danish development assistance in countries, sectors and forms of assistance.

Thematic evaluations are usually based on relatively large samples of development activities implemented over a relatively long period of time. Much of the study contends with older documents on file that are incomplete or marginal to the theme; as a result, fieldwork concentrates on the more recent development activities, aims at verifying information and acquiring a developing an understanding that is meaningful in current contexts.

For all their difficulties, thematic evaluations have proved useful instruments in generating specific knowledge and recommendations at the highest level of aggregation, i.e. the policy level.

Impact evaluation

Impact evaluations do not focus on the type of intervention that is being evaluated, i.e. the evaluand, but rather on results different from outputs or outcomes. Impact evaluations assess the positive and negative changes in terms of people's welfare, produced by a development intervention, directly or indirectly, intended or unintended.

By examining whether changes in the well-being of individuals can be attributed to a specific intervention, an impact evaluation can help understand whether or to what extent an intervention has produced the intended effect, e.g. increased incomes, decreased maternal mortality, increased literacy, etc.

An impact evaluation can also help to judge the value of an intervention in relation to its cost and provide information to decide whether it should be continued, scaled-up, altered or otherwise.

CHAPTER 3
DANIDA'S EVALUATION PROCESS

Evaluation programme

Danida's evaluations are planned on the basis of a rolling two-year pro-
gramme that is updated annually by the Evaluation Department. All
relevant departments and representations are consulted and invited to
propose evaluations to be included in the programme.

The Evaluation Department ensures that the programme reflects the
distribution of Danish development assistance in countries, sectors,
assistance instruments and priority areas.

The evaluation programme is posted on the Evaluation Department's
website www.evaluation.dk.

The main steps in Danida's evaluation process, from the initial decision
through to dissemination, use and follow up are described in the follow-
ing sections.

Launching an evaluation

The Evaluation Department ensures that all involved parties are inform-
ed about the evaluation in due time before it is initiated. In the case of
larger evaluations this is done through an approach paper that provides
information on:

1. The background for the evaluation
2. The main objective
3. Outputs (reports etc. to be produced)
4. Major issues (themes or aspects to be covered)
5. Approach (methodological considerations)
6. Organisation and management
7. Preliminary work plan with a time schedule

The approach paper indicates the extent to which the partners will be involved in preparation and implementation of the evaluation, as well as possible co-operation with other donors. It is distributed for information and comments to relevant Danida departments and representations as well as partner institutions.

Where there are many interests involved, a workshop may be organised as a basis for the approach paper, and a reference group for the evaluation appointed. The composition of the reference group reflects the topic and purpose of the evaluation and is decided in consultation with relevant Danida departments.

Steps, roles and responsibilities in preparing the evaluation

Evaluation step	EVAL
1. **Approach paper**	a. Prepares draft. b. Ensures involvement of relevant MFA staff and other stakeholders. c. Finalises paper.
2. **Terms of Reference**	a. Prepares draft. b. Ensures involvement of relevant MFA staff and other stakeholders. c. Finalises Terms of Reference.
3. **Tender process**	a. Requests support from ERH. b. Prepares Terms of Reference for tender consultant. c. Participates in tender committee.

BFT is Danida's Technical Assistance Services
ERH is Danida's Contracts Department

Terms of Reference

The Terms of Reference provide the background for the evaluation, define its objectives and scope, the composition of the evaluation team and the timing.

The Evaluation Department drafts the Terms of Reference based on the approach paper, and ensures that the evaluation is consistent with Danida's evaluation policy and practices.

Other MFA staff	Consultant	Others
Embassy and BFT: Consultation.		*Partners, reference group and possibly other donors:* Consultation.
Embassy and BFT: Consultation.		Partners, reference group and possibly other donors: Consultation.
ERH a. Announces pre-qualification. b. Announces tender. c. Hires tender consultant. d. Chairs the tender committee. e. Enters contract with consultant.	Prepares and submits bid.	*Tender consultant:* a. Prepares specifications. b. Assesses Letter of Interest. c. Assesses proposals. d. Participates in tender committee *Partners:* Participate in tender committee.

It is essential to consider the information needs of all involved parties. To this end a workshop or seminar may be organised to bring together the Evaluation Department, the representation involved, Danida's Technical Assistance Services (BFT), the partner country (the line ministry), the project's management (in single intervention evaluations) and other development agencies (in sector programme evaluations). Since the partner should be involved as a matter of priority, such workshops may be organised in the partner country.

The Terms of Reference include the following chapters:

1. Background
2. Objective
3. Output
4. Scope of work
5. Methodology
6. Work plan
7. Composition of evaluation team
8. Documents available

Evaluations are usually carried out in several phases such as desk study, fieldwork, and synthesis. In such cases overall Terms of Reference provide the framework for the evaluation and one of the objectives of the desk study is to formulate detailed fieldwork plan for the evaluation.

Assigning the evaluation team

Danida, with the participation of the partner country, establishes the evaluation team by contracting with individual consultants or, as is typically the case, with consulting organisations. Competence in the field and experience relevant to the task are the key selection criteria.

All Danida evaluations are tendered according to EU procurement directives and invitations to bid are disseminated internationally, i.e. beyond the EU. Tenders are appraised on the basis of topical content, professional composition and competence of the evaluation team, and

price. This tender procedure requires a preparation period that may take three to six months more than the process for evaluations below that threshold.

In cases of partner-led joint evaluations, the lead partner's procurement policies and processes may apply.

The success of an evaluation depends on the composition of the evaluation team and the competence and personal abilities of the team members. This applies in particular to the team leader who should be the one concerned with the overall perspective, able to organise and co-ordinate the work of the team members, assess the quality and relevance of their contributions and act as a spokesperson for the team.

Members of the evaluation team are selected to represent relevant professional areas and to include professional expertise from the programme country whenever possible.

To safeguard impartiality, members of the evaluation team may not have been personally involved in the activities to be evaluated; as well, organisations conducting evaluations may not have been involved in the preparation or implementation of those activities.

Overseeing the evaluation

The Evaluation Department is responsible for managing the contractual relationship with the selected evaluation team to make sure that the evaluation is consistent with its Terms of Reference Danida's evaluation policy, good evaluation practise and that it is carried out in cost-effective and timely fashion.

As well, the Evaluation Department is responsible for ensuring that the evaluation process and products meet the DAC Evaluation Quality Standards.[8]

8) Annex 2.

The DAC Standards are intended to:

• Provide standards for the process (conduct) and products (outputs) of evaluations;
• Facilitate the comparison of evaluations across countries (meta-evaluation);
• Facilitate partnerships and collaboration on joint evaluations;
• Better enable member countries to make use of each others' evaluation findings and reports (including good practice and lessons learned); and streamline evaluation efforts.

The Evaluation Department may set up a reference group for each evaluation that includes representatives from the main users of the evaluation and serves to provide guidance and feedback to the external evaluation team. In the case of joint evaluations the structure and mandate of the reference group may differ.

Furthermore, to strengthen quality assurance, the Evaluation Department may use external peer reviewers.

Once the evaluation team has been assigned, the Evaluation Department monitors the conduct of the evaluation in each of the following phases:

Inception
During this first phase the evaluation team works on the basis of the Terms of Reference with the Evaluation Department to develop a detailed plan for the field study phase.

The evaluation team discusses any methodological adjustments with the Evaluation Department, whose role it is to consider significant proposed changes to what was described in the Terms of Reference and reach agreement with the evaluation team on the course of action.

At the end of this first phase, the evaluation team produces an inception report, i.e. a detailed operational plan for the conduct of the evaluation fieldwork that is submitted for approval to the Evaluation Department.

The inception report is reviewed by the Evaluation Department and the reference group and, where appropriate, consultations take place with other key partners and stakeholders before the Evaluation Department signs off on the inception report.

Fieldwork
This second phase encompasses typically data collection, analysis, discussion of findings, conclusions and recommendations through to the drafting of the final report.

The role of the Evaluation Department is to facilitate the gathering of information and general conduct of the fieldwork by the evaluation team, while monitoring that the evaluation plan is being followed and changes discussed and agreed to with the team.

As well, the Evaluation Department is available to provide guidance and assistance to the evaluation team in dealing with management, logistic and methodological issues as they arise. It acts to facilitate coordination of the evaluation with the country partner, Danish representation in the country, collaborating organisations, users and stakeholders.

Together with the reference group, the Evaluation Department considers and discusses with the team the evaluation's findings and the preliminary conclusions and recommendations before these are included in the draft final report.

It makes sure that all relevant users and stakeholders are included in the draft discussion process as appropriate, while supporting the evaluation team's independence.

Use of evaluation

The users of evaluations are Danida, partner country authorities, stakeholders, the media, politicians, the public, and external resources (researchers, consultants, professional agencies, etc.).

Making information available is essential to general awareness, interest and support for development assistance. The evaluation report provides a status of strengths and weaknesses of development interventions and suggests solutions to major problems. To the extent that it does so, evaluation is a piece of technical assistance in itself which is appreciated as a test of the viability of development activities and the bilateral co-operation between the partner authorities and Danida.

Follow-up of evaluations

At the conclusion of an evaluation, a follow-up memo is prepared, taking note of Danida's position on the conclusions and recommendations as well as identifying which departments are responsible for the agreed follow-up activities. The follow-up memo is discussed in the Programme Committee and signed off by the State Secretary. The Evaluation Department undertakes to monitor the implementation of the follow-up activities at regular intervals.

The Evaluation Department contributes actively to the dissemination of Danida's own as well as other organisations' evaluation experience via workshops and seminars for staff in co-operation with the Ministry's education section. Further, the Evaluation Department assists Danida's Centre for Competence Development in the dissemination of evaluation experience.

On the basis of Danida's own and others' evaluation experiences, "best practices" will be compiled and formulated. Furthermore, the Evaluation Department will contribute to the incorporation of evaluation experience in policies, strategies and guidelines, etc.

After the final report has been received, Danida decides what actions to be taken in the light of conclusions and recommendations presented. Salient issues are brought to the attention of Danida's senior level management. This may eventually be used to further develop and improve overall policies and methods of work. It is the task of the Evaluation

Department to systematise such experience and formulate practical proposals on that basis.

Summaries of evaluation reports are distributed to all Danida staff members. It is the responsibility of Danida's operational departments and Embassies to ensure that relevant past experience is built into the design and preparation of future activities.

The significance of an evaluation report is different for different users. It is essential that the final users of the information are able to utilise the evaluation system for their own needs. This implies that for the evaluation system to function properly, the demand for information needs to be carefully analysed. In an evaluation system designed to respond to demands rather than having a control function, the information is more likely to be utilised.

In order to ensure this linkage, the Evaluation Department has a role in assessing the needs of users and developing evaluation to respond to shifting policy and modes of development co-operation. Also, it is vital for the Department to assess continuously the quality of evaluation activities and explore mechanisms for feedback and improved learning.

The main feedback linkages are to make use of evaluation in planning of new development interventions, to manage existing activities, to develop policy and strategies, and to train staff members and external resources.

Apart from the evaluation report itself, feedback is provided through the summary of evaluation, the follow-up memorandum, specific studies and reports produced by the Evaluation Department, seminars and workshops, and guidelines issued by the Evaluation Department. Brief summaries of evaluation reports are submitted to the database of the OECD/DAC Network on Development Evaluation, and all Danida's evaluations are published and made public in the form of printed reports, summaries, and electronically on www.evaluation.dk.

Dissemination

Whether evaluation reports should be made public has been a matter of debate. Evaluation activities were originally not meant to be a means of accountability for development assistance. In the early 1980s, however, the policy was that the activities to inform the public to some extent should make use of the results of evaluations. The policy has since changed successively, and since 1989 all evaluations reports are available to the public.

The reaction has been generally positive. The availability of information has lifted the professional debate regarding Danish development assistance. Researchers, students, companies and individuals have benefited. Some of the debate in mass media has been more informed with the availability of evaluation reports.

Apart from distributing the evaluation report itself, the most common ways to disseminate evaluation information are through the evaluation summaries, annual reports, bibliographies, thematic reports, the web, seminars, press releases, and public debate.

Whatever channel is preferred, the best way to ensure dissemination of lessons learned and knowledge gained in evaluations is to improve both the content of reports and the presentation of material. A key benefit of good dissemination practises is transparency of development interventions and public insight into their value.

Code of conduct

Danida has developed a code of conduct to ensure the independence and credibility of evaluations by clarifying the responsibilities, proper conduct and mutual relations of consultants, the Evaluation Department, other Danida staff, and other parties responsible for the activities under evaluation. Other Danida staff includes both Embassy staff and Headquarters staff, while other parties responsible for the activities under evaluation include partner organisations and institutions in partner countries (e.g.

line ministries, private sector actors, and civil society organisations), organisations in Denmark (e.g. Danish NGOs), and multilateral organisations and other international institutions supported by Danida.

Evaluators have cited two particularly difficult phases in the evaluation process where uncertainty about roles may threaten their independence: organisation of fieldwork and discussion of the draft report. The code, therefore, concentrates on issues relating to these two phases in the evaluation process, while also addressing a range of other potential conflicts.

Other key actors in the evaluation process are the target group of the activities under evaluation, other persons interviewed and contacted during the process, reference groups and peers. When the evaluation is undertaken as a joint evaluation involving several donors and partners, the described role of the Evaluation Department is also applicable to the role of other donors (the lead agency).

The code for consultants can be found in Chapter 5 – Evaluation team tasks.

Code for Evaluation Department staff

The Evaluation Department is the client, and in the case of joint evaluations with other partners: organises the assignment, controls the quality of the evaluation method and the quality of the report and is responsible for dissemination and for monitoring the follow-up of evaluation results. The Evaluation Department prepares approach papers and Terms of References, identifies consultants, facilitates contacts to Danida staff and other relevant resource persons, facilitates the collection of documents, ensures thorough assessment of the evaluation method, approves the inception report, assesses the quality of the draft report and approves the final report.

• The Evaluation Department is responsible for briefing the consultants at the inception stage on the operations, the expected role of all parties involved in the process, relevant documents and data sources and as

often as necessary during the evaluation process. Contact between the Evaluation Department and consultants may take place at the initiative of either party.

- While consultants have chief responsibility for preparing the fieldwork, the task of the Evaluation Department is to facilitate the process and provide quality control of the final evaluation methodology and work plan. Consultants suggest sites for field visits, samples of interview persons, etc., but final decisions on methodology are subject to approval by the Evaluation Department. After the inception report has been approved, adjustments of the methodology described are subject to agreement between the Evaluation Department and the consultants.

- Evaluation Department staff participating in a scoping mission, the initial part of the fieldwork or in post-fieldwork debriefing meetings are responsible for facilitating the evaluation and assuring that the evaluation is proceeding correctly with regard to methodology, timing, etc. The participation and role of the Evaluation Department, and possible stakeholders, in meetings, interviews and field visits must be agreed with the consultant team leader. Stakeholders may include Embassy staff, Danida advisers and company advisers, partner representatives, other parties responsible for the activities under evaluation, and others.

- The Evaluation Department provides quality assurance of the report by checking the validity of evidence, the analytical rigour, the consistency of analyses, conclusions and recommendations, that requirements in Terms of Reference have been adequately dealt with, that the report is clearly structured and well-written, and that the report does not contain factual errors and inaccuracies.

- It is the responsibility of the Evaluation Department to synthesise comments received from Danida staff and other stakeholders into a set of consolidated comments pointing out which of the comments received are considered the most important, citing necessary changes to be made (comments relating to the issues above as well as further suggested changes, which the consultant may consider when preparing the final report).

- It is the responsibility of the Evaluation Department to ensure that relevant Danida staff and other stakeholders are kept informed about

the progress of the evaluation. After the draft report has been presented, all communication between consultants and Danida staff and other stakeholders should go through the Evaluation Department.

- The Evaluation Department is required to react to all requests for assistance relating to situations in which consultants feel their independence threatened. Thus, during all steps of the evaluation process, it is an important task of the Evaluation Department to protect consultants against undue pressure from Danida staff and other stakeholders. If consultants encounter insufficient assistance or outright resistance during the evaluation process, including when carrying out the fieldwork, it is the responsibility of the Evaluation Department to contact the persons involved and ensure that proper cooperation is established. If the atmosphere during discussions of the draft report turns hostile, or arguments are not professionally sustained, it is the task of the Evaluation Department to intervene. In the extreme situation where comments can be interpreted as subtle or overt pressure against consultants to achieve specific conclusions, the Evaluation Department has a particular responsibility to take immediate action.
- If consultants suspect mismanagement, corruption or other illicit practices, the Evaluation Department must ensure that the information is passed on to relevant departments for appropriate action.

Code for Danida staff (Embassy and Headquarters staff)

Danida staff has a complex role in the evaluation process: they can request an evaluation, be the object of an evaluation, function as key resource persons during the evaluation, and/or be users of the results. They have a key role in identifying relevant evaluation topics and in assuring the usefulness of evaluation findings for the learning processes of Danida and its partners. They facilitate evaluations, and they contribute to quality assurance of the reports by pointing out factual errors and inaccuracies. Danida staff, thus, has the following responsibilities:

- When facilitating evaluations by providing contacts, references, information about activities and logistical support to the consultants, Danida staff is expected to respect the integrity of the evaluation team

in making its own decisions about where to go and whom to see. Danida staff may certainly provide comments or background information on suggested sites for field visits and persons to be interviewed, but the final decision to visit the site rests with the evaluation team.

- Danida staff should assist with the identification of relevant documents, even if the material has not been specifically requested.
- When contributing to quality assurance of the report, Danida staff should observe the right of consultants to make conclusions and recommendations, which may not be shared by Danida. It is the responsibility of Danida staff to contribute to a fruitful exchange of information and views and to avoid any forms of communication, which could be perceived by consultants as undue pressure or even threats to produce certain conclusions.
- Reservations regarding the competence of evaluation team members, the quality of the fieldwork, the quality of analyses, etc. should be reported immediately to the Evaluation Department. If such reservations have not been expressed in due time, Danida staff should refrain from criticising the quality of the evaluation after its publication.
- After the draft report has been produced, all contact between consultants and Danida staff should go through the Evaluation Department. If consultants meet with Danida staff, the Evaluation Department should be present, and all correspondence between consultants and Danida staff should be copied to the Evaluation Department.

Code for other parties responsible for the activities under evaluation

With the increasing alignment of Danish development cooperation with partner management structures and with the emphasis put on partner ownership, actors in partner countries such as line ministries, private sector actors, and civil society organisations are key parties to the evaluation process. Other parties responsible for the activities under evaluation may include organisations in Denmark (e.g. Danish NGOs), multilateral organisations and other international institutions supported by Danida. Similar to Danida staff, these stakeholders have the complex role of being both the object of an evaluation, key resource persons during the evalua-

tion, and/or users of the results. They have a key role in assuring the usefulness of evaluation findings for their own learning processes. They facilitate evaluations, and they provide quality assurance of the reports by pointing out factual errors and inaccuracies. Their responsibilities are as follows:

• When facilitating evaluations by providing contacts, references, information about activities and logistical support to the consultants, stakeholders are expected to respect the integrity of the evaluation team in making its own decisions about where to go and whom to see. Stakeholders may certainly provide comments or background information on suggested sites for field visits and persons to be interviewed, but the final decision to visit the site should be made by the evaluation team.

• Stakeholders should assist with the identification of relevant documents, even if the material has not been specifically requested.

• When providing quality assurance of the report, stakeholders should observe the right of consultants to make conclusions and recommendations, which may not be shared by their organisation. It is the responsibility of the stakeholders to contribute to a fruitful exchange of information and views and to avoid any forms of communication, which could be perceived by consultants as undue pressure or even threats to produce certain conclusions.

• Reservations regarding the competence of evaluation team members, the quality of the fieldwork, the quality of analyses, etc. should be reported immediately to the Evaluation Department. If such reservations have not been expressed in due time, stakeholders should refrain from criticising the quality of the evaluation after its publication.

• After the draft report has been produced, all contact between consultants and stakeholders should go through the Evaluation Department. If consultants meet with stakeholders, the Evaluation Department should be present, and all correspondence between consultants and stakeholders should be copied to the Evaluation Department.

CHAPTER 4
EVALUATION QUESTIONS

Evaluation criteria

The OECD/DAC definition of evaluation has been adopted by Danida and all major development agencies internationally. The definition contains five evaluation criteria that should be used in assessing development interventions: relevance, efficiency, effectiveness, impact and sustainability.

These are general criteria that should be used as a basis for developing evaluative questions through the full range of evaluations topics, i.e. from single intervention through to thematic, and ways of conducting the evaluation, e.g. joint evaluation.

Taken together, these five criteria should provide the decision-maker with the essential information and clues to understand the situation and determine what should be done next.

To the extent that specific evaluations have specific purposes, that there is no one right way to conduct an evaluation and that these criteria are interdependent and not mutually exclusive, their relative meaningfulness for a specific evaluation should be assessed and trade-offs discussed in each case to ensure that key questions are addressed and to avoid unnecessary effort and expense.

In addition, use of the five criteria does not exclude that other criteria might be used as well to better focus the evaluation on specific characteristics of the intervention and its context.

The criteria for the evaluation of humanitarian assistance are a case in point: because of some unique features of humanitarian intervention, the Active Learning Network for Accountability and Performance in Huma-

nitarian Action, ALNAP www.alnap.org, has introduced three additional
evaluation criteria: connectedness, coherence and coverage.[9]

Evaluation criteria

Relevance   The extent to which the objectives of a development interven-
tion are consistent with beneficiaries' requirement, country
needs, global priorities and partners' and donors' policies.

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

Effectiveness   The extent to which the development intervention's objectives
were achieved, or are expected to be achieved, taking into
account their relative importance.

Impacts   The 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 long-term benefits. The resilience to risk of the
net benefit flows over time.

9)   *Beck, T. (2006). Evaluating humanitarian action using the OECD-DAC criteria:
An ALNAP guide for humanitarian agencies. London, UK: Overseas Development
Institute.*

Additional criteria for evaluation of humanitarian action

Connectedness The need to ensure that activities of a short-term emergency nature are carried out in a context that takes longer-term and interconnected problems into account.

Coherence 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 rights considerations.

Coverage The need to reach major population groups facing life-threatening suffering wherever they are.

Relevance

Relevance is a measure of the extent to which development interventions meet population needs and country priorities, and are consistent with donor policies.

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 new crops, farmers' responses to the various programme initiatives, etc.

A change in society's policies or priorities could imply that the development interventions are now accorded lower priority, or lose some of their rationale. Once an endemic disease has been eradicated, for instance, it could mean there is no longer any need for a special health programme.

In other words, relevance is basically a question of usefulness; in turn, the assessment of relevance leads to higher level decisions as to whether the development activities in question ought to be terminated or allowed to continue. And, if the latter is the case, what changes ought to be made, and in what direction? Are the agreed objectives still valid, and do they represent sufficient rationale for continuing the activities?

These questions should be addressed at various levels with reference to the partner country:

- At the higher level it concerns the relationship between the development activities and the development policy of the partner country, as well as whether the activities are in keeping with the priorities of the donor country.
- At the middle level it is a question of how development activities fit into a larger context, e.g. in relation to other development interventions and development efforts within a larger programme or sector.
- At the lower level it is a question of whether the development activities are directed towards areas accorded high priority by the affected parties.

Efficiency

Efficiency is a measure of the relationship between outputs, i.e. the products or services of an intervention, and inputs, i.e. the resources that it uses.

An output is a measure of effort; it is the immediate observable result of intervention processes over which the managers of the intervention, i.e. the implementers, have some measure of control. An intervention can be thought of as efficient if it uses the least costly resources that are appropriate and available to achieve the desired outputs, i.e. deliverables, in terms of quantity and quality.

The quality of the inputs and the outputs is an important consideration in assessing efficiency: the most economical resource is not necessarily the

most appropriate and the trade-offs between the quantity of outputs and their quality are a key factor of overall performance.

Furthermore, assessing the efficiency of an intervention generally requires comparing alternative approaches to achieving the same outputs and this will be easier for some types of intervention that for others.

In practise, the extent to which intervention activities are standardised or not, i.e. the factors of production are well known or not, usually determines how efficiency is measured and assessed.

In a road building project for example, where the methods of construction are fairly well established, a typical measure of efficiency would be the cost per km per class of road. As well, because other projects and jurisdictions are also likely to use that same measure of efficiency, among others, the bases for comparison and assessment, or benchmarks, are readily available in most cases.

On the other hand, a national initiative on women's rights for example is not standardised across countries. In such cases, relevant measures of efficiency typically address waste in the process, either at the level of inputs, i.e. economy – obtaining appropriate resources at least cost or fair market value, or at the level of process, i.e. duplication-triplication – etc. of activities, conflicting processes, throughputs that do not link to outputs. As well, good practices, i.e. lessons learned from similar endeavours, can be used as benchmarks for assessing efficiency.

Some examples of useful and practical criteria for assessing the efficiency of a programme or a project are:

- Appropriate resources acquired with due regard for economy
- Activities carried out as simply as possible
- Decisions made as close to where the products or services are delivered
- Overhead as low as possible
- Duplication or conflicts addressed and resolved
- Deliverables achieved on time and on budget.

Effectiveness

Effectiveness is a measure of the extent to which the intervention's intended outcomes, i.e. its specific objectives – intermediate results – have been achieved.

Explicitly, effectiveness is the relationship between the intervention's outputs, i.e. its products or services – its immediate results – and its outcomes, meaning usually the intended benefits for a particular target group of beneficiaries.

As such, an intervention is considered effective when its outputs produce the desired outcomes; it is efficient when it uses resources appropriately and economically to produce the desired outputs.

For example, a teaching programme is considered effective if students learn, i.e. acquire intended knowledge, skills and abilities; it is considered efficient if it provides instruction, i.e. teaching time and materials, economically and of quality.

An efficient intervention is not necessarily effective. Teaching may be provided economically and efficiently, but if it is not of good quality, e.g. appropriate to the needs and interests the students, intended learning outcomes will not be achieved, i.e. it will not be effective.

Evaluating the effectiveness of an intervention involves three steps:

1. Measuring for change in the observed outcome, e.g. did the students learn something;
2. Attributing the change in the observed outcome to the intervention, did the students learn something because of the teaching;
3. Judging the value of the teaching to the learning, e.g. by using comparisons such as targets, benchmarks, similar interventions, the assessments of teachers, students, others, etc.

Interventions have no control per se over outcomes; at best, a program strives to produce those outputs that have the greatest likelihood of producing the intended outcomes.

As such, an intervention's effectiveness is driven primarily by two things: its design and its implementation, i.e. its management.

In many cases, attribution, i.e. internal validity, is typically the central challenge to assessing effectiveness of interventions, i.e. how and to what extent can observed changes in outcomes be attributed validly to interventions.

Where the internal validity of the programme is well established, e.g. an immunisation programme, attribution of outcomes, e.g. beneficiaries protected from disease because they have been vaccinated against that disease, can be fairly straightforward.

However, in the case of many development interventions, internal validity is not well established and attribution can become a significant challenge. For example, attributing validly a change in the incidence of human rights violations in a country to an intervention or set of interventions might be difficult for most evaluations.

Generally, the more the evaluand, i.e. what is being evaluated, is large and complex, the more attribution is likely to be difficult.

The reality of methodological and resource constraints in carrying out practical evaluation means that often attribution is expressed in terms of likelihood rather than proof, and that ultimately the test of validity is credibility.

Other challenges to assessing effectiveness are typically:

- Non-existent or poorly defined objectives, e.g. intended outcomes are not stated as measurable change over time in target groups
- Unrealistic and/or conflicting objectives
- Lack of targets or measures of success.

Impact

Impact is a measure of all significant effects of the development intervention, positive or negative, expected or unforeseen, on its beneficiaries and other affected parties.

Whereas effectiveness focuses on the intended outcomes of an intervention, impact is a measure of the broader consequences of the intervention such as economic, social, political, technical or environmental effects; locally, regionally, or at the national level; on the target group and other directly or indirectly affected parties.

For example an HIV/AIDS prevention and treatment programme targeting vulnerable groups could have broader effects both positive, such as a reduction in the incidence of tuberculosis on other groups, and negative, such as a reduction of funding to malaria prevention. Effects may also be economic in nature, e.g. size of the workforce, political, e.g. state budget allocation, and so on.

A broad assessment of impact is essential in a comprehensive evaluation, however there are two central challenges to assessing impact: boundary judgment, i.e. deciding what effects to select for consideration, and attribution, i.e. what effect is due to what.

Because, on the one hand, effects can be numerous and varied, and on the other they are typically the result of complex interactions, assessing impact is difficult in most circumstances.

As is the case for effectiveness, the assessment of impact poses a particular challenge with regard to attribution; in most cases, it is difficult to attribute rigorously broad effects on different groups and at different levels over time to a specific intervention or set of interventions.

Systems theory approaches typically provide more appropriate and useful tools for dealing with complex adaptive systems, e.g. societies, than simple linear causal approaches.

As well, a useful principle for dealing pragmatically with the issue of selection of effects, levels and groups for the evaluation, is making choices consistent with the intended use of the evaluative information.

In the final analysis, usually and at best, evaluations estimate impact on probability-based inferences derived from assumptions of simplified cause and effect relationships.

In the case of an impact evaluation, i.e. one in which measurement and assessment of impact is given priority, one must estimate the counterfactual – that means what would have happened if the intervention had not taken place. This can be done by choosing a control or comparison group – a group of individuals, households, etc., that are identical to the project group except for participation in the project or programme.

As well, baseline data, i.e. information about the state of groups before the intervention, is useful in order to measure differences after the intervention has taken place.

These methodological requirements pose particular challenges to the conduct of impact evaluation.

Sustainability

Sustainability is a measure of whether the benefits of a development intervention are likely to continue after external support has been completed.

While the four preceding criteria concern specific development interventions, the assessment of sustainability addresses the effects of the development process itself over the long term.

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 used and provide benefits 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.

Sustainability is in many ways a higher level test of whether or not the development intervention has been a success. Far too many development initiatives tend to fail once the implementation phase is over, because either the target group or the responsible parties do not have the means or sufficient motivation to provide the resources needed for the activities to go further.

Sustainability is becoming an increasingly central theme in evaluation work since many development agencies are putting greater emphasis on long term perspectives and on lasting improvements.

As a result, capacity-development of communities and organisations is a common objective of development interventions, consistent with the overall goal of promoting increased autonomy and self-reliance of partner countries for the provision of public services.

Useful questions for assessing sustainability address the extent to which capacity has been successfully developed, e.g. through participation, empowerment, ownership, local resources are available and sustained political support exists.

As well, the sustainability of development interventions depend to a large extent on whether the positive impacts justify the required investments and whether the community values the benefits sufficiently to devote scarce resources to generating them.

Because sustainability is concerned with what happens after development activities are completed they are measured ideally some years afterwards. It is difficult to provide a reliable assessment of sustainability while activities are still underway, or immediately afterwards. In such cases, the assessment is based on projections of future developments based on available knowledge about the intervention and the capacity of involved

parties do deal with changing contexts. It requires an analysis of the contextual setting – its capabilities and restraints – and future scenarios.

Experiences of donor agencies conclude that a development intervention's sustainability hinges mainly on seven areas, also called sustainability factors.

These seven factors should be taken into account all along the implementation cycle, and should be used as a checklist during the evaluation to identify relevant questions:

Policy support measures
The recipient's commitment is one of the most commonly identified factors affecting success of development interventions. Commitment is expressed in terms of agreement on objectives, the scope of support to responsible organisations and the willingness to provide financial and personnel resources. Country commitment is also shaped by perceptions of mutuality of interests versus perceptions of predominantly donor-driven interests.

Choice of technology
The partner country's financial and institutional capabilities are crucial determinants for the technology chosen and that the technology is accepted with mechanisms for its maintenance and renewal. Evaluation teams should consider the effects of technology in society and the costs of providing and maintaining the technology versus the benefits generated.

Environmental matters
The importance of environmental considerations is now widely recognised. Although environmental effects may be negligible seen in a narrow, short-term perspective, the broader effects may be significant in a long-term perspective. Evaluations will frequently have to look specifically at environmental policy, incentives and regulatory measures, the interests of different stakeholders, and the effects of development interventions.

Socio-cultural aspects
Social and cultural factors influence the adaptability and relevance of
various development activities. They also influence motivation among
the target group members and hence whether individuals and groups will
participate and accept responsibilities in the development process.

Sustainability factors

Policy support measures	Policies, priorities, and specific commitments of the recipient supporting the chances of success.
Choice of technology	Choice and adaptation of technology appropriate to existing conditions.
Environmental matters	Exploitation, management and development of resources. Protection of the environment.
Socio-cultural aspects	Socio-cultural integration. Impact on various groups (gender, ethnic, religious, etc.).
Institutional aspects	Institutional and organisational capacity and distribution of responsibilities between existing bodies.
Economic and financial aspects	Economic viability and financial sustainability.
External factors	Political stability, economic crises and shocks, overall level of development, balance of payments status and natural disasters.

Development interventions that are consistent with local traditions or do
not assume major changes in behaviour patterns stand a better chance of
success. Danida requires special attention to the roles of both women and
men in implementing development interventions, particularly their

access to means of production and support services, as well as their rights and benefits.

Institutional aspects
The strength of institutions and the capacity of organisations are the most important factors in the success of development interventions. Current trends to change the division of roles between public and private organisations may raise the issue of how development interventions affect the co-operation and co-ordination of participating bodies. At the more detailed level, assessment may include considerations of managerial leadership, administrative systems and the involvement of beneficiaries.

Economic and financial aspects
Evaluations should focus essentially on three aspects of economic and financial performance. Firstly, the cost effectiveness of the intervention strategy. Secondly, the economic and financial benefits of investments as compared with the funds and resources spent. Finally, the financial sustainability of operations in the future to explore whether funds are or will be sufficient to cover future operations, maintenance and depreciation of investments.

External factors
Development assistance takes place in the context of political, economic and cultural environments that are beyond its control yet can influence it significantly. Factors such as political stability, economic crises and shocks, overall level of development, balance of payments status and natural disasters can have a determining influence on the sustainability of development interventions.

Additional criteria for evaluating humanitarian action
(Adapted from Beck, T. 2006)

Evaluation of humanitarian action, in response to natural disasters and to conflicts, has been the subject some attention over the last few years with a view to improving its quality. References for evaluators and for evaluation managers, such as "Evaluating Humanitarian Assistance Programmes

in Complex Emergencies",[10] "Guidance for Evaluating Humanitarian Assistance in Complex Emergencies"[11] and "Evaluating Humanitarian Action using the OECD/DAC Criteria", are key to furthering the understanding and the quality of evaluation practise in this area.

Because of some of the distinct characteristics of humanitarian action, ALNAP proposes three additional criteria (to OECD/DAC's five): connectedness, coherence and coverage.

Evaluation of humanitarian action (EHA) is defined by ALNAP as "a systematic and impartial examination of humanitarian action intended to draw lessons to improve policy and practice and enhance accountability."

That they are undertaken usually during severe disruptions, often prolonged in the case of complex emergencies, give EHAs some distinct characteristics that can make access to data and information difficult, for example:

• Polarisation of perspectives in conflict situations that reduces the "space" for fair and balanced views;
• High turnover of staff working in humanitarian action that affects "organisational memory";
• Reactive and quick implementation of humanitarian action that affects planning and the identification of performance measures;
• The context of humanitarian action may be disordered, with rapid changes in circumstances, invalidating assumptions that may apply to usual social and physical conditions.

10) Hallam, A. (1998), *Evaluating Humanitarian Assistance Programmes in Complex Emergencies. London, UK: ODI. Good Practice Review 7.*
12) OECD (1999), *Guidance for Evaluating Humanitarian Assistance in Complex Emergencies. Paris.*

Connectedness

Connectedness is defined as "The need to ensure that activities of a short-term emergency nature are carried out in a context that takes longer-term and interconnected problems into account."

Connectedness is adapted from the concept of sustainability and is a measure of the relationship between humanitarian interventions and longer term goals, in particular the linkages between humanitarian action, recovery and development.

Some of the issues to consider when addressing connectedness are relative expenditure on relief and recovery, the nature of partnerships, in particular between international and national NGOs, and the extent to which local capacity is supported and developed.

Coherence

Coherence is defined as "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 rights considerations."

The focus of this criterion is on the extent to which the policies of different actors, e.g. military, civilian, political, were complementary or contradictory. Policies may be of any type such as promoting gender equality, participation or environmental protection.

Coherence may be the most difficult criterion to evaluate, in part because it is often confused with coordination. The evaluation of coherence focuses mainly on the policy level while that of coordination more on operational issues.

Evaluation of coherence is important where there are many actors and increased risk of conflicting mandates and interests. Questions on the

degree of coherence observed, or otherwise, are also important and useful to address.

Coverage

Coverage is defined as "The need to reach major population groups facing life-threatening suffering wherever they are." The key questions this criterion generates are who was supported by humanitarian action and why.

Evaluation of coverage usually takes place at three levels: international, national or regional, and local, and the ecaluation should consider whether assistance was provided proportionally according to the need at each level.

Whether protection needs have been met is an important question to address and so are issues of inclusion and exclusion bias at regional and local levels.

Political factors often determine coverage so that analysing them is key to understanding the nature and extent of coverage of groups, including issues of protection and humanitarian space.

Finally, equity questions are central to the assessment of coverage and are addressed through geographical analysis and the organisation of data by socioeconomic categories such as gender, socioeconomic groupings, ethnicity, age and ability.

CHAPTER 5
EVALUATION TEAM TASKS

This chapter explains the expectations Danida and its Evaluation Department have of the external evaluation team with regard to the content and quality of the evaluation it is responsible for.

The information in this chapter intends to provide the external evaluation team with a clear understanding of the framework of principles, standards, roles, responsibilities and practices that Danida applies both to the methodologies and the management of its evaluations.

As such, the information is presented for each of the three main phases of the evaluation process that the external consultant team is responsible for: inception, fieldwork and reporting.

Inception: Planning the evaluation

The purpose of the inception phase is for the evaluation team to prepare a detailed operational plan, i.e. the inception report, for the next phases of the evaluation: fieldwork and reporting.

Proper planning is essential to identifying those activities required to provide well-supported answers to the evaluation questions and to avoiding other unnecessary activities and related expenditures of time, effort and money.

The planning phase provides the evaluation team with the opportunity, and responsibility, to discuss methodological specificities, fieldwork activities and reporting strategy with, and where required obtain approval from, the Evaluation Department, and as well to consult with other stakeholders.

The inception report should present:

- An overall logic model of the intervention (the evaluand), depicting the linkages between resources (inputs), intervention activities (processes), intervention results (outputs or deliverables), intended outcomes (intervention objectives), overall impacts, and their relationships in terms of the criteria of relevance, efficiency, effectiveness and impact; an explanation of how the sustainability criterion is defined and operationalised.
- The methodology: design, approach, sufficiency and appropriateness of evidence, data collection strategy and methods, analytical framework and reporting outline.
- The hierarchy of evaluation questions starting from the general ones that are presented in the Terms of References through to the specific ones that will produce data and information.
- For each specific question the basis for assessment, i.e. indicator of minimum acceptable performance.
- A matrix indicating for each specific question the nature and source of evidence.
- A schedule of activities.
- A communication and consultation plan (with stakeholders).
- In the case of evaluations with complex evaluation team organisation and logistics, e.g. joint evaluations, a systematic management plan that addresses key issues of management, coordination, authorities, responsibilities, etc.

It is expected that the external evaluation team presents, applies and documents its quality control and assurance process from the beginning of its work, i.e. at the start of the inception phase.

It should address key questions of methodology, e.g. reliability and validity issues, as well as of project management, e.g. security and confidentiality of data and information.

As well, as part of its quality control and assurance process, it is expected that the evaluation team defines, documents and presents to the Evaluation Department the structure and organisation of its evaluation file, i.e.

the way in which all the information relating to the evaluation is organised and kept.

The inception phase and its attendant tasks, roles and responsibilities are organised as follows:

Inception: Steps, roles and responsibilities

Evaluation step	EVAL
1. Briefing of consultants	a. Clarifies Terms of Reference b. Clarifies role of consultant c. Presents expectations to inception report d. Adjusts time-table e. Provides background materials f. Provides writing instruction and Code of Conduct
2. Desk study	a. Facilitates study by providing relevant documents and contacts b. Ensures involvement of relevant MFA staff and other stakeholders
3. Preliminary field visit	Facilitates visit by providing contacts, information and logistical support.
4. Inception report, incl. evaluation matrix, proposed methodology, proposed sources of information, work plan and communication strategy	a. Ensures involvement of relevant MFA staff and other stakeholders. b. Provides consolidated comments on the report c. Approves the report

Other MFA staff	Consultant	Others
	Raise uncertainties concerning interpretation of assignment and role of consultants for clarification by the Evaluation Department.	
Embassy, BFT and other departments: Facilitate the study by providing relevant documents and contacts.		*Partners and possibly other donors:* Facilitate the study by providing relevant documents and contacts.
Embassy: Facilitates visit by providing contacts, information and logistical support.	Undertakes the visit.	*Partners*: Facilitate visit by providing contacts, information and logistical support.
Embassy and BFT: Comment on the report.	Prepares the report.	*Partners, reference group and possibly other donors:* Comment on the report.

Inception: Methodological considerations

The Terms of References of the evaluation stipulate usually, and to varying degrees, the methodology to be employed as well as the organisation of the evaluation activities, e.g. budget, schedule and travel.

However, as a result of discussions with the Evaluation Department and other stakeholders, desk study, preliminary field visit, etc., and its own evaluation expertise, the evaluation team is expected to address thoroughly key issues and questions related to methodology and management, in order to propose defensible activities, changes and/or alternatives, that will improve the quality of the evaluation.

The DAC Evaluation Quality Standards require that sound methodology be used and explained in the evaluation report.

The purpose of methodology, and the basis on which its soundness is assessed, is to produce reliable data that allow for valid evaluative judgments that are useful for learning and making decisions.

While evaluation methodology is grounded firmly in traditional social science approaches, evaluation is more than the application of social science methods to study social problems.

To make judgments and facilitate decisions the evaluation team is required to engage with stakeholders and users, in order to integrate empirical evidence with standards and values to reach evaluative conclusions.

The methods used for one evaluation can differ greatly from the next, depending on what the evaluation intends to accomplish or answer and the theories and preferences that are used.

For example, if the priority is to make sure audiences use the evaluation, a utilisation-focused approach might be employed; if the priority is answering as unequivocally as possible "what works," a randomized trial may be chosen; if the priority is engaging stakeholders and building evaluation capacity, an empowerment or participatory approach may be privileged.

Increasingly evaluators use "mixed methods" to optimise the configuration of components that make up the evaluation's methodological framework as the evaluation's methodology encompasses choice, and allows for mixing of:[12]

- Design, i.e. experimental, quasi-experimental, non-experimental;
- Approaches, e.g. theory-based, goal free, constructivist, empowerment, utilisation-focused, etc.);
- Qualitative and quantitative methods;
- Tools and techniques (e.g. logic modelling, strategic planning, concept mapping, etc.).

Methodology is also the basis on which validity and reliability of the evaluation is determined. As such, it is expected that the evaluation team develop methodology with due regard to minimising threats to validity and addressing issues of reliability.[13]

Validity
Validity is a measure of the extent to which, taken together, the evaluation's design, data collection methods and analyses provide a reasonable basis for conclusions about the evaluation's questions.

Four types of validity are important to consider in planning and carrying out evaluations:

- External validity – the extent to which issues related to the generalisation of conclusions are addressed; this is a particularly significant issue whenever sampling is used.
- Construct validity – the extent to which inferences can be made legitimately from the operationalisations, e.g. the specific questions and indicators, to the constructs on which they are based, i.e. the general evaluation criteria.

12) Adapted from Weiss, H. B. (Summer 2005). The Evaluation Exchange: Editorial. Harvard Family Research Project.
13) Adapted from Trochim, W. M. K. (2002). Research Knowledge Database.

- Internal validity – whether observed changes can be attributed to the intervention, i.e. the cause and not to other possible causes (sometimes described as alternative explanations, rival hypotheses or counterfactuals, for outcomes and impacts).
- Conclusion validity – the degree to which conclusions reached about relationships in the data are reasonable.

Reliability
Reliability is a measure of the quality of measurement; information is reliable if the measurement procedure yields the same results if applied repeatedly.

Reliability is a key factor for the quality of the evaluation and, as such, the evaluation team is expected to incorporate into its methodology ways of estimating the reliability of the data it gathers.

One of the most commonly used techniques for approximating the reliability of data is triangulation, applying the same measurement procedure to three or more different sources to obtain data that can be compared for similarity.

This type of reliability estimate is called internal consistency reliability; other forms of reliability estimates are:

- Parallel-forms reliability, e.g. where triangulation is applied to three different categories of data: verbal, documentary and observation.
- Test-retest reliability, used to assess the consistency of a measure from one time to another, e.g. where the same household survey is administered at periodic intervals.
- Inter-Rater or Inter-Observer Reliability, used to assess the degree to which different raters/observers give consistent estimates of the same phenomenon, e.g. an assessment of the fairness and freedom of elections by different observers.

The evaluation team is expected to consider the relationship, i.e. the trade-offs, between issues of validity and reliability, as part of the development of the methodology.

For example, although a design that incorporates randomised control groups might be the most appropriate for dealing with issues of validity, availability and reliability of data might preclude such an approach (comparison group data, baseline data, surveys that could be substituted for baseline data, measures that are valid and reliable and so on). As well, the reality of the contexts in which most development interventions take place is such that even if such data are available they should be treated with caution.

As a result of these kinds of considerations an approach that makes use predominantly of qualitative data through interviews, observations and documentary reviews might often be more appropriate. What is expected of the evaluation team is that it undertake, and document in the evaluation file, these considerations systematically.

Sufficient and appropriate evidence
When making choices about the amount and nature of data to gather, it is expected that the evaluation team will collect only the information required to answer the evaluation questions.

Sufficiency has to do with the amount of information required to provide persuasive support for the contents of the evaluation report, i.e. will the collective weight of the evidence be sufficient to persuade a reasonable person that the observations and conclusions are valid, and the recommendations appropriate.

Some of the factors to consider when judging sufficiency are:

- The quality of the data, i.e. its relevance, reliability and validity;
- The significance of the finding and conclusion the data are intended to support, e.g. how important is it?
- How much assurance is intended, e.g. is the evaluation important for accountability purposes?
- What is the risk of making an incorrect observation or reaching an invalid conclusion?
- What is the cost of obtaining additional information in relation to its additional benefits, i.e. in terms of support for observations and conclusions?

Appropriateness of data includes questions of reliability and validity, but as well of relevance, i.e. the extent to which information bears a clear and logical relationship to the evaluation criteria and questions.

General categories of data are verbal, documentary and observational; as a general rule of thumb, observations are considered the most robust type of data followed by documentary and verbal. To the extent possible and appropriate, more than one type of data should be used to answer evaluation questions.

Overall, planning the conduct of the evaluation, and in particular the fieldwork, is an exercise in professional judgment, i.e. how best to answer the evaluation questions in light of methodological and resource constraints. (See "Shoestring Evaluation: Designing Impact Evaluations under Time, Budget and Data Constraints" by Michael Bamberger et al.)

Fieldwork: Collecting information

The purpose of fieldworks is to collect systematically and in accordance with the plan, the information required to support the formulation of conclusions, i.e. answers to the evaluation questions.

Fieldwork includes applying data collection methods, coordinating with the Evaluation Department, stakeholders and users, making adjustments to data collection as needed in response to constraints, to maintain and where possible improve the quality of the evaluation, undertaking analysis by formulating findings of fact based on the data and clearing these with concerned parties.

Data collection and choice of methods should take into account the specific challenges that may be posed by different languages in different contexts. As well, cultural sensitivities should be considered and respected both in the choice of data collection methods and in the way information is gathered.

Examples of data collection methods

1. Literature search — Economic and efficient way of obtaining information. Difficult to assess validity and reliability on secondary data.

2. Key informant interviews — Flexible, in-depth approach. Easy to implement. Risk of biased presentation/interpretation from informants/interviewer.

3. Direct measurement — Registration of quantifiable or classifiable data by means of analytical instrument. Precise, reliable and often requiring few resources. Registers only facts, not explanations.

4. Direct observation — Involves inspection, field visits, observation to understand processes, infrastructure/services and their utilisation. Dependent on observer's understanding and interpretation.

5. Group interviews — Low-cost, efficient. Direct contact with those affected. Susceptible to manipulation and less suitable for sensitive issues.

6. Informal survey — Involves quantitative surveys of small samples. Reasonable and rapid. Risk of sampling errors/biases. Less suited for generalisation.

7. Case studies — 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.

8. Observation — In-depth observations over an extended period of time, participatory or non-participatory. Well-suited for understanding processes but with limited potential for generalisation.

9. Formal survey — Oral interviews or written questionnaires in a representative sample of respondents. Data collection is demanding but often produces reliable information.

10. Story telling/ collection — Obtaining participants' and communities' experiences of change by collating their observations and stories.

Because evaluations often produce controversial results and may be subject to disagreements, the choice of methods of data collection and analysis may be a focus of criticisms.

It is expected that the evaluation team take into account, and communicate in its report, the strengths and weaknesses of the various methods employed, bearing in mind the considerations discussed during inception.

Fieldwork: Analysis

The purpose of analysis is to transform data into credible evidence about the development intervention and its performance. Typically, the analytical process involves three steps:

1. Organising the data for analysis, i.e. data preparation
2. Describing the data, e.g. generating findings of fact
3. Interpreting the data, e.g. assessing the findings against criteria

Fieldwork: Steps, roles and responsibilities

Evaluation step	EVAL
1. **Fieldwork, incl. debriefing in partner country**	Possible participation during initial part of fieldwork.
2. **Debriefing in Denmark**	a. Participates in debriefing meeting b. Approves debriefing report

Preparation – Cleaning and classifying the data
Prior to and during fieldwork the instruments are used to collect information in the form of raw data, time-series, individual statements, etc. The first step in the analytical process involves preparing the data so that they meets standards of quality, i.e. validity and reliability, and that they are organised according to the evaluation questions and criteria they address.

Findings – What we saw, read and heard
The second step is to aggregate the data to constitute findings that are relevant to the evaluation questions at hand. Although judgment is involved in their formulation and selection, findings describe facts and do not constitute nor contain evaluative judgments in of themselves.

For example, as a result of collecting data on primary education a finding might be that the number of girls enrolled during a given school year was X. This might be a relevant finding to addressing the question of gender balance in the school system and contributing to forming an overall

| Other MFA staff	Consultant	Others
Embassy: a. Facilitates fieldwork b. Participates in debriefing based on debriefing report	a. Undertakes fieldwork b. Prepares and presents debriefing report	*Partners*: a. Facilitate fieldwork b. Participate in debriefing based on debriefing report
BFT and other relevant departments: Participate in debriefing meeting.	Presents debriefing report.	*Reference group:* Participates in debriefing meeting.

assessment of a national education policy for example. However, the finding in of itself is neither good nor bad, it is only a fact.

Conclusions – What we think

The third step in the analytical process is to interpret the findings, i.e. formulate conclusions, meaning usually in the case of evaluation to judge the findings in relation to a reasonable expectation, e.g. a standard, criterion, benchmark, target, indicator, good practice, etc.

For example, a finding that 40 per cent of girls enrolled in primary school complete their education over a given period, may be interpreted as "good" if the target was 40 per cent, "excellent" if the target was 30 per cent, "poor" if it has remained unchanged over the period and the policy was to increase the graduation rate of girls, etc.
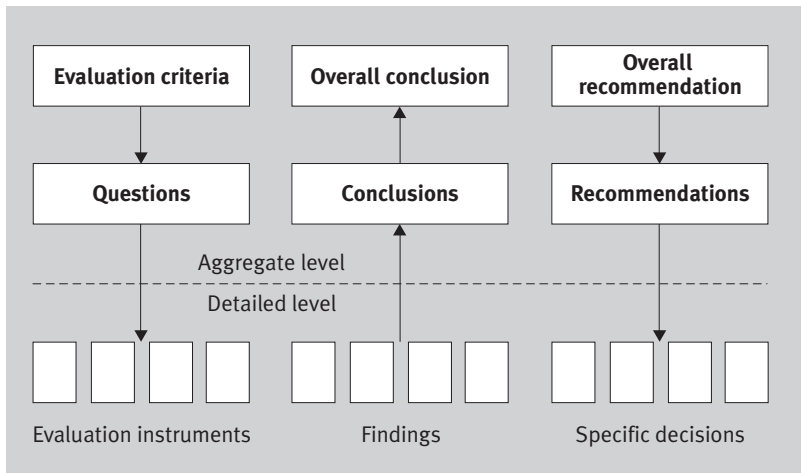
The distinction between findings and conclusions, and its application to the processes of analysis, of clearance and of reporting are a cornerstone of Danida's approach to ensuring the quality of its evaluation and constitute a fundamental expectation of the evaluation team's work.

The ability for the evaluation team, the Evaluation Department as well as all other interested parties, to clearly distinguish between the agreed to facts and their interpretive conclusions, reduce the risk of general disagreement, focus critiques on specific aspects of evaluative judgment so that these can be dealt with, e.g. ultimately "agree to disagree", contribute to the transparency and objectivity of the evaluation process as well as the credibility of the report.

Recommendations – What we suggest
The evaluation team is then expected to translate the conclusions into main recommendations which can be considered by management and subsequently translated into more specific and detailed decisions.

The evaluation team is usually not 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.

10.3 Distinction between conclusions, recommendations and lessons learned

Evaluation reports must distinguish clearly between findings, conclusions and recommendations. The evaluation presents conclusions, recommendations and lessons learned separately and with a clear logical distinction between them. Conclusions are substantiated by findings and analysis. Recommendations and lessons learned follow logically from the conclusions.

(DAC Evaluation Quality Standards)

Reporting: Preparation

Reporting the evaluation starts with its Terms of Reference. The evaluation team is bound by contract to ensure that the issues in the Terms of Reference are adequately addressed in the report.

The team leader normally drafts a table of contents at the inception stage of the evaluation, based on Danida's requirements for reporting, the Terms of Reference and the introductory discussions with Danida. This also helps to distribute responsibilities between the members in the evaluation team.

The draft table of contents is a convenient means to organise information as the work proceeds. This includes information from initial desk studies, special studies, etc. Normally, the team prepares the outline of the draft report before departing for the fieldwork. This helps to focus the field-work to collect missing information, verify information, and draw and discuss conclusions and recommendations.

Usually, the team splits after the fieldwork. As a result, team members are normally obliged to complete their inputs to the draft report while in the field. This helps the team leader present a draft summary for discussion with country officials and the Embassy before leaving.

The draft of the report is reviewed by the Evaluation Department and Danida's technical advisors and operational staff before it is distributed for wider consultation. They assess the quality of the report and submit comments regarding factual information, conclusions, and recommen-dations to be considered by the team before the final draft version is prepared.

This is usually done in a meeting or workshop with all relevant stake-holders where the evaluation team is debriefed. Danida then obtains comments on the final draft report from all partners and makes them available to the team. Such comments may be annexed to the report.

The team leader is responsible for submitting the final report and a summary (four pages). The structure and content of the report is dis-cussed in the following sections. The final report should be in accordance with Danida's layout for evaluation reports. The layout is available from the Evaluation Department.

Evaluation report

The evaluation team has the final responsibility for the contents of the report. Any disagreements among the evaluation team or between the evaluation team, Danida, and relevant partners that are significant to conclusions and recommendations must be reflected in the report, either in the form of comments in the text, footnotes or as a special section.

Evaluation reports must be brief and concise, and the presentation must be clear and adjusted to the target group. The normal language is English, and where relevant reports are translated into French, Spanish or Portuguese.

For every evaluation a brief summary in Danish, English and other relevant languages must be compiled with a view to publication. The summary must cover the most important observations and conclusions of the evaluation. The presentation must be made in language accessible to non-professionals.

If an evaluation is thought to be of interest to a broader audience Danish and local language versions will be produced (possibly abbreviated) and edited to communicate effectively to readers that are not professional specialists.

Danida Evaluation Policy

Reporting: Steps, roles and responsibilities

Evaluation step	EVAL
1. **Final draft of sub-reports (country or project reports)**	a. Ensures involvement of relevant MFA staff and other stakeholders b. Provides consolidated comments on the reports c. Approves the reports
2. **Final draft of synthesis report**	a. Ensures involvement of relevant MFA staff and other stakeholders b. Provides consolidated comments on the report c. Approves the report

Reporting: Structure

Because the focus of evaluations varies between sector programme, country evaluations, etc., there is no common format for reporting. However, Danida expects the evaluation team to report consistent with the DAC Evaluation Quality Standards and the following principles:

• The main report is relatively short. It contains the aggregated information and highlights. All detailed information is referred to annexes.
• Information is organised in a simple format which is easily accessible to the reader. This means that discussion of the evaluation criteria is a focus in the report.

The evaluation report consists of three levels of information. First, the executive summary is written as a self-contained paper that provides the bare essentials for decision-makers on background, major conclusions, recommendations and lessons learned.

Other MFA staff	Consultant	Others
Embassy, BFT and other relevant departments: a. Comment on reports b. Possible participation in meeting about reports	Prepares final drafts (language checked, and copy edited).	*Partners, reference group and possibly other donors:* a. Comment on reports b. Possible participation in meeting about reports
Embassy, BFT and other relevant departments: a. Comment on report b. Possible participation in meeting about report	Prepares final draft (language checked, and copy edited).	*Partners, reference group, and possibly other donors:* a. Comment on report b. Possible participation in meeting about report

The second level is the main report (40-50 pages) of which a substantial part is made up of the main conclusions and recommendations. These are substantiated with more detailed information only to the extent necessary.

Detailed findings are referred to the annexes. Conclusions and recommendations in the main report have references to the relevant findings in the annexes.

The third level in the report is the annexes. These provide all information necessary to substantiate major conclusions and recommendations in the main report. The Terms of Reference, the team's itinerary, list of persons met, and list of documents used should also be annexed.

Danida's evaluation reports are published in print. Annexes are included in the CD-ROM that accompanies the report and are also available on the website: www.evaluation.dk.

The main report should include

- Executive Summary
- Introduction, with the background for and implementation of the evaluation as well as a presentation and discussion of the methodology used
- The setting in which assistance interventions are implemented, in terms of geography, policy, sectors and organisations, as appropriate
- Description of the planning, and implementation of development activities. This includes the presentation of objectives, target groups, components, financing, management, etc.
- The findings
- The conclusions based on findings and analysis, including efficiency, effectiveness, impact, relevance, and sustainability
- Lessons learned
- Recommendations

Reporting: Drafting and editing

Danida expects the evaluation team to maintain a professional standard in reporting. The evaluation team must employ the services of a professional copy editor.

Consultants are expected to present a draft report, which has been thoroughly copy edited, hence comments from the Evaluation Department and other stakeholders on presentation, language and structure can be reduced to a minimum. An unfinished draft is likely to provoke frustration among the readers and may result in tensions between consultants and stakeholders.

Readers have different needs for information. Some need highly aggregated information – others are interested in details. People also have different professional and educational background. The report is expected to meet the needs of all groups.

The language, use of terminology, illustrations, etc., are the main assistances for efficient communication. The report is written keeping target groups with the least specialist knowledge in mind.

Some general advice to make reports easier to read:

Brevity
- A rule of thumb is that the shorter a text, the greater the number of people who actually read it.

Structure
- The structure of the report is simple with titles and subtitles precisely identifying the topics discussed.
- A "horizontal" structure is used with not more than two levels of headings.
- The text is broken down in relatively small thematic or sequential parts, with simple and clear subtitles.

Editing
- The first sentence of paragraphs is used to make the main point, and the remainder to supplement, substantiate and discuss the main point.
- Illustrations, graphs, tables, etc. are used extensively to illustrate and elaborate on points that are emphasised.
- The details of tables and diagrams are not being repeated in written form. Instead, the reader's attention should be focused on the important points that are derived, and related to the problem under discussion.
- Jargon and difficult words are avoided. As far as possible, the report uses words that the readers understand. Essential technical terms that may be new to the reader are defined in the text, and in a glossary.
- Long and complicated sentences are broken down in several short sentences. It is useful to give lists of items or conditions step-by-step.
- Active sentences are preferred. The more complicated passive form is used more to focus attention on a specific topic.
- Ideas are not crammed together in complex sentences with commas and brackets. Separate main ideas are presented in separate sentences.

- The meaning of abbreviations and colloquial words is explained. Abbreviations are used as little as possible.
- Simple link words are used to split sentences and indicate the direction in which the argument is moving. Link words are simple, such as: also; even so; on the other hand; in the same way. Long words like: moreover; nevertheless; notwithstanding, are avoided.

Numbers
- Only data tables or diagrams contain detailed numbers. The written text highlights only the most important numbers and says what they mean. Percentages are in most cases rounded to the nearest whole number.
- It should be possible for the reader to get the main message from a table without consulting the text. Every table has a title, table number, reference to the source of information, sample size, and full description of what each figure refers to.
- Pictures, diagrams and illustrations are used to make the report lively and interesting. These are often more effective than written text for communicating straight-forward messages.
- Use space around the text. Ease of reading and understanding is more important than reducing the volume of writing.

Feedback from the evaluation team

The evaluation team will draft a brief comment on its experience of the evaluation process so that the Evaluation Department may draw lessons on how it worked from the perspective of the consultants.

Quality control and assurance

Danida expects the evaluation team to have a systematic process for controlling and assuring the quality of its evaluation process and product. Danida contracts require that a quality manager be identified.

The DAC Evaluation Quality Standards constitutes a frame of reference for all Danida evaluations; the evaluation team must ensure that its evaluation process and product are consistent with these standards.

Expectations of quality in methodology, and to a lesser degree in management of the evaluation, have been addressed to some extent earlier in this chapter.

Another important dimension of quality has to do with the behaviour of professionals involved in Danida evaluations; the following two sections deal with overlapping expectations of professional practise and behaviour with stakeholders and other parties – Evaluation Ethics – and more specifically with Danida and related parties – Code of Conduct for Consultants.

Code of Conduct for consultants

Consultants have cited two particularly difficult phases in the evaluation process where uncertainty about roles may threaten their independence: organisation of fieldwork and discussion of the draft report. The code, therefore, concentrates on issues relating to these two phases in the evaluation process, while also addressing a range of other potential conflicts.

Evaluations are carried out by independent consultants. It is the responsibility of the consultants to collect the information necessary to fulfil the tasks set out in the Terms of Reference for the assignment and to present findings, conclusions and recommendations in a well-structured and well-written report. Consultants prepare and undertake the fieldwork. They prepare an evaluation method for the collection of information, organise logistics, undertake surveys and studies and consult with relevant Danida and partner staff.

Conclusions and recommendations of the evaluation report are discussed with the Evaluation Department, other Danida staff, partners and possibly a reference group, but it is the consultants who have ultimate responsibility for conclusions and recommendations.

- Consultants should be prepared to engage in discussions with Danida staff and other stakeholders without perceiving any difference of opinions as a form of pressure. There is an inherent tension in any evaluation assignment, and consultants should expect their interpretations to be challenged. Professional discussions can be difficult, and consultants should avoid feeling too easily offended or overly sensitive to criticism.
- Consultants, who believe that they have been subject to undue pressure, inappropriate behaviour or even threats from Danida staff or other stakeholders (including those responsible for the activities under evaluation) during fieldwork or while preparing the report, should report this immediately to the Evaluation Department.
- It is the responsibility of the team leader to decide, whether it is appropriate that the Evaluation Department and possible stakeholders participate in meetings, interviews and field visits. Stakeholders may include Embassy staff, Danida advisers and company advisers, partner representatives, other parties responsible for the activities under evaluation, and others. Issues relevant to the team leader's decision include ensuring the impartiality of the evaluation, ensuring openness of discussions, and avoiding imbalance in numbers (too many 'evaluators' present can make stakeholders uncomfortable).
- If consultants encounter or suspect corruption or other illicit practice, while carrying out their evaluation, it is their responsibility to report this to the Evaluation Department, which will ensure that Danida can take appropriate action. Only cases sustained by legally valid evidence can be reported in evaluation report.
- During the preparation and the undertaking of the fieldwork, the consultant is responsible for informing the Evaluation Department about the time schedule of the evaluation and the persons to be contacted and interviewed. This information ensures coverage of all important sources of information and helps to avoid duplication and the evaluation becoming a burden upon a small group of key resource persons.
- After the draft report has been produced, consultants should avoid any direct contact to Danida staff and other stakeholders without prior agreement with the Evaluation Department. If consultants meet with Danida staff, the Evaluation Department should be present, and all

correspondence between consultants and Danida staff should be copied to the Evaluation Department.

- Consultants should correct all factual errors and inaccuracies and make changes related to report structure, consistency, the analytical rigour, the validity of evidence, and requirements in Terms of Reference, as pointed out by the Evaluation Department. However, consultants should only agree to make changes to conclusions and recommendations of the evaluation that they regard as qualitative improvements. Where consultants disagree with changes proposed by the Evaluation Department, other Danida staff, and other stakeholders, they must present counter arguments. In cases of serious disagreement, the alternative assessments and counterarguments by the consultant should be presented in the report, if appropriate as footnotes. In cases of minor issues, the consultant must explain in writing to the Evaluation Department why suggested changes are not considered relevant or appropriate.

- Statements should not be made on behalf of the evaluation team if all team members have not had the opportunity to express their agreement. Unless disagreement has been clearly expressed in the report, all consultants involved in an assignment are expected to be loyal to the conclusions of the report. If internal team disagreement has not been reported, it is not appropriate that team members subsequently criticise the quality or the conclusions of the evaluation after it has been published.

- A particular case of internal team disagreement regards the situation in certain partner countries, where criticism can have serious consequences for national members of the evaluation team. Therefore, the team leader should ensure that national consultants either endorse a report in its entity or that their limited roles are clearly described in the report.

- Participation by consultants in the dissemination of evaluation findings after the publication of the evaluation report is appreciated. If consultants contribute to a public debate about the interpretation of the evaluation, however, this has to be agreed with the Evaluation Department. Particular attention should be paid to the issue of confidentiality of information obtained during the evaluation process, and consultants should not orally or in writing present information, which is not included in the report.

Evaluation ethics

Evaluation is a profession of persons with varying interests and differences in training, experience and work setting. Despite the diversity, the common ground is that evaluators aspire to provide quality information and assessment. Also that the evaluator has a responsibility versus the commissioning organisation, and the institutions, groups and individuals involved or affected by the evaluation.

Evaluators are expected to abide by international ethical standards. The following are adapted from those developed by the Canadian International Development Agency (CIDA, 1990):

Cultural intrusion
Evaluators should ensure that they are familiar with beliefs, manners and customs before going into the field. Both in data collection and reporting, it is imperative that customs regarding dress, personal interaction, religious beliefs and practices are respected.

Anonymity/Confidentiality
Evaluators must respect people's right to provide information in confidence and ensure that sensitive data cannot be traced to its source.

Responsibility for evaluations
In developing countries, criticism can have serious consequences for a national. Evaluators must take care that those involved as local evaluators either endorse a report, or that their restricted roles are clearly described in the report. Statements should not be made on behalf of the evaluation team if other team members have not had an opportunity to disagree.

Right to privacy
Evaluators should realise that people can be extremely busy and their participation in evaluations can be burdensome. Therefore, evaluators should provide ample notice and minimise demands on time.

Supremacy of fundamental values
There is a delicate balance between certain cultural practices and the deprivation of fundamental rights and freedoms. While evaluators are expected to respect other cultures, they must also be aware of the values affecting minorities and particular groups. In such matters, the United Nations Universal Declaration of Human Rights (1948) is the operative guide.

Omissions
Ethically, evaluators have a responsibility to bring to light issues and findings which may not relate directly to the Terms of Reference. Certain other issues can cause difficulties for the evaluator and should be acknowledged and discussed with the Evaluation Department as necessary.

Evaluation of individuals
Performance evaluation is not normally a part of evaluations, though reports will touch issues such as leadership and management competence that border on evaluation of individuals. The evaluator is not expected to evaluate individuals and must balance an evaluation of management functions with this general principle.

Evidence of wrongdoing
Evaluations sometimes uncover evidence of wrongdoing. This can be complicated by an interpretation in the culture of the partner country that is different from the interpretation in the donor country. The evaluation team should consult the Evaluation Department when there are questions about what should be reported, how and to whom.

ANNEX 1
BIBLIOGRAPHY

1. Bamberger, M., Rugh, J., Church, M., & Fort, L. (2003), *Shoestring Evaluation: Designing Impact Evaluations under Time, Budget and Data Constraints*. American Journal of Evaluation 2004; 25: 5-37. (See also recent publication: Ref. No. 15 below)

2. Beck, T. (2006), *Evaluating humanitarian action using the OECD-DAC criteria: An ALNAP guide for humanitarian agencies*. London, UK: Overseas Development Institute.
   www.alnap.org/publications/eha_dac/pdfs/eha_2006.pdf

3. Breier, H. (2005), *Joint Evaluations: Recent Experiences, Lessons Learnt and Options for the Future*, DAC Evaluation network Paper, OECD, Paris.
   www.oecd.org/dataoecd/28/9/35353699.pdf

4. Danish Ministry of Foreign Affairs, *Guidelines for Program Management*. Copenhagen. Updated regularly only on
   www.amg.um.dk/en

5. European Commission (2001), *Guide to Ex Ante Evaluation*. Brussels.
   www.ec.europa.eu.int/comm/budget/evaluation/pdf/ex_ante_guide_en.pdf

6. Freeman, T. (2004), *Managing the External Evaluation Team During Joint Evaluations: Challenges and Responses*. Ottawa, Canada: Goss Gilroy Inc.
   www.ggi.ca

7. Hallam, A. (1998), *Evaluating Humanitarian Assistance Programmes in Complex Emergencies*. London, UK: ODI. Good Practice Review 7.

8. OECD (2006), *DAC Evaluation Quality Standards*. OECD, Paris.
   www.oecd.org/dataoecd/30/62/36596604.pdf

9. OECD (1991), *Principles for Evaluation of Development Assistance*. OECD, Paris.
www.oecd.org/dataoecd/secure/9/11/31779367.pdf

10. OECD (2006), *DAC Guidance for Managing Joint Evaluations*. OECD, Paris.
www.oecd.org/dataoecd/28/14/37484787.pdf

11. OECD (2002), *DAC Glossary of Key Terms in Evaluation and Results Based Management*. OECD, Paris.
www.oecd. www.oecd.org/dataoecd/29/21/27594804.pdf

12. OECD (1999), *Guidance for Evaluating Humanitarian Assistance in Complex Emergencies*. OECD, Paris.
www.oecd.org/dataoecd/9/50/2667294.pdf

13. *Paris Declaration on Aid Effectiveness: Ownership, Harmonisation, Alignment, Results and Mutual Accountability*.
www.oecd.org/dataoecd/11/41/34428351.pdf

14. Trochim, W. M. K. (2002), *Research Knowledge Database*.
www.socialresearchmethods.net/kb/

15. Weiss, H. B. (Summer 2005), *The Evaluation Exchange: Editorial*. Harvard Family Research Project.

16. World Bank (2006), *Conducting Quality Impact Evaluations under Budget, Time and Data Constraints.*
www.worldbank.org/ieg/ie/

ANNEX 2
DAC EVALUATION QUALITY STANDARDS

Introduction

1. The DAC Evaluation Quality Standards identify the key pillars needed for a quality evaluation process and product. They have been prepared by DAC members in order to define member countries' expectations of evaluation processes, and evaluation products. The Standards are not binding on member countries, but a guide to good practice and aim to improve the quality of development intervention evaluations. They are intended to contribute to a harmonised approach to evaluation in line with the principles of the Paris Declaration on Aid Effectiveness.[14]

2. The Standards are intended to:

   - Provide standards for the process (conduct) and products (outputs) of evaluations.
   - Facilitate the comparison of evaluations across countries (meta-evaluation).
   - Facilitate partnerships and collaboration on joint evaluations.
   - Better enable member countries to make use of each others' evaluation findings and reports (including good practice and lessons learned).
   - Streamline evaluation efforts.

3. The Standards support evaluations that adhere to the DAC Principles for the Evaluation of Development Assistance, including impartiality and independence, credibility and usefulness, and should be read in conjunction with those principles.

---

14) *Paris Declaration on Aid Effectiveness: Ownership, Harmonisation, Alignment, Results and Mutual Accountability.*

The Principles focus on the management and institutional set up of the evaluation systems within development agencies and remain the benchmark against which OECD Members are assessed in the DAC Peer Reviews. By contrast the Standards provide guidance on the conduct of evaluations and for reports. While the Standards are not binding on every evaluation, they should be applied as widely as possible and a brief explanation provided where this was not possible.

4. The term 'development intervention' is used in the Standards as a general term to refer to the subject of the evaluation and may refer to an activity, project, programme, strategy, policy, topic, sector, operational area, institutional performance etc.

5. The Standards recognise that the product of an evaluation may be in a variety of different forms, including oral or written reports, presentation and community workshops. The term 'evaluation report' is used to cover all forms of evaluation products.

1.   Rationale, purpose and objectives of an evaluation

1.1   The rationale of the evaluation
Describes why and for whom the evaluation is undertaken and why it is undertaken at a particular point in time.

1.2   The purpose of the evaluation
The evaluation purpose is in line with the learning and accountability function of evaluations. For example the evaluation's purpose may be to: Contribute to improving an aid policy, procedure or technique; Consider a continuation or discontinuation of a project/programme; Account for aid expenditures to stakeholders and tax payers.

1.3   The objectives of the evaluation
The objectives of the evaluation, specify what the evaluation aims to achieve.

For example:

- To ascertain results (output, outcome, impact) and assess the effectiveness, efficiency and relevance of a specific development intervention;
- To provide findings, conclusions and recommendations with respect to a specific policy, programme etc.

2.   Evaluation scope

2.1   Scope of the evaluation
The scope of the evaluation is clearly defined by specifying the issues covered, funds actually spent, the time period, types of interventions, geographical coverage, target groups, as well as other elements of the development intervention addressed in the evaluation.

2.2   Intervention logic and findings
The evaluation report briefly describes and assesses the intervention logic and distinguishes between findings at the different levels: inputs, activities, outcomes and impacts. The report also provides a brief overall assessment of the intervention logic.

2.3   Evaluation criteria
The evaluation report applies the five DAC criteria for evaluating development assistance: relevance, efficiency, effectiveness, impact and sustainability. The criteria applied for the given evaluation are defined in unambiguous terms. If a particular criterion is not applied this is explained in the evaluation report, as are any additional criteria applied.

2.4   Evaluation questions
The questions asked, as well as any revisions to the original questions, are documented in the report for readers to be able to assess whether the evaluation team has sufficiently assessed them.

3.   Context

3.1   The development and policy context
The evaluation report provides a description of the policy context relevant to the development intervention, the development agency's and partners' policy documents, objectives and strategies.

The development context may refer to: regional and national economy and levels of development.

The policy context may refer to: Poverty reduction strategies, gender equality, environmental protection and human rights.

3.2   The institutional context
The evaluation report provides a description of the institutional environment and stakeholder involvement relevant to the development intervention, so that their influence can be identified and assessed.

3.3   The socio-political context
The evaluation report describes the socio-political context within which the intervention takes place, and its influence on the outcome and impact of the development intervention.

3.4   Implementation arrangements
The evaluation report describes the organisational arrangements established for implementation of the development intervention, including the roles of donors and partners

4.   Evaluation methodology

4.1   Explanation of the methodology used
The evaluation report describes and explains the evaluation method and process and discusses validity and reliability. It acknowledges any constraints encountered and their impact on the evaluation, including their impact on the independence of the evaluation. It details the methods and

techniques used for data and information collection and processing. The choices are justified and limitations and shortcomings are explained.

4.2 Assessment of results

Methods for assessment of results are specified. Attribution and contributing/confounding factors should be addressed. If indicators are used as a basis for results assessment these should be SMART (specific, measurable, attainable, relevant and time bound).

4.3 Relevant stakeholders consulted

Relevant stakeholders are involved in the evaluation process to identify issues and provide input for the evaluation. Both donors and partners are consulted. The evaluation report indicates the stakeholders consulted, the criteria for their selection and describes stakeholders' participation. If less than the full range of stakeholders was consulted, the methods and reasons for selection of particular stakeholders are described.

4.4 Sampling

The evaluation report explains the selection of any sample. Limitations regarding the representativeness of the evaluation sample are identified.

4.5 Evaluation team

The composition of evaluation teams should posses a mix of evaluative skills and thematic knowledge, be gender balanced, and include professionals from the countries or regions concerned.

5. Information sources

5.1 Transparency of information sources

The evaluation report describes the sources of information used (documentation, respondents, literature etc.) in sufficient detail, so that the adequacy of the information can be assessed. Complete lists of interviewees and documents consulted are included, to the extent that this does not conflict with the privacy and confidentiality of participants.

5.2 Reliability and accuracy of information sources
The evaluation cross-validates and critically assesses the information sources used and the validity of the data using a variety of methods and sources of information.

6. Independence

6.1 Independence of evaluators vis-à-vis stakeholders
The evaluation report indicates the degree of independence of the evaluators from the policy, operations and management function of the commissioning agent, implementers and beneficiaries. Possible conflicts of interest are addressed openly and honestly.

6.2 Free and open evaluation process
The evaluation team is able to work freely and without interference. It is assured of cooperation and access to all relevant information. The evaluation report indicates any obstruction which may have impacted on the process of evaluation.

7. Evaluation ethics

7.1 Evaluation conducted in a professional and ethical manner
The evaluation process shows sensitivity to gender, beliefs, manners and customs of all stakeholders and is undertaken with integrity and honesty. The rights and welfare of participants in the evaluation are protected. Anonymity and confidentiality of individual informants should be protected when requested and/or as required by law.

7.2 Acknowledgement of disagreements within the evaluation team
Evaluation team members should have the opportunity to dissociate themselves from particular judgements and recommendations. Any unresolved differences of opinion within the team should be acknowledged in the report.

8.   Quality assurance

8.1   Incorporation of stakeholders' comments

Stakeholders are given the opportunity to comment on findings, conclusions, recommendations and lessons learned. The evaluation report reflects these comments and acknowledges any substantive disagreements. In disputes about facts that can be verified, the evaluators should investigate and change the draft where necessary. In the case of opinion or interpretation, stakeholders' comments should be reproduced verbatim, such as in an annex, to the extent that this does not conflict with the rights and welfare of participants.

8.2   Quality control

Quality control is exercised throughout the evaluation process. Depending on the evaluation's scope and complexity, quality control is carried out either internally or through an external body, peer review, or reference group. Quality controls adhere to the principle of independence of the evaluator.

9.   Relevance of the evaluation results

9.1   Formulation of evaluation findings

The evaluation findings are relevant to the object being evaluated and the purpose of the evaluation. The results should follow clearly from the evaluation questions and analysis of data, showing a clear line of evidence to support the conclusions. Any discrepancies between the planned and actual implementation of the object being evaluated are explained.

9.2   Evaluation implemented within the allotted time and budget

The evaluation is conducted and results are made available in a timely manner in relation to the purpose of the evaluation. Unenvisaged changes to timeframe and budget are explained in the report. Any discrepancies between the planned and actual implementation and products of the evaluation are explained

9.3   Recommendations and lessons learned
Recommendations and lessons learned are relevant, targeted to the intended users and actionable within the responsibilities of the users. Recommendations are actionable proposals and lessons learned are generalizations of conclusions applicable for wider use.

9.4   Use of evaluation
Evaluation requires an explicit acknowledgement and response from management regarding intended follow-up to the evaluation results. Management will ensure the systematic dissemination, storage and management of the output from the evaluation to ensure easy accessibility and to maximise the benefits of the evaluation's findings.

10.   Completeness

10.1 Evaluation questions answered by conclusions
The evaluation report answers all the questions and information needs detailed in the scope of the evaluation. Where this is not possible, reasons and explanations are provided.

10.2 Clarity of analysis
The analysis is structured with a logical flow. Data and information are presented, analysed and interpreted systematically. Findings and conclusions are clearly identified and flow logically from the analysis of the data and information. Underlying assumptions are made explicit and taken into account.

10.3 Distinction between conclusions, recommendations and lessons learned
Evaluation reports must distinguish clearly between findings, conclusions and recommendations. The evaluation presents conclusions, recommendations and lessons learned separately and with a clear logical distinction between them. Conclusions are substantiated by findings and analysis. Recommendations and lessons learned follow logically from the conclusions.

10.4 Clarity and representativeness of the summary

The evaluation report contains an executive summary. The summary provides an overview of the report, highlighting the main conclusions, recommendations and lessons learned.

ANNEX 3
DANIDA'S EVALUATION POLICY

1.   Introduction

Evaluation has been used systematically since the early days of Danish development co-operation. In 1982 Danida established a special unit responsible for evaluation. The use of evaluations has developed successively in four main stages:

Prior to 1982 evaluations focussed essentially on individual projects and programmes. Most of these were mid-term or phase evaluations conducted as the project moved from one phase to the next. Only a few end-evaluations were conducted, and only occasionally were ex post evaluations carried out to study the long-term effect of projects.

In the period 1982-87, after Danida's Evaluation Unit was established, it was agreed to use evaluations for more systematic studies to improve the quality of Danida financed development activities. Also in this period most evaluations were mid-term or phase evaluations of individual projects. The trend was to replace mid-term evaluations with internal reviews and increase the number of end-evaluations. The use of evaluations was more systematic in the sense that it was guided by an annual evaluation programme to ensure that the sample of evaluated projects and programmes were representative for Danish bilateral cooperation.

During 1987-97 the number of individual project evaluations was reduced and the number of thematic and sector evaluation increased. As a principle, all evaluation reports were made public. In 1992, informing the public was included as an essential goal for evaluation in accordance with DAC principles. In this period evaluations became more experimental and included a number of impact evaluations as well as use of participatory methods. All evaluations were conducted by external, independent consultants.

In 1997 an evaluation policy was formulated and the Evaluation Secretariat was established as a separate, independent entity within the Ministry of Foreign Affairs (in 2003 the name was changed to the Evaluation Department). A Review of Evaluation in Danida by external international consultants found in 2003 that Danida's evaluation system basically is sound in comparison to general international standards. Based on the public hearing held in 2004 the institutional arrangement of the Evaluation Department was confirmed.

While the policy has provided the overall framework for specific Danida evaluation, practice has developed considerably since 1997. In particular the move towards sector wide approaches has required donor evaluation departments and partners to work together to conduct evaluations jointly. Danida has been at the forefront in this move towards joint evaluations. In recent years about half of evaluations have been conducted jointly with partners.

The Rome, Marrakech, and Paris Declarations on ownership, alignment, harmonisation, managing for results, and the Good Humanitarian Donorship Principles have reinforced this trend. Fortunately, the donor evaluation community has had a good basis for harmonising evaluation work: the DAC Principles for Evaluation which has provided a common frame of reference for evaluation across donors and countries.

The new evaluation policy captures these changes and makes collaboration and partnership central principles for Danida's evaluation.

2. Objectives

An evaluation is an assessment, as systematic and objective as possible, of on-going or completed development activities, their design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, developmental efficiency, effectiveness, impact and sustainability.[15]

---

15) DAC Principles for Aid Evaluation, OECD-DAC, 1991. (the definition was reconfirmed in DAC's Glossary of Key Terms in Evaluation and Results Based Management in 2002).

Evaluations are carried out to *generate knowledge* and *accountability* information about development interventions:

Evaluations shall contribute to the improvement of development cooperation by collation, analysis and dissemination of experience from current and completed development activities. They shall seek the causes and explanations as to why activities succeed or fail to succeed and produce information to help improve more relevance and effectiveness of future activities. The target group is Danida's management and staff, government authorities and other concerned parties in partner countries, stakeholders of a particular organisation under evaluation, and Danish and foreign development professionals.

Further, evaluations shall provide parliamentarians and the general public in Denmark and partner countries with professional documentation for the use and results of development cooperation. They shall also contribute to a better understanding of development cooperation and of its potential and limitations as an instrument for economic and social change.

While evaluation is Evaluation is distinct from review, which is a management tool to monitor whether an activity or programme is on track and produces the immediate output.

3.   Guiding Principles

These principles reflect the core values that anchor Danida's evaluation work. They are interdependent and mutually reinforcing and, as such, they form an overarching frame of reference with which all aspects of an evaluation must be consistent.

The guiding principles articulate fundamental expectations of evaluators, of evaluation processes and products, of methodology and of those managing the evaluation.

Danida's guiding principles draw on, and are consistent with, the DAC Principles for Aid Evaluation as well as DAC Evaluation Standards.[16]

16) Annex 2.

All Danida evaluations must be consistent with the principles of:

Independence – the evaluator's judgments are not influenced by pressure or conflict of interest. Members of evaluation teams must not in person have been engaged in the activities to be evaluated. Companies involved in the preparation or implementation of the activities to be evaluated cannot be evaluator of these activities.

Impartiality – the personal preferences of the evaluator do not affect the evaluation. Evaluations must give a balanced presentation of strengths and weaknesses. Although evaluators are responsible for all conclusions, different views of interested parties should be reflected in the evaluation report.

Objectivity – the evaluation rests on verifiable findings of fact. Judgments must be clearly separated from factual statements.

Transparency – features, issues and decisions significant to the evaluation are identified and explained clearly. Relevant parties in Denmark and the partner country must be consulted during the preparation and implementation of evaluations, drafting of Terms of Reference and discussion of the draft report.

Partnership – in conformity with the Rome and Paris declarations on harmonisation and partnership as well as the Good Humanitarian Donorship Principles evaluations should to the extent possible be undertaken in partnership with stakeholders in partner countries and other donors.

Feasibility – the appropriate methodology and resources required by the evaluation are available.

Propriety – the evaluation does not harm individuals.

Cost-efficiency – the evaluation is realised at least cost.

Accuracy – the data do not contain errors of significance to the evaluation.

Fairness – evaluations give a balanced presentation of strengths, weaknesses and different views.

Credibility – the evaluation is conducted in such a way that the results are believed.

Usefulness – the evaluative process and the information it produces are utilized by users and stakeholders.

4.   The Evaluation Department

The Evaluation Department is responsible for evaluating the performance of development activities to which Denmark has contributed. It provides feedback to the Ministry of Foreign Affairs about development cooperation processes, results, and accountability to Parliament and the public for the results of Danish development assistance.

The Evaluation Department is an independent department in the Ministry of Foreign Affairs. The Head of the Evaluation Department refers to the Head of the South Group – the State Secretary. While evaluation is part of the Performance Management Framework, the Evaluation Department holds no responsibility for daily administration, implementation and monitoring of development cooperation.

Core responsibilities and duties of the Evaluation Department include:

Programming, formulating and managing evaluations of development activities funded or co-funded by Denmark, including multilateral and NGO activities, to the extent possible in the form of joint evaluations with national authorities in partner countries and other donors.

Contributing to the learning process within the Ministry of Foreign Affairs and partner countries by providing feedback to operational departments and management about relevance, impact and operational performance of development activities. Participation in the Programme

Committee is a key vehicle for feeding back information to operational departments and embassies.

Contributing to increased accountability of Danish development assistance by disseminating evaluation results to the Danish public, Parliament, the Ministry's management and staff, the Danida Board, partner countries, other donors and other interested parties.

Developing and improving evaluation methodology and guidelines, as well as methods for disseminating results.

Participating in international co-operation on evaluation principally in DAC, EU and Nordic contexts.

Analysing Programme and Project Completion Reports, and maintaining the filing system for them.

Contributing to the development of evaluation capacity in partner countries through bilateral and multilateral co-operation.

Contributing to the development of evaluation capacity in NGOs and the Danish resource base.

5.   Evaluation Programming

The Evaluation Department is responsible for preparing two-year rolling programmes on the basis of the strategic and operational needs of Danida and partners. The programme is prepared in consultation with relevant stakeholders: the Ministry of Foreign Affairs, partner authorities and other donors with a particular view to carry out evaluations jointly with these partners.

The evaluation programme must achieve a suitable coverage of geographical areas, large and small partner countries and thematic areas over time:

Danish development cooperation with a particular country, one or more sectors in a partner country, principally in partnership with relevant stakeholders.

Cross cutting issues (gender equality, environment, human rights and democracy) and priority themes (HIV/AIDS, children and young people, sexual and reproductive health, etc.) as well as other themes such as planning and implementation of policies, strategies and programmes.

Development cooperation instruments (programme and project assistance, assistance via NGOs, mixed credits, budget support, humanitarian assistance, research support, fellowship programmes).

Multilateral development cooperation, principally in co-operation with other donors. Evaluation of multilateral organisations is expected to be gradually replaced by joint donor assessments of the evaluation function of multilateral organisations.

Ongoing or completed projects and programmes to which Denmark has contributed.

The evaluation programme shall contain a brief rationale for each evaluation: the primary objective and which features in particular the evaluation shall comprise. It will be presented to The Board of Danida and to Parliament for information and comment.

6.   The Evaluation Process

6.1   Preparation
The management and staff of Danida as well as the relevant partners (the country, other donors, NGOs or multilateral organisations) should be involved in the preparation of evaluations by participating, to the greatest possible degree, in the formulation of Terms of Reference with a view to focusing the evaluation on relevant subjects. Before an evaluation is set in motion, the Evaluation Department prepares, in consultation with relevant interested partners, a description of the evaluation that covers:

The main objective of the evaluation.

What the evaluation shall cover: the entire or part of a programme or project, a sector, an instrument of development cooperation or a particular theme.

Specific issues or features to be covered by the evaluation.

Prospective approach and methodology.

Time schedule.

On this basis the actual Terms of Reference are prepared which forms the basis for the selection of the evaluation team and the conduct of the evaluation. The Terms of Reference shall specify which specific questions the evaluation shall seek to answer.

An evaluation should normally consider the following issues:

Relevance – Are the development interventions relevant to Danish and partners' development policies, goals and strategies as well as global priorities: poverty reduction, a sustainable environment, gender equality and democratisation and human rights? Is the activity relevant in relation to the needs and priorities of the intended beneficiaries?

Effectiveness – Achievement of objectives: Have the primary objectives which have been defined for the activity been achieved? Have the specific results planned been achieved?

Efficiency – How economically have resources/inputs (funds, expertise, time, etc.) been converted to results? Are the investment and recurrent costs justified? Could the same results have been achieved with fewer resources?

Impact – What positive and negative, primary and secondary long-term effects have been produced by a development intervention, directly or indirectly, intended or unintended?

Sustainability – The probability of long-term benefits. Will the intended benefits continue when development cooperation is terminated? Is local ownership established?

6.2 Selection of Evaluation Team
Evaluation teams are selected through international competitive bidding in accordance with prevailing regulations.

The criteria for the selection of the evaluation team are professional competence, experience in relation to the task, independence (no conflict of interest) and the quality of the evaluator's proposal. The team must constitute a representation of the relevant professional expertise. Professional expertise from the partner country shall as far as possible be represented on the team.

6.3 Implementation
Evaluations are typically carried out in three phases. In the first phase the evaluator prepares an operational evaluation plan, consistent with the Terms of Reference, that details specific questions, assessment criteria, approach, design, data collection methods, analytical framework, preliminary findings based on documentary review and interviews, provisional conclusions, report outline, and a detailed work plan for the second phase. This may cover further documentation studies, interviews, primary data collection, field observations etc. Any changes from the Terms of Reference are justified and agreed with the Evaluation Department.

On approval of the detailed operational plan by the Evaluation Department or, in the case of joint evaluations the steering committee, the evaluator proceeds to the second phase that includes collecting, consolidating and analysing data, establishing and clearing findings of fact, formulating conclusions and recommendations, preparing and clearing the draft report. The third phase entails finalising the report.

6.4 Independent Reporting
Danida favours a participatory approach to the conduct of evaluations. Danida management and staff as well as relevant representatives of partners should be involved in the implementation of evaluations to the

greatest extent possible especially at (but not limited to) key stages of the evaluation process: preparation and planning, clearance of factual findings, and discussion of conclusions and recommendations. This involvement takes place through regular and systematic communication with the Evaluation Department and stakeholders through meetings, workshops and seminars as needed over the course of the evaluation.

The evaluation team has the final responsibility for the contents of the report. Any disagreements among the evaluation team or between the evaluation team, Danida, and relevant partners that are significant to conclusions and recommendations must be reflected in the report, either in the form of comments in the text, footnotes or as a special section.

Evaluation reports must be brief and concise, and the presentation must be clear and adjusted to the target group. The normal language is English, and where relevant reports are translated into French, Spanish or Portuguese.

7. Dissemination

All evaluations to which Danida is a partner are published in the form of printed reports and summaries as well as electronically on the Evaluation Department's website.

For every evaluation a brief summary in Danish, English and other relevant languages must be compiled with a view to publication. The summary must cover the most important observations and conclusions of the evaluation. The presentation must be made in language accessible to non-professionals.

If an evaluation is thought to be of interest to a broader audience Danish and local language versions will be produced (possibly abbreviated) and edited to communicate effectively to readers that are not professional specialists.

The Evaluation Department contributes actively to disseminating the evaluation experience of Danida as well as other development organisations via workshops and seminars for staff and partners. Furthermore, the Evaluation Department assists Danida's Centre for Competence Development in the dissemination of evaluation experience.

The Evaluation Department will contribute to the incorporation of evaluation experience in policies, strategies and guidelines etc. The participation in the Programme Committee is a key vehicle for this.

The Evaluation Department reports annually directly to the Board of Danida and to the Foreign Affairs Committee of Parliament about the activities of the Evaluation Department, the findings of evaluations, and the follow up on previous years' evaluations.

In addition, the Evaluation Department contributes to Danida's Annual Report and the Annual Performance Report.

8.   Follow-up

At the conclusion of an evaluation, a follow up memo is prepared by the relevant department/embassy. This takes note of Danida's position on the conclusions and recommendations as well as identifying which departments are responsible for the agreed follow up activities. The follow up memo is discussed in the Programme Committee and signed off by the State Secretary. The Evaluation Department undertakes to monitor the implementation of the follow up activities at regular intervals.

9.   Monitoring implementation of the policy

Implementation of the Evaluation Policy is guided by Danida's Evaluation Guidelines which specify in greater detail how evaluations are conducted, the quality standards required and codes of conduct for parties to an evaluation.

Implementation of the policy will be reported in conjunction with the annual report on evaluation activities submitted to the Board of Danida and to the Foreign Affairs Committee of Parliament. Implementation of the evaluation programme will be part of the annual results contract between the Evaluation Department and Management.

The Evaluation Policy should be assessed and, if needed, revised after five years.

12406 / DESIGNGRAFIK