DANIDA’s Evaluation Guidelines

EVALUATION GUIDELINES

- February 1999 (2nd edition, revised 2001).

The main objective of these guidelines is to operationalise Danida’s Evaluation Policy by clarifying principles, concepts, terms and processes of evaluation as applied by Danida.

Table of Contents

PREFACE

The main objective of these Guidelines is to operationalise Danida's Evaluation Policy by clarifying principles, concepts, terms and processes of evaluation as applied by Danida.

Contents | Next

The Guidelines shall be applied to evaluation of development activities financed by Danida. They should guide specific evaluations and serve as a source of general information for the staff of Danida, for concerned parties in partner countries, for external consultants, and for Danida's technical assistance personnel in the field.

The Guidelines are divided into eight chapters. The first two chapters discuss the basic principles of evaluation and different types of evaluation. Chapters three to six describe the evaluation process, criteria, methods, and quality assurance. The final two chapters outline reporting requirements and utilisation of evaluations.

The present Guidelines replace the 1994 Danida Guidelines for Evaluation, they will be subject to improvements in the light of experience from practical application. To this end comments and suggestions from the users would be welcome. They should be addressed to Danida's Evaluation Secretariat.

1.1 WHAT IS EVALUATION?

Contents | Previous | Next

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

OECD, through its Development Assistance Committee (DAC), has recommended the definition quoted below for its member countries. This definition has been adopted by all major donor agencies, including Danida.

Danida makes the distinction between "evaluation" and "review". The former is an inquiry that will be available to the public, while the latter is an internal, periodic assessment.

Danida reserves the term “evaluation” for retrospective studies of what has happened and why. Some agencies also use the term evaluation for prospective studies conducted prior to a project’s approval. Such studies are termed “appraisals” by Danida.

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

Evaluations are carried out at different stages. Evaluations of ongoing development activities are called interim evaluations, and usually take place midterm in the implementation period or at the end of a distinct phase.

Interim evaluations will typically focus on operational activities, but will also take a wider perspective and possibly give some consideration to longterm effects. *Interim Evaluations* should not be confused with Reviews.

Endevaluations aim to establish the situation when external aid is terminated and to identify the possible need for followup activities either by Danida or the partner country.

Expost evaluations are carried out two to five years after external support is terminated. The main purpose is to assess what lasting impact it has had or is likely to have and to extract lessons of experience.

The present chapter provides definitions of key concepts used in evaluation activities and discusses the purpose of evaluation in Danish development assistance.

A discussion of differences and similarities in evaluating different types and aspects of bilateral aid is included in Chapter 2. The administrative procedures involved in planning and carrying out an evaluation are discussed in Chapter 3.

The definition quoted above specifies five specific evaluation criteria to be used, i.e. relevance of objectives, development efficiency, effectiveness, impact, and sustainability. The application of these evaluation criteria is explained in Chapter 4.

In order to make the best use of resources, the scope of the evaluation has to be defined and the right methods applied. This is discussed in Chapter 5.

Evaluations are intended to provide reliable information to political authorities, the general public and development agencies. In Chapter 6, the professional standards to be applied are discussed in terms of quality of information and analysis, interdisciplinarity, impartiality, professional ethics, etc.

Evaluation is not an aim in itself. The true value of evaluation lies in using the results and conclusions to improve the quality of development assistance. Chapter 7 deals with the dissemination and transfer of the results of evaluation work.

Finally, the use of evaluation findings as seen in relation to the various target and user groups involved is discussed in Chapter 8.

1.2 DANIDA'S EVALUATION POLICY

Contents | Previous | Next

Evaluation has been used systematically since the early days of Danish development aid. 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 1982-1987, after Danida’s Evaluation Unit was established, it was agreed to use evaluations for more systematic studies to improve the quality of Danida financed aid 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 aid. As a policy, all evaluation teams should include one Danida staff member.

In the period 1987-1997 the number of individual project evaluations was reduced and the number of thematic and sectoral evaluation increased. As a policy, 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. In general all evaluations were conducted by external, independent consultants.

In 1997 a new evaluation policy (See Annex II) was formulated. At the same time the Evaluation Unit was renamed Evaluation Secretariat and established as a separate, independent entity within the Ministry of Foreign Affairs. The policy is the basis for the present guidelines and its principles and concepts are elaborated in the following chapters.

The tasks of the Evaluation Secretariat are to:

- *plan, formulate and carry out evaluations of Danida-financed aid activities*
- *disseminate and contribute to the dissemination of evaluation results and experience to Danida staff, authorities in partner countries, other interested bodies and to the public*
- *contribute to the development of strategy and policy in Danida on the basis of own and other donors' experience*
- *build and maintain an electronic database of evaluations*
- *develop and improve methods of evaluation and evaluation guidelines*
- *contribute to the development of evaluation capacity in Danida Programme countries*
- *participate in international co-operation on evaluation.*

(Danida policy)

1.3 PURPOSE OF EVALUATION

Danida evaluations are intended as a management tool that should serve two main purposes:

1. Accumulation of Experience
Evaluations shall contribute to the improvement of aid by the collation, analysis and dissemination of experience from current and completed aid activities. Evaluations shall seek the causes and explanations as to why activities succeed or fail to succeed and produce information that helps make future activities more relevant and effective.

2. Documentation

Evaluations shall provide political decision makers and the general public with professional documentation as to the use and results of aid resources and also contribute to a better understanding of development aid, its potential and limitations as an instrument for economic and social change.

(Danida policy)

Both objectives must be satisfied in Danida's evaluation activities, though individual evaluations do not necessarily have to fulfil both objectives to the same degree. An implication of this is that it may be useful to make the distinction between two types of evaluations depending on their main purpose: those designed essentially to provide documentation, and those designed to accumulate experience. These will have distinct different features and be aimed at different target groups, as described in Chapters 1.4 – 1.5.

The target group is Danida's management and staff, partners in partner countries and the resource base of Danish and foreign aid professionals. But evaluation activities should also serve the information needs of political decisionmakers and of the general public in Denmark and the partner countries.

Evaluations are expected to produce high quality, reliable information which can help guide future decisions and thus improve Danish development assistance.

Evaluations are also expected to be impartial and carried out by professionals who are independent of the development activities under study. This will therefore help ensure accountability for the results of public expenditures both at the partner and the donor side.

1.4 EVALUATIONS TO ACCUMULATE EXPERIENCE

A main concern for Danida is that the experience gained in evaluations be employed in improving international development assistance. To this effect it is necessary that the evaluations respond to needs in Danida and the partner countries, and that they are credible. In some previous evaluations, conclusions and recommendations have lacked realism.

This may have been caused by lack of communication between evaluators and Danida about what is feasible and possible. In order to improve upon this problem it has been decided that operative personnel should be more actively involved as resource persons in connection with future evaluations designed to accumulate experience.

The target group for evaluations to accumulate experience is primarily Danida's management and staff, relevant institutions in partner countries and the external professional resources and institutions used by Danida in Denmark and other countries.

Evaluations to Accumulate Experience

- **Aim**
  Draw lessons to improve Danish development assistance
- **Target group**
  Danida and partner institutions, and external professional resources
- **Evaluators**
  Interdisciplinary teams of external evaluators
- **Terms of Reference**
  Based on an "approach paper" and a workshop to identify topics and focus the evaluations

An evaluation designed to draw lessons from experience needs to be interdisciplinary and comprehensive enough to generate the necessary perspective and insight for learning. Such evaluations will often be thematic or comprehensive in the sense that they focus on a specific type of development assistance.

The perspective could also be wide – focussing on country or sector programme – or narrow, looking at a specific aspect of a project or programme. (See Chapter 2.1)

This type of evaluation should be implemented by external, independent evaluators. However, since the purpose is to generate experience more than ensuring accountability, personnel from Danida and partner institutions should participate more actively than in documentation evaluations.

Such broader participation is essential to ensure that the experiences generated through evaluations are applied in the formulation of policies and strategies, and in the design and initiation of new development activities. Broader participation will also help ensuring that the evaluations are relevant with respect to the problems and challenges that Danida and partner organisations are facing today.

The quality and implementation of development assistance is a matter of concern for both the donor and the partner organisations. The extent to which evaluations seek to accumulate experience should therefore be determined jointly by Danida’s management, the Regional Department, the internal Technical Advisory Service (TSA), the Embassies, and the relevant authorities in the partner countries.

These evaluations will therefore be initiated on the basis of a process to determine which areas or topics would be relevant. The first step is to formulate an “Approach Paper” which provides an overview of possible topics and angles. This will be followed by a workshop designed to identify topics and focus the evaluation. Such workshops can be internal, or include representatives from the partner country and external expertise. Based on the recommendations of the workshop, the Terms of Reference for the evaluation will be formulated by the Evaluation Secretariat, and the evaluation carried out.

1.5 DOCUMENTATION EVALUATION

Evaluations are analytical assessments addressing results of public policies, organisations and programmes, which emphasise reliability and usefulness of findings. A key role is to improve information and reduce uncertainty, since much of the information generated when development activities are implemented relies significantly on judgement.

Documentation evaluations aim specifically to account for aid through making operations and results transparent. A main concern is to produce reliable information. They therefore require impartiality and independence, and a relatively high degree of precision. They must be carried out in a systematic manner, building on a combination of scientific methods, expert knowledge, and common sense.

Documentation Evaluations

- **Aim**
  Produce reliable information
- **Target group**
  The public and political decisionmakers

- **Evaluators**
  External, impartial, independent
- **Process**
  Data collection and analysis phase followed by a verification phase

Documentation evaluations are defined to provide an adequate coverage of geographic areas, major and minor partner countries, and different sectors. They will be carried out by independent external expertise. They will be based on a specific Terms of Reference specifying the issues on which to focus.

The Terms of Reference is the key instrument to guide the evaluation. However, it should always be formulated with a flexibility that encourages the evaluators to include other topics not included in the Terms of Reference if the analysis indicate that this is important.

Documentation evaluations are typically implemented in two phases. Firstly the evaluators undertake a preliminary “desk review” based on studies of documents, interviews in Denmark, and interviews in the partner country. The report provides hypotheses and preliminary conclusions, and it identifies areas that need to be further documented.

On this basis, a detailed plan for the evaluation is developed by the evaluators and Danida’s Evaluation Secretariat. The plan focuses on the information that the evaluation team should collect while in the field. It could involve different evaluations tools, such as further document studies, interviews, collection of primary data, field observations, focus group studies, user studies, etc.

In the second phase, when information has been collected and analysed, the evaluators draw their conclusions and formulate recommendations. The draft report is then discussed with relevant staff in Danida and the partner countries. Normally, controversies may occur, and this phase is important in sorting out controversial issues, supplementing necessary information, correcting factual errors and possible misunderstanding, etc., before the final report is written.

The information provided and the views expressed in the final report are solely those of the evaluators and not those of Danida or the partner countries.

The target group for documentation evaluations is the public and the political decision-makers. The evaluation report should therefore be written in a down-to-earth language, avoiding as far as possible specialised terminology.

1.6 THE MAIN PARTIES TO AN EVALUATION

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

Evaluations are commissioned by Danida through its Evaluation Secretariat, jointly with other donors, or with the partner in the partner country. The evaluation is normally part of an annual evaluation programme established after consultations with the departments involved. The evaluation programme usually covers a range of evaluations: policy, programme and project activities as well as sectors, themes, and cross-cutting issues. (See Chapter 2.1).

In some cases several donors collaborate to undertake evaluations. Joint donor evaluations are common in case of co-financing. It is useful to reduce the administrative burden on the partner and to enhance usefulness for donor agencies. Usually one of the collaborating agencies will be responsible for commissioning the evaluation study. (See Chapter 2.7).

The Main Parties to an Evaluation

- **Commissioner**
  Danida, other donor agencies (joint evaluation) or partner authorities
- **Evaluators**
  External expertise from private or public institutions
- **Users**
  Decision-makers with a formal role in following up the evaluation, policy-makers, auditors, etc.
- **Stakeholders**
  Beneficiaries, affected institutions, individuals, etc., with an interest in the evaluation

The *evaluators* are those organisations or individuals collecting and analysing data and judging the value of the evaluated subject. These are commonly selected among Danish or international research institutions, directorates, consulting firms, industry, etc., to ensure independence and impartiality.

They may be supplemented by Danida internal staff members and representatives of partner authorities as resource persons. Involving internal staff provides an opportunity for learning by doing and to strengthen skills and capabilities in Danida and partner organisations.

The *users* of evaluations may be the policy makers; Danida policy and programme managers with staff; auditors and all parties with a formal role in relation to the development activities under evaluation. In order to have an impact on decision-making, evaluations must be relevant to the users and responsive to their particular needs for information. Evaluations must also be timely in the sense that they should be available at the time which is appropriate for the decision-making process.

The *stakeholders* are those individuals or organisations that informally have an interest in the policy or programme being evaluated and the findings of the evaluation. Stakeholders may sometimes be the same as users. In order to ensure usefulness of evaluations, the views and expertise of groups affected ought to be considered and integrated whenever appropriate.

Stakeholders can for instance be involved through steering or advisory groups. Participatory evaluation methods can be used to create consensus and ownership in relation to the development activities. Dialogue with stakeholders can help improve understanding and responsiveness to their needs and priorities. Feedback to both users and stake-holders is essential to make evaluations useful.

1.7 GUIDING PRINCIPLES FOR EVALUATIONS

The quality of evaluations can be judged differently. Some of the most commonly used criteria are credibility, impartiality and independence, as well as cost effectiveness. To be useful, evaluations must be relevant and timely.

Credibility

The credibility of evaluations depends on the expertise and independence of the evaluators and the degree of transparency of the evaluation process. It requires that the evaluation report successes as well as failures. Partner countries should participate in the evaluation in order to promote commitment.

Credibility also requires that a sufficiently large resource pool of professional evaluation staff exists to provide enough expertise in the various fields of activities and to ensure a certain degree of market competition among evaluators.

Credibility also requires transparency of the evaluation process. This applies both to the planning and implementation of the evaluation itself, as well as making the results of the evaluation widely available.

Impartiality and Independence

Impartiality contributes to the credibility of evaluations and the avoidance of bias in findings, analysis and conclusions. Independence provides legitimacy to evaluations and reduces the potential for conflict of interest which could arise if policymakers and managers were solely responsible for evaluating their own activities.

The evaluation process should therefore be impartial and independent in its function from the process concerned with policymaking, the delivery, and the management of development assistance. The requirement for impartiality will also apply to all stages of the evaluation process, including the planning of the evaluation programme, the formulation of the Terms of Reference, and the selection and approval of evaluation teams.

Involvement of internal staff as resource persons in evaluations will not compromise the requirement on impartiality and independence. The team members are external and independent, and the evaluation report is issued in the name of the authors.

Cost Effectiveness

The cost of evaluations is usually a minor part of the budget of the development activities under study. Evaluations may have considerable effect on management of aid activities and a substantial cost saving potential in terms of improved performance.

The key value of evaluations is that they allow for in-depth study of performance and independent assessment of the effectiveness of other performance management instruments. Potential benefits are the greatest for large policies and programmes. On the other hand, experience shows that evaluations can be too costly and time consuming compared to their real use and effect.

As a principle, benefits of evaluations should out-weigh their costs. Both costs and benefits can be largely affected by careful focusing of evaluations and by choosing the appropriate evaluators and the best suited evaluation methods.

2.1 EVALUATING DIFFERENT TYPES OF AID

Danida's evaluation programme for bilateral aid includes sector evaluations, evaluations of country programmes, project evaluations, thematic evaluations and evaluations of mode of co-operations. These are either designed for documentation purpose, for extracting experience, or both. Some evaluations are conducted jointly with other donors or partner governments.

The trend in Danida in recent years to shift the focus from evaluation of project aid towards more complex modes of development assistance – for instance sector, country-wise or thematically evaluations – has generally been considered positive. For instance, they have proved to be increasingly cost effective and with more impact on policy decisions.

Sector and thematic evaluations generally have considerably larger impact on decision-making than single project evaluations:

- *The Danida evaluation of the agricultural sector in 1994 resulted in Danida formulating a separate policy for support to agriculture, and a considerable strengthening of support to this sector*
- *The WID-evaluations in 1987 and 1994 resulted in a more consistent approach to include the gender aspect in mainstream development interventions*
- *The environmental evaluation in 1996 established a failure to operationalise Danida's environmental policy and strategy, and resulted in a renewed focus on environmental impact analysis*
- *The poverty evaluation in 1996 caused an increased attention to poverty analysis and monitoring*

The emergence of new types of aid interventions and the focus of specific priority themes in development aid poses new challenges to aid evaluators. Some of those challenges are methodological: how to evaluate support for good governance, human rights, civil service reform and privatisation.

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

In particular, it is a challenge to decide on the most effective use of evaluation resources and to find workable ways and means to focus evaluations correspondingly. Some of these challenges, as well as differences and similarities in various types of evaluations, are discussed in Chapters 2.2 – 2.7 below.

Types of Evaluations

- **Project/Component**
  Evaluation of individually planned undertakings designed to achieve specific objectives within a given budget and time period.

- **Sector**
  Usually applied cross-country wise.

- **Sector Programme Support**
  Evaluation of a coherent set of activities in terms of policies, institutions or finances

- **Aid forms**
  Evaluation of a specific instrument or channel for development aid (research, NGO’s, humanitarian aid, balance of payment support, technical assistance, etc.). Usually applied cross-country and cross-sector wise.

- **Country**
  Evaluation of the combined cross-sectoral Danish support to a partner country, notably one of Danida’s programme countries.

- **Thematic**
  Evaluation of selected aspects of different types of development aid (poverty, environment, choice of technology, gender aspects, sustainability, etc.).

- **Joint**
  Evaluation of large parts of development activities where Danish aid represent only a share of the total. Conducted jointly with other donors and the partner.

2.2 PROJECT EVALUATION

A development project is defined as a planned undertaking, designed to achieve certain specific changes within a given time through the use of specified resources. The limited scope of the project makes it relatively easy to focus the evaluation. That is, if the objectives against which the project should be evaluated are specifically expressed and in quantifiable terms.

A complicating factor is that the objectives are often complex, confounding, unrealistic, ill-defined, or may be changed as the project develops. This problem is discussed further in Chapter 5.2.

The success of projects in industrial countries is usually measured in a limited perspective. The main issue is whether outputs have been achieved quantitatively in time and within budgetary limits. Development projects are usually evaluated in a wider perspective with a main focus on their objectives, in accordance with the model presented in Chapter 4.1. The objectives of a project go beyond what is immediately achievable, and the chance of success is therefore correspondingly reduced.

However, project evaluation even goes beyond what is planned and desirable (the objectives) and attempt to judge the project also on the basis of its foreseen and unforeseen effects in society (the impact).

The reason for such a broad perspective in evaluations of projects is essentially the need to learn from experience and improve upon the design of similar projects in the future. However, it will often limit the chances that the overall assessment of the project is favourable.

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

The above mentioned also applies to evaluations of sector programme components.

2.3 SECTOR EVALUATIONS

These evaluations focus on several Danida-supported projects within a given sector in one or several partner countries. This is a cost-effective way to evaluate clusters of smaller projects together with larger projects. Sector evaluations will normally cover completed as well as on-going projects.

A sector may be defined broadly: the education sector, or more narrowly: vocational training. Irrespective of the definition, a sector evaluation will cover policies and strategies applied and a sample of projects representing typical interventions.

A main advantage of sector evaluations is that they tend to have considerable impact on accumulating experience and developing policies.

2.4 SECTOR PROGRAMME EVALUATIONS

The shift of Danida policy in 1992 to a move from individual projects towards sector programme support represents a new challenge for sector evaluations. Project evaluations focus on an a priori defined cause-effect chain leading from project inputs to outputs and fulfilment of objectives . In sector programme support, however, the development of institutions becomes a key facilitating factor for development. Sector programme evaluations must therefore focus more on questions of institutional performance, processes, changes and interrelationships, as well as the development in the sector involved.

Sector Programme Evaluations

Sector programme evaluations are more complex than project evaluations. There is not a single set of consistent objective, and there are often several donors involved.

Ownership and partner responsibility are key issues in sector programme support. This would make it necessary to involve the partner more actively in evaluation. The concern of the donor is to improve aid delivery and be accountable for aid provisions. The concern of the partner institutions is to improve development in the sector.

A key issue in evaluation is therefore how to secure involvement of the partner in sector programme evaluations and ownership to the results. In line with the emphasis on development partnership, local ownership and good governance, the donors will increasingly have to use joint programming of sector assistance as a vehicle to assist in developing an evaluation culture by building joint evaluations into sector programmes. To succeed in this, Danida will need to co-ordinate joint evaluations and ultimately let evaluations be co-ordinated by the partner countries.

A complicating factor is that there will often be several donors involved in the same sector. This makes it difficult to identify and assess each donor's individual contribution. However, if the evaluation is carried out jointly by several donors, it will reduce the administrative burden on partner authorities and organisations and be more cost-effective than individual evaluations.

2.5 COUNTRY EVALUATIONS

Country programme evaluation focus on the entire Danish aid to one of the main partner countries (programme countries).

Such evaluations may have as its primary aim to document past activities and performance in order to provide Danida and the partner government with a basis for deciding on possible changes in the co-operation strategy, country programme and sector interventions. The primary aim may also be to compile experience of general and specific character, which may be useful to improve future assistance to the country or to gain general experience with development support to national country programmes.

The evaluations are truly cross-cutting in the sense that they cover all sectors and modes of co-operation with the country. Focussing the evaluation will therefore be equally difficult. This also highlights the need that the evaluation team is interdisciplinary, with expertise reflecting the issues to be focused by the evaluation.

Country Evaluations

Country 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 aid.

The country evaluation will look 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 will review the instruments used in the bilateral co-operation, the modalities, and the relative weight given to assistance for economic and social development. The field work will normally focus on all levels, from policy level to project level, and include observations and interviews with government officials, project staff, beneficiaries and affected parties as appropriate.

Considerable work will therefore be invested in focussing the evaluation and make it relevant for all users and stakeholders.

Country evaluations are important for policy planning at the highest level and to provide a basis for country strategies and bilateral negotiations. They also represent an opportunity to focus on specific relevant issues such as the country's dependency on aid, institutional capabilities, democracy, human rights, gender, and environment in a wide context.

2.6 EVALUATION OF AID FORMS

Evaluations of aid forms focus on the various modes and funding channels of development assistance, such as balance of payment support, project support, technical assistance, export promotion, credit schemes, NGO's, etc. Such evaluations are usually cross-sectoral and cross-country and designed to assess the efficiency of a particular aid form as a means of promoting development.

Evaluations of aid forms may be initiated to document past activities or to extract experience to improve future performance. The focus will usually be on the efficiency of aid delivery, but also to some extent on the effect of aid delivered under the given mode of co-operation.

The methodological approach will vary according to which form of development aid being considered. Evaluating humanitarian assistance will for instance require quite a different approach from evaluating research assistance or balance of payment support.

Evaluation of Aid Forms

The focus is on efficiency and effectiveness of specific aid forms. The evaluation is complex in the sense that it may involve many sectors and countries, as well as a multitude of different objectives.

One problem is that the individual cases to be studied will often have very different objectives. This will make it difficult to explore the effects of development assistance systematically. It will make it even more difficult to explore the impact of aid. The focus will therefore usually be more on the performance of the various parties involved in disbursing development assistance and making use of the resources made available.

Evaluation of aid forms therefore has a rather narrow perspective while at the same time allowing experience to be drawn from a larger sample of aid interventions. This makes it possible to draw lessons that can have wide ranging implications both in terms of policy and in making specific improvements in the bilateral co-operation.

2.7 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 aid over the years. These themes are often borne out of policy statements and often termed “cross-cutting issues”. This implies that they are to be systematically integrated in all sectors of aid, in all stages of the project cycle and in all aid forms. Some of the factors that have been highlighted in recent years are the 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.

Thematic Evaluations

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

This has provided a wealth of information which may be useful when a thematic evaluation is undertaken.

In a thematic evaluation the subject may be analysed both in the context of Danida's development policy, but also international conventions, and the partner's priorities and strategies. Clearly there will be methodological constraints in distinguishing between the impact of Danish assistance vis-à-vis the impact of other donor initiatives, national activities and policies.

Usually, thematic evaluations cut across different sectors, countries and aid forms. As illustrated above, these evaluations represent the highest level of complexity of the type of evaluations employed by Danida. For that reason, they have proved useful instruments in providing experience and recommendations at the highest level of aggregation, i.e. the policy level.

In the preparatory phase, a main methodological challenge will be to draw a sample of development activities that can reflect Danish development aid both in terms of countries, sectors and forms of aid. Thematic evaluations are usually based on relatively large samples of development activities implemented over a relatively long period of time. Much of the study will therefore have to be based on older documents on file which may be incomplete or pay inadequate attention to the theme under study. Field work will essentially concentrate on the more recent development activities and be aimed at verifying information and acquiring a deeper understanding.

2.8 JOINT EVALUATIONS

The changing modes of 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 and on partner country leadership roles. Effective evaluations of such new types of aid call for more collaboration. Clearly, it would be a waste of resources on both the donor and partner side if several donors performed individual evaluations of their sector support to the same sector.

A joint evaluation is not a specific type of evaluation. A main challenge is to involve the partner at all stages.

Joint evaluations have focussed vital areas and helped consolidating international responses and development policy:

- *The Nordic agricultural program in Mozambique (MONAP) was evaluated jointly by Denmark, Norway and Sweden in 1991 (Aid under Fire), 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, in particular the country programming process.*
- *The international response to conflict and genocide in Rwanda was evaluated in 1996 jointly by 38 bilateral donors, UN agencies and NGOs. It identified key issues and problems of international humanitarian assistance and has induced changes of policies and strategies of humanitarian agencies.*

The aim is to ensure that the evaluation becomes an efficient learning tool, helps promote good governance, enables the partners to be fully accountable, and be cost-effective.

The aim is to ensure that the evaluation becomes an efficient learning tool, helps promote good governance, enables the partners to be fully accountable, and be cost-effective.

Joint evaluations do take place but are still infrequent and primarily concern donors' jointly financed programmes. Common to most of these evaluations is that developing countries have played a minor role, if any, in their planning and execution.

In joint evaluations including larger groups of donors, the involvement of the majority is typically limited to a few key stages such as review and approval of the design. Management and co-ordination is typically delegated to one or a few donor agencies taking a lead responsibility or to an external organisation designated to manage the study. It will be a considerable challenge for donors in the near future to organise joint sector evaluations in cases where several donors are supporting the same sector. In this endeavour it will be useful to draw on experience from previous joint evaluations of programme support. A key challenge will be to secure local ownership for such evaluations by focussing on what is relevant for the partner.

A key advantage of joint, multi-donor evaluations is that they have greater credibility and broader ownership of findings among the broader donor community than would be the case with single donor evaluations. Their value lies in the comprehensiveness of findings and thus in the validity of lessons learned. They generally yield useful lessons, focus policy discussions among donor agencies, and achieve an impetus for joint action that is out of reach of smaller evaluations.

Co-ordination of aid interventions among donors is notoriously difficult. The same applies to evaluations. Evaluation programmes are prepared in response to agency needs for lesson learning and accountability, and are geared to the planning and programming cycle of the agency. If evaluation is 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.

3.1 DANIDA'S EVALUATION PROGRAMME

Danida’s evaluations are planned on the basis of a rolling three year programme. The programme is updated annually by the Evaluation Secretariat on the basis of the status of the existing evaluation programme. All relevant departments in the ministry and the embassies in partner countries are consulted and invited to propose evaluations to be included in the programme.

According to Danida’s evaluation policy, the programme shall cover:

- ongoing or completed projects and programmes
- Danish aid to a particular country, one or more sectors in a partner country or a sector in a number of partner countries
- aid instruments (project and programme aid, aid via NGOs, mixed credits, balance of payments support, emergency aid, research support, fellowship programmes)
- multilateral aid, primarily in co-operation with other donors

The Evaluation Secretariat ensures that the programme reflects the distribution of Danish aid on countries, sectors, aid forms, and special priority areas.

The evaluation programme contains a brief rationale for each evaluation; the primary objective (documentation or accumulation of experience), the target group, and which features in particular the evaluation shall comprise.

Documentation evaluations are planned in order to achieve a suitable coverage of geographical areas, large and small partner countries and sectors over time.

Evaluations to accumulate experience are planned on the basis of the strategic and operational needs of the aid administration. They can cover the above subjects but will, in addition, often incorporate multidisciplinary subjects (i.e. strategies, capacity building, planning and implementation of projects and programmes, consideration of the environment, gender aspects) or elements of these.

When the programme has been approved, the individual desk officers in the Evaluation Secretariat will be assigned the responsibility for specific evaluations.

The evaluation programme will be distributed widely to Danish and international consulting firms which may then indicate their interest for competing for contracts for specific evaluations.

The main steps in the evaluation process from the initial decision, until the final report is made public, are described in Chapters 3.2 – 3.9.

3.2 DECISION TO LAUNCH THE EVALUATION

The Evaluation Secretariat ensures that all involved parties are informed about the evaluation in due time before it is initiated. In the case of larger evaluations this is done through a brief 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 (methods to be applied)
6. Preliminary work plan with a time schedule

The approach paper should indicate the extent to which representatives of the partner will be involved in preparation and implementation of the evaluation, as well as possible co-operation with other donors.

The approach paper will be distributed to Danida's Regional Departments, the Embassies, and the partner for information and comments.

Where there are many interests involved, an internal 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 should reflect the topic and purpose of the evaluation and be decided in consultation with the Regional Departments, the Embassies and the partner.

3.3 THE TERMS OF REFERENCE

The Terms of Reference provides the background for the evaluation, defines its objectives and the scope of work, the composition of the evaluation team and the timing.

The Evaluation Secretariat is responsible for drafting the Terms of Reference based on the approach paper, and to ensure that the evaluation is commensurate with Danida's evaluation policy and approved practices.

It is essential to ensure that the need for information of all involved parties is considered. This may require a workshop or seminar which involves the Evaluation Secretariat, the Regional Department, the Embassy involved, the evaluation team, Danida's Technical Advisory Service, the partner country (the line ministry), the project management (in project evaluations) and other donors (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 will 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

Larger evaluations may be carried out in several phases: desk study, field work, and synthesis. In such cases an overall Terms of Reference provides the framework for the evaluation. Part of the objective of the desk study is to formulate detailed Terms of Reference for the field work.

3.4 ASSIGNING THE EVALUATION TEAM

The evaluation team may be established by Danida engaging individual consultants or, as is typically the case, by a consulting company to which the evaluation is contracted.

The composition of the evaluation team is also carried out in consultation with the involved parties. The Evaluation Secretariat has the initiative. The criteria for the selection of the evaluation team are professional competence, impartiality, and experience in relation to the task. The professional competence must be relevant for the topic under study. Professional expertise from the partner country should as a matter of principle be represented on the team. The participation of representatives from the partner country will improve the quality of evaluation. It also increases the credibility of evaluation findings for host country managers, promotes the acceptance of findings and recommendations.

Danida's evaluation policy also encourages a closer involvement of Danida staff and partner officials in the evaluations. This can be as resource persons or by actively participating in specific phases of the evaluation. This increases the relevance of the resulting report, and promotes the utilisation of findings.

The success of an evaluation ultimately 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 spokesman on the team.

In order to develop and strengthen the evaluation resources in Denmark and partner countries, it may be useful to include trainees on the team. Consulting firms will often on their own initiative include trainees in their proposals. The language used to communicate evaluation findings need special attention, and in larger evaluations it may be necessary to include a copy editor in the team.

3.5 PRELIMINARY STUDIES

There is usually a wealth of information available, and it is worth while to do a thorough review of this information. This will provide an insight to the topic under study and make it possible to establish facts, describe processes, and establish some key hypotheses.

The evaluation's field work will usually not permit the evaluation team to develop or collect all necessary data. Collection of data will therefore have to be organised and partly completed well in advance of the team's arrival in the field. This would include existing data, interviews, surveys or special studies. Much of this can be done by local consultants, the institutions under study or the responsible line ministry. The advantages are that local resources are utilised and developed, the team can focus more on analysis than data collection during field work and costs are reduced.

Data collection in Denmark will usually focus on existing data, interviews with relevant Danida departments, and other institutions that could provide relevant material or insight. The team is expected to make use of Danida's documentation centre, which operates a computerised bibliography, and is on-line with the most important data bases.

Documentation studies abroad may involve interviews at the Embassies, the institutions under study, the Ministries, with target groups and affected groups, etc. It may also involve larger surveys, impact studies, tracer studies, etc. Danida is prepared to finance, if required, local research institutions or the like to do such preparatory research. Such studies need to be endorsed by the Evaluation Secretariat.

3.6 FIELD WORK

The core activities during field work will be to analyse available information, and to test and verify hypotheses and preliminary conclusions through interviews with key informants and groups of individuals. This presupposes that most of the data have already been collected.

In connection with evaluations to accumulate experience, Danida management and staff plus the relevant representatives of partners should be involved in the implementation of evaluations to the greatest possible degree to participate in the analysis and in the formulation of lessons learned, conclusions and recommendations.

On the basis of documentation studies and the information gathered, the evaluation team presents the problems and/or the hypotheses with a view to discuss and analyse the material produced together with relevant staff. This will often take place in a number of workshops or seminars in the course of the evaluation.

The evaluation team will have a catalytic role rather than performing an independent analysis.

In evaluations with a documentation objective, the participation of Danida staff will typically be less intensive: to produce documentation and to participate in interviews, debriefings and discussion of draft reports.

In country and sector evaluations, representatives of the partner should be involved to the extent possible, partly in order to have their views reflected, and partly to promote local involvement and follow up of evaluations.

3.7 DRAFT REPORT

The evaluation thus presupposes active participation of Danida and other involved parties. It is nevertheless the responsibility of the evaluation team to formulate the report. Major differences in opinion between the evaluation team, the partner authorities, and Danida should be included in the report, either as remarks in the text or as separate paragraphs.

The draft report should usually be presented to the Evaluation Secretariat two-four weeks after the field work is completed, depending on the type and scope of evaluation. The report shall fulfil the requirements in the Terms of Reference.

The Evaluation Secretariat will then review the document and decide on what is needed to finalise the report. This may require one or more meetings between the evaluation team, the Evaluation Secretariat, and relevant Danida departments. It could also be done through a more comprehensive workshop.

In all evaluations, a debriefing should be made in the partner country including partner authorities, the Embassy, and other donors. The main intention would be to promote involvement and follow-up of the evaluation.

The final title of the evaluation report is to be endorsed by the Evaluation Secretariat. At the same time, a plan for disseminating the final report should be prepared.

3.8 FINAL REPORT

The team leader is responsible for the evaluation report. The individual team members will usually prepare their separate parts, which will then be edited together by the team leader. He/she is also responsible for quality assurance and that Terms of Reference is fully covered in the report.

Making the report easy to read should have a high priority. Complicated technical discussions should be referred to the annexes. The main report should only be about 40-50 pages. The text should be reviewed and edited to ensure an easily understandable language and effective communication aids. See also Chapter 7.3. When relevant or necessary, the team should involve a language editor during the final write-up of the report.

Evaluations that may be of interest to a broad audience, should be published in a Danish and possibly abridged version, edited specifically for non-professional readers. Such a popular version should still be loyal to the original version. It must be endorsed by the team leader before it is published. Preparing such a Danish version should be part of the evaluation contract.

The final report is submitted to the Evaluation Secretariat when an agreement on its form and presentation has been reached. This should include a summary both in Danish and English, and a summary of recommendations for the follow-up memorandum. The final report should be in accordance with Danida's layout for evaluation reports. The layout is available from the Evaluation Secretariat.

3.9 FOLLOW-UP AND PUBLICATION

After the final report has been received, a follow-up memorandum is prepared which summarises Danida's response to conclusions and recommendations. The Evaluation Secretariat will periodically check if agreed action has been taken.

The Evaluation Secretariat will promote the dissemination of the experience gained through workshops and seminars, in co-operation with the Ministry's training section. Seminars can also be held in the partner countries on the basis of the final evaluation report if the result of the evaluation is considered of wider importance.

All Danida's evaluations are published both in printed reports and are made openly available electronically through the web-site of the Ministry of Foreign Affairs.

The Evaluation Secretariat will also undertake an internal assessment of the performance of the company or institution responsible for the evaluation as a reference for future evaluations and composition of evaluation teams.

Finally, the Evaluation Secretariat will make an assessment of the outcome of the evaluation and list the favourable and unfavourable experiences as an input to future planning and implementation of evaluations.

4.1 EVALUATION CRITERIA

Danida's definition of evaluation is identical to the definition recommended by OECD/DAC. The same definition has been adopted by all major donor agencies internationally. It stipulates five specific evaluation criteria that should be used in assessing development interventions: efficiency, effectiveness, impact, relevance, and sustainability. These are generally applicable analytical measures that can be used on all administrative or aggregation levels – whether project, programme, sector, or country level – and all forms of aid whether commodity support, technical assistance, credit schemes, etc.

Taken together, these five criteria used in combination should provide the decision-maker with the essential information and clues to make correct diagnosis and determine what should be done next.

Aid interventions can be viewed in different perspectives. The operational perspective is the narrowest. It is concerned about the delivery of aid and is the perspective of the implementing party; for instance, whether the agreed outputs have been produced in a project or funds have been disbursed as planned in a credit scheme. This is measured in terms of efficiency.

The wider and more ambitious perspective is the tactical perspective. This is the perspective of the users. It is concerned about the next step in the sequence from the input of funds and resources to the fulfilment of objectives. In general terms this is measured in terms of effectiveness.

The widest perspective is the strategic perspective. This is assessing the aid interventions from the society's point of view. It takes into account not only needs satisfaction of the primary beneficiaries but also its impact on other groups in society.

The strategic perspective will also have to consider the relevance of the aid intervention, i.e. its consistency with needs and priorities in general. Finally, the sustainability, or the long-term effects will have to be taken into account.

In total, these five criteria cover all three perspectives. There are several methodological challenges for the evaluator. The first is to disaggregate these general criteria to specific evaluation questions relevant to the situation under study. The second is to find the answers to these questions on the basis of reliable information. The third challenge is to provide an aggregate conclusion on the basis of these answers to each of the five evaluation criteria. Chapters 4.2 – 4.6 discuss how to apply the evaluation criteria in an evaluation.

Evaluation Criteria

- **Efficiency**
  The productivity of the implementation process
- **Effectiveness**
  The extent to which the objective has been achieved
- **Impact**

All positive and negative changes and effects caused by the aid intervention

- **Relevance**
  Whether the objectives are still in keeping with the donor's and local and national priorities and needs

- **Sustainability**
  Whether the positive effects will continue after external support has been concluded

4.2 MEASURING EFFICIENCY

Efficiency is a measure of the "productivity" of the aid intervention process, i.e. to what degree the outputs achieved derive from efficient use of financial, human and material resources. In principle, then, it means comparing outputs against inputs. It takes the limited operational perspective of the key implementing institutions.

For example, in a road project efficiency is measured in terms of the construction of the road (the physical output); based on its quality, cost and the timeliness of construction. In a sector programme to support agriculture, it could be measured in terms of disbursement of loans to local farmers, development of policy and institutional support structures, etc.

The main difficulty in measuring efficiency lies in determining what standard to adopt as the reference. This will depend on the type of aid intervention, what can realistically be expected under prevailing circumstances, etc. The standard will rarely be predetermined, so in practice it will be up to the evaluation team to specify what standard they are measuring against, for instance:

- as against comparable aid activities
- on the basis of best practice standards
- on the basis of specified criteria of what is considered reasonable.

Efficiency will usually be a concern already at the appraisal stage, when budgets and inputs are being decided. It will also be a concern when activities are implemented.

In the course of events, relevant information will therefore usually be available to undertake preliminary assessments of efficiency. The final verdict, however, can only be made after the aid components have been delivered.

In a project, efficiency will be measured as an aggregate of the efficiency associated with individual outputs. In larger evaluations at sector or country level the efficiency will have to be expressed as an aggregate of the efficiency associated with all the aid interventions involved.

Measuring Efficiency

- **What to measure**
  The delivery of aid
- **Who's perspective**
  The implementers
- **Point of reference**
  Similar interventions/best practice standards
- **Methodological challenge**
  What standard to use as reference
- **Key questions**
  To what degree have aid components been delivered as agreed? Could it have been done better, more cheaply and more quickly?

4.3 MEASURING EFFECTIVENESS

Effectiveness is a measure of the extent to which the formally agreed objectives have been achieved, or can be expected to be achieved. It takes the perspective of the users or those formally designated as the beneficiaries of the aid interventions: the target group.

For example, in a road project, effectiveness could be measured in terms of traffic flow, or transport of different commodities and the users of the road. In a sector programme to support agriculture, effectiveness could be measured in terms of the scope of productive investments made by the farmers involved, development of human resources through training and institutional development in the sector.

Assessing effectiveness is also a key issue when aid activi-ties are implemented. The necessary information is therefore likely to be available when the evaluation is carried out.

However, assessing effectiveness presupposes that the objectives have been unambiguously and realistically defined so as to make verification possible. Often, however, this will not be the case: objectives may be unclear and highly general, such as “improved health conditions” or “improved living conditions”.

In such cases, the evaluation team will have to try to op-erationalise the objectives on the basis of existing documentation.

If objectives are unrealistically ambitious seen in relation to the resources applied, the scope of activities and the time perspective, the evaluation team would have to stipulate what would seem to be a realistic objective as the point of reference for measuring effectiveness.

Measuring Effectiveness

- **What to measure**
  Achievement of objectives
- **Who’s perspective**
  The target group
- **Point of reference**
  Agreed objectives
- **Methodological challenge**
  Unclear, multiple, confounding or changing objectives
- **Key questions**
  To what extent have agreed objectives been reached? Are activities sufficient to realise agreed objectives?

In a project the immediate objective can usually be expressed in precise and verifiable terms. In aid interventions at sector or country level, where several development activities are involved, a large number of objectives may be involved. The analysis of effectiveness will be correspondingly complex and will have to be expressed in terms of an aggregate of the assessment of effectiveness in the individual aid activities.

Usually, the development intervention is only one of several factors likely to contribute to achieving set objectives. The evaluation team should therefore pay attention to all major causes behind fulfilment of objectives. This includes the external or contextual factors that also influence the chances of success. Here attention should be paid to whether these may have changed to such a degree that reformulating the objectives should be considered.

4.4 MEASURING IMPACT

Effectiveness is a narrow evaluation criterion, that focuses only on specific positive and planned effects expected to accrue to the involved parties, expressed in terms of the immediate objective. By contrast, the concept of impact is a far broader one, as it includes both positive and negative consequences, whether these are foreseen and expected, or not.

A broad assessment of impact is essential in a comprehensive evaluation. It will often be the most difficult and demanding part of the evaluation work.

In assessing impacts, the point of reference is the status of affected parties prior to aid interventions. The question to ask is: Which effects – whether positive or negative, expected or unforeseen – have come about as a result of the aid activities? These may be economic, social, political, technical or environmental effects – locally, regionally, or at the national level. We are now also concerned with effects for other affected parties than the target group, directly or indirectly, for men and women, ethnic groups, etc.

For example, in a road project, the impact may be measured in terms of economic benefits and losses that accrue to changes in traffic flow and transport capacity, through new settlement patterns, opening up new areas to agricultural production, increased pollution, etc. In a sector programme to support agriculture, impact could be measured in terms of increased productivity of farmers, positive and negative employment effects, environmental effects, economic dependency, etc.

The combined impact is the result of complex causal conditions that are difficult to analyse. It may be especially problematic to prove that observable changes can be ascribed to a given aid intervention. In order to understand causal connections, the evaluator may need to employ process-oriented methods such as participant observation.

Measuring Impact

- **What to measure**
  Intended and unintended positive and negative effects
- **Who's perspective**
  The society
- **Point of reference**
  Status of affected parties prior to intervention
- **Methodological challenges**
  Lack of information about affected parties
  Cause and effect linkages
- **Key questions**
  What are the positive and negative effects?
  Do positive effects outweigh negative effects?

Different impacts may appear at varying times, and it may be necessary to distinguish between the short-term and the long-term ones.

How much time must elapse until the impacts are felt will vary with the type of aid intervention. An agricultural project, for instance, may produce considerable impacts after only a few months – whereas a sector programme to support education will usually yield the greatest effects several years after it has been completed.

Attempts to assess impact are often done in interim evaluations before aid interventions are terminated. In such cases, it may be difficult to predict the long-term impacts as these can usually be evaluated only long after. Although such assessments of impact may be hinging on uncertain assumptions, they may still produce valuable insight which can be used to guide decisions.

4.5 MEASURING RELEVANCE

By relevance is meant an overall assessment of whether aid interventions are in keeping with donor and partner policies, as well as with local needs and priorities.

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

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 potentials. In a sector programme to support agriculture, relevance could be assessed in terms of domestic market responses to new crops, farmer's' responses to the various programme initiatives, etc.

In other words, relevance is basically a question of usefulness, and in turn leads to higher-level decisions as to whether the aid 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? This will have to be asked at various levels with reference to the partner country:

- At the highest level it concerns the relationship between the aid 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 next level it is a question of how aid activities fit into a larger context, e.g. in relation to other aid interventions and development efforts within a larger programme or sector.
- Then at the lower level it is a question of whether the aid activities are directed towards areas accorded high priority by the affected parties.

The question of relevance is a key issue from the earliest stage when aid interventions are appraised. It should also be a major issue throughout the implementation of activities.

Measuring Relevance

- **What to measure**
  Appropriateness in relation to policies, needs and priorities
- **Who's perspective**
  The society
- **Point of reference**

Needs and priorities of donor and partner

- **Methodological challenge**
  Lack of consensus regarding needs and priorities

- **Key questions**
  Are objectives in keeping with needs and priorities?
  Should the direction be changed?
  Should activities be continued or terminated?

4.6 MEASURING SUSTAINABILITY

Assessing the sustainability of development interventions involves more than merely asking whether it has succeeded in contributing to the objectives set. Sustainability is an indication whether the positive impacts are likely to continue after external assistance has come to an end.

The four evaluation criteria described above primarily concern the specific aid intervention, whereas assessment of sustainability is more a matter of the development process itself, in a longer-term perspective. This largely depends on whether the prerequisite local resources are available, and how partners view the aid interventions.

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

Thus, sustainability is in many ways a higher-level test of whether or not the aid 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.

The sustainability of any aid intervention will depend to a large extent on whether the positive impact justify the investments necessary and whether, say, the local community values the project highly enough to be willing to devote scarce resources to continuing it.

Measuring Sustainability

- **What to measure**
  Likelihood of benefits to continue
- **Who's perspective**
  The society
- **Point of reference**
  Projected, future situation
- **Methodological challenge**
  Hypothetical answers
- **Key questions**
  To which extent does the positive impact justify investments?
  Are the involved parties willing and able to keep facilities operational and continue activities on their own?

In other words, sustainability is concerned with what happens after the aid activities are completed and should ideally be measured some years afterwards. It will be difficult to provide any sure assessment of sustainability as long as activities are still underway, or immediately afterwards. In such cases, the assessment will have to be based on projections about future developments on the basis of whatever knowledge is available about the aid intervention and the involved parties. It will require an analysis of the contextual setting – its capabilities and restraints.

Experience of donor agencies conclude that a project's sustainability hinges mainly on six main areas, also called sustainability factors. These factors must be taken into account all along the implementation cycle, and should be used as a check list during evaluation to identify relevant evaluation questions related to the project's sustainability. A brief description of the sustainability factors is included in Chapter 6.4.

5.1 THE SCOPE OF EVALUATION

The scope of evaluation is a question of breadth of the study and its depth. The breadth of the evaluation is essentially determined by the type of intervention(s) the evaluation is intended to cover. A project evaluation has a narrow perspective and a country study a much broader perspective, as discussed in Chapter 3. The depth is essentially determined by the evaluation criteria applied, as discussed in Chapter 4, and the purpose of the evaluation. An evaluation to generate experience usually digs deeper than an evaluation designed to document achievements, as discussed in Chapter 1.

The less comprehensive evaluation would then be the one that focuses on a single project or credit scheme for the purpose of documenting achievements. The most comprehensive evaluation would be the one focusing on the development assistance to a country with the purpose to extract experiences.

The most common approach in evaluation of development assistance is the deductive approach whereby hypotheses are formulated at an early stage on the basis of common understanding and available information. These are then tested against reality when the evaluation is carried out. The advantage of this approach is that it takes as its point of departure existing experience, which makes it easier to choose indicators and interpret data. The main criticism against this approach is that it may delimit both themes and problem areas so that important aspects or conditions may be overlooked.

Both in single projects and in more complex types of aid interventions, objectives will often be unclear or inadequately formulated. Also, development interventions may have many unforeseen impacts, both positive and negative – which may well be overlooked if evaluation focuses only on what has been formally agreed.

Overcoming such problems will depend on the Terms of Reference and the experience of the evaluation team and their ability to define analytical frameworks for the investigation which can capture the major impacts of the activities in question.

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

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

5.2 REVIEWING THE STRATEGY

The intervention strategy is the starting point for the evaluator, since it identifies the anticipated achievements against which assessments will be made. Evaluations are primarily concerned about the broad overall assessment of development interventions, and less with the implementation of activities as such. In other words, the significance of the process, rather than the process itself.

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

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

Danish development assistance is planned according to the Logical Framework method, where the formally agreed objectives are established. If changes have been made in these objectives, they should be analysed specifically. Complex programmes may have many different objectives. In such cases, it may be useful to organise the objectives in a logical cause/effect sequence. This helps to provide an initial overview of the complexity of the strategy, which in many cases will reveal a picture of confounding, even conflicting, objectives. It will often demonstrate that the objectives have gradually changed from what was initially intended. This has often been done without being discussed explicitly and approved between the parties involved.

The analysis should consider whether the objectives are realistic. Often objectives are too ambitious seen in relation to the resources available and the agreed time frame of the aid intervention. Unrealistic objectives are a problem for two reasons. Firstly, if they are used as the formal reference for the evaluation the final assessment will be that the aid activities have failed. Secondly, the analysis necessary to assess the fulfilment of ambitious objectives would be unduly expensive. Also, the evaluation would not be able to demonstrate that the changes can be attributed to the aid intervention and to separate them from other factors that might have caused major changes.

In order to assess the realism of a project, the evaluators will have to consider the probabilities of fulfilling the various formally agreed objectives at the earliest time before the evaluation is planned. This would make it possible to identify realistic points of references when the scope of the study and the research methodology for the evaluation is chosen.

5.3 STUDY DESIGN

In research, the ultimate design to study the effects of an intervention is to establish the situation both before and after the intervention – in the target group and a control group as well. This is a precise procedure, but costly and complicated.

Evaluation is not research, nor is it hearsay. The evaluation must satisfy a minimum standard of quality of information in terms of validity and reliability (See Chapter 6.2). The time and resources available are usually severely restricted.

A commonly applied design in the evaluation of development assistance is to study the target group only in retrospect and without a control group that could help provide evidence of changes. Such a simple design does not provide a basis for drawing firm conclusions.

1. “Before and after”-study with control group.
2. (Complex, costly)
3. “Before and after”-study without control group.
4. (Simple, less costly)
5. “After”-study with control group.
6. (Simple, less costly)
7. “After”-study without control group.
8. (Very simple, low cost)

Four different study designs. The most commonly used design in evaluation is the simple “after”-study without control group.

The lack of reference data means that the evaluator must make judgements on the basis of general experience. Provided that this is adequate, the method may well yield useful information, but validity will remain low. Other supporting information must therefore be made available to verify the findings.

Because of this, the preferred method of information is tri-angulation. This means to compensate the use of simple data collection methods and a simple study design with the use of several information sources and different methods simultaneously to generate information about the same topics.

For instance, information from a simple survey may be supplemented with general experience data from similar interventions, and interviews with a variety of key informants to provide contextual information. In a situation that affects several parties with different interests, representatives of all parties as well as some neutral parties should be interviewed. This provides a triangulation effect which largely helps to verify information, cut through conflicting evidence and acquire insight in a costeffective way.

In some cases, for instance in health programmes, a more complex design may be used, with a separate study prior to the intervention (baseline study) against which to measure changes. However, experience has shown that baseline data often provide an uncertain basis for ascribing the impact of a specific aid intervention. An alternative is to use a control group after the aid intervention. The advantage is in its ease of application and cost. However, there is an ethical dilemma involved in studying a control group if this means that its members will not benefit from the aid intervention.

5.4 THE ANALYTICAL PROCESS

The process that takes place during an evaluation is depicted below. In the preparatory phase the first step is to translate very general evaluation criteria into specific evaluation questions, that relate specifically to the aid activities under study. There will always be several evaluation questions under each evaluation criterion, in order to substantiate conclusions and provide enough information for decision-making. The evaluation questions will constitute a set of indicators for the evaluation. The disaggregation process leading to the evaluation questions is discussed in Chapters 4.2 – 4.6.

The next step is to make the evaluation questions operational by turning them into evaluation instruments. In this context the evaluation instrument means the evaluation question plus the information collection method applied. The evaluator must ensure that the evaluation instruments produce valid and reliable information. This is discussed in Chapter 6.2.

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

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

The strength of the conclusions will rely on the amount and quality of underlying information. In some cases the analysis of information will reveal conclusions that are linked to other evaluation questions, or other evaluation criteria, than the one in focus. Such conclusion will add to the picture and can provide opportunities for verifying or rejecting other conclusions.

The analytical part of the evaluation is the one that transforms the detailed pieces of information into conclusions at a more aggregate level, and finally makes use of these to draw the overall conclusion at the highest level, linked to the evaluation criterion in question.

The evaluator 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 evaluator 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.

The purpose of analysis is to transform data into credible evidence about the aid intervention and its performance. Individual data are of little use unless they can be extracted or aggregated to address the questions in focus. This assumes that the evaluation design and instruments are appropriate, that the data are valid and reliable, and are analysed properly. In Chapter 5.5 some methods to extract information are discussed briefly.

5.5 COLLECTING INFORMATION

Evaluations often produce controversial results. Therefore, the evaluation itself is often under criticism, and the choice of methods in data collection and analysis may be in focus.

The basic distinction is between the *formal* methods typical of academic research and the simple, *less formal* methods.

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

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

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

In evaluation of development assistance, the most frequently used data collection method is interviews with key informants. This will inevitably produce subjectively biased information that will have to be verified. The evaluator therefore has a need for supplementary methods that maintain a reasonable level of precision without being dependent on very time-consuming procedures to collect and analyse vast amounts of data.

The quality of information, especially in terms of validity and reliability, should be a main concern for the evaluator. If the expected degree of precision is low, the evaluator may simultaneously employ various methods and sources of information in order to cross-validate data (triangulation). In the report, particular mention must be made of the choice of instruments and any limitations due to the quality of data and the analysis.

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. **Focus group interviews**
   For analysis of specific, complex problems, to identify attitudes and priorities in smaller groups. Reasonable in terms of cost, and efficient. Stimulates the generation of new ideas. Risk of one-sidedness on the part of participants and the moderator.

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

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

8. **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.

9. **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.

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

5.6 QUANTITATIVE AND QUALITATIVE

The level of sophistication of analysis is a matter of concern in evaluation. Tables, percentages and averages often give a clear picture of the sample data particularly for non-specialists, and many users will only be interested in this level of analysis. In addition, measures of spread, including percentiles and standard deviations, may add valuable information on how a variable is distributed throughout a sample population. There are also a wealth of more sophisticated research methods that can be applied. However, much of the evaluation work can be done using very basic methods. After all, evaluation is not scientific research and the quality of data and the size of samples are often such that it does not warrant sophisticated methods.

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

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

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

One of the main advantages of statistical analysis is that it can be used to summarise the findings of an evaluation in a clear, precise and reliable way. However, not all information can be analysed quantitatively.

For example, responses to an open-ended interview survey may provide lengthy descriptions which may be very difficult to categorise, let alone quantify, without losing subtle differences in the responses.

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

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

Qualitative data includes detailed descriptions, direct quotations in response to open-ended questions, the transcript of opinion of groups, and observations of different types.

Qualitative analysis typically include content analysis, analysis of case studies, inductive analysis and logical and sequential analysis. All methods may produce *descriptions* (patterns, themes, tendencies, trends, etc.), and *interpretations* and explanations of these patterns. The data analysis should include efforts to assess the reliability and validity of findings.

The analysis of qualitative data help broaden the view of the phenomena of interest in an evaluation, but can also increase depth and detail, where needed. The process of analysing qualitative information is often inductive and without any particular guiding theory.

Qualitative analysis rely on the evaluator's professional judgement concerning the relevance and significance of available data to evaluation issues at hand. The evaluator must therefore be both knowledgeable concerning the evaluation issues and attuned to potential biases.

Qualitative analysis is best done in conjunction with the statistical analysis of related (quantitative or qualitative) data. The evaluation should be designed so that the two sorts of analysis, using different but related data, will be mutually reinforcing.

6.1 QUALITY OF INFORMATION

The international community of evaluators have over several years developed what is now agreed to be a common standard of what constitutes quality in evaluations. The standards are intended to reflect the best practise among evaluators. They were set in order to encourage useful, feasible, ethical and sound evaluations, which in turn will contribute to improved interventions in diverse settings. The standards are organised around four important attributes of an evaluation: *utility* (to serve the users information needs); *feasibility* (to be realistic, prudent, frugal and diplomatic); *propriety* (to be conducted legally, ethically, and with regard to those involved and affected); and *accuracy* (to produce trustworthy, valid and reliable information).

These four attributes are necessary and sufficient for sound and fair evaluation. The standards relate to each other. An evaluation which is not feasible, is not likely to yield accurate conclusions, and conclusions that are not accurate are not likely to be used. Similarly, an evaluation which is conducted according to high standards of propriety will generally have much higher utility than one with short-comings in these respects.

The purpose of the present guidelines is to improve the quality of evaluation. A Danida evaluation is a selective investigation aimed at collecting and analysing data, formulating observations and making recommendations of practical relevance to the operation of Danish development assistance.

A main concern in evaluation is to learn about the factors causing success or failure of particular development activities. As much should be learned from successes as from failures, and the reports should present both types of findings in a balanced manner that reflects the situation under study. Evaluations should be critical, but criticism should be constructive.

Evaluations are not fault-finding, and the purpose is not to attribute praise and blame to individuals.

Evaluations must be useful. If they are not, they are a waste of money. The evaluation team has the primary responsibility to make sure that the evaluation confirms with high accuracy standards. It is those who have commissioned the evaluation, who process the draft and communicate the final report who have the main responsibility to make sure that the exercise is useful.

International Evaluation Standard

- **Utility**
  To ensure that an evaluation will serve the information needs of intended users.
- **Feasibility**

To ensure that an evaluation will be realistic, prudent, diplomatic, and frugal.

- **Propriety**
  To ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results.

- **Accuracy**
  To ensure that an evaluation will reveal and convey technically adequate information about the features that determine worth or merit of the program being evaluated.

The team leader of an evaluation team appointed by Danida is bound by contract to present a quality certificate to Danida when the final draft report is submitted. In this certificate the team leader will have to indicate the extent to which:

- Danida guidelines, including the present one, have been used
- the use of methods have been discussed explicitly
- the focus, analyses and results are according to standard
- the language, layout, illustrations, etc. are according to standard
- the professional input of team members is acceptable
- Some of the requirements to ensure evaluation quality are discussed in Chapters 6.2 – 6.5.

6.2 ENSURING VALIDITY AND RELIABILITY

The most basic aspect of information quality is the extent to which it corresponds to what we want to measure. Validity is the closeness or fit between and intellectual construct and the things we measure empirically as indicators or proxies for that abstraction. If our measurements do not correspond to the abstraction, then the entire chain of reasoning between the analytical and empirical domain collapses.

Ensuring *validity* is important in programme evaluation, since social concepts are often multi-dimensional. One specific measure might emphasise one dimension at the expense of another: for instance per capita net income as a measure of economic welfare as opposed to family or gross income.

Too often, a decision is based solely on data that just happened to be available or readily obtainable, but which yield measurements that others would not accept as meaningful in the context of the study.

Validity cannot be verified or tested directly, but will have to be based on the judgement of the evaluator. However, validity can often be improved by using several different measures for the same phenomenon, such as several different measures of income in the example above. In the context of an evaluation this requires that the evaluation instruments are checked as to whether they measure what they intend to measure. Firstly, that the evaluation questions correspond to the evaluation criterion being focused. Secondly, that questionable validity is compensated by using several information sources or methods. Thirdly, that the evaluation questions in combination are sufficient to provide a valid overall conclusion.

Information is *reliable* if the measurement procedure yields the same results if applied repeatedly; in other words, that the evaluation instrument with a certain degree of confidence can ensure that the measurement results correctly represent empirical relationships. For instance, an I.Q. test would be unreliable if it was administered to the same person twice and produced different scores.

Validity and Reliability: An Example
In studying women's influence in decision-making there are two choices: either to register the number of women attending the relevant meetings over the past year; or to observe what actually goes on at these meetings. The former approach would yield information with a high degree of reliability but low validity (no information of women's role in the meeting). The latter approach is less reliable but will observe directly and have greater validity. The best solution is to combine the two approaches, supplemented by interviewing a few participants (triangulation).

The evaluation team will have to consider the reliability of instruments used. Evaluation reports should always discuss reliability and validity of evaluation instruments and data and should indicate any weaknesses and the implication of these, as appropriate.

With limited time and resources, the evaluator needs to make use of whatever documentation is available, including international or national research and evaluation material as well as consultant reports from comparable projects. In many instances, the material may be of uncertain quality and difficult to check.

Time, cost, and data quality are interrelated, and the evaluator must consider the trade-offs between them. There is no automatic correlation between the cost of information and its value. The most expensive option is not automatically best. A thorough scrutiny of evaluation instruments usually helps to improve results and save money.

6.3 ENSURING INDEPENDENCE

Lack of credibility undermines the use of evaluation findings. Factors influencing credibility include the competence and credibility of the evaluator, mutual trust between the evaluator and those evaluated, consultation and involvement of stakeholders and methodological quality of evaluation. Hence the requirements that evaluations to the extent possible must be objective, impartial and independent.

Evaluations Must Be:

Objective and transparent – Assessments must be clearly separated from factual statements, and be based on reliable data or observations. Relevant parties in Denmark and the recipient country must be consulted in connection with the preparation and implementation of evaluations and the discussion of the draft report.

Impartial – Evaluations must give a balanced presentation of strengths and weaknesses. In the event that interested parties have different views these should be incorporated in the evaluation.

Independent – 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 are not eligible.

(Danida policy)

In order to achieve a high degree of objectivity, agencies have frequently relied on external, independent evaluation teams, and in some cases also established independent evaluation secretariats outside the donor agency. This type of objectivity has a cost.

The evaluation is more easily perceived as a control measure with a negative bias on shortcomings. Being external, it is often felt that the focus is not relevant and the results are not internalised in the organisation.

Danida's Evaluation Secretariat is internal to the organisation with a certain degree of autonomy and independence. Also, Danida's own staff is involved in evaluations. An acceptable degree of objectivity is ensured by the methodological standards, the competence of the evaluator, and the transparency of the evaluation process.

Objectivity implies that different angles are given due consideration to avoid biases. The objective approach may result in conclusions that are both positive and negative. It is essential that controversial conclusions are substantiated. Criticism should be expressed in constructive terms. Criticism that is insufficiently substantiated and expressed negatively may undermine the credibility of the report and reduce the chance that it will be used. Statements that are unduly negative or positive may also have as a consequence that they are lifted out of the context by the mass media and thus may distort the information flow to the public.

6.4 ENSURING AN INTERDISCIPLINARY PERSPECTIVE

Development interventions affect societies in various and complex ways. Any attempt to evaluate the effect of these interventions will require an interdisciplinary approach. Six factors are of particular importance to all aid interventions. These so-called sustainability factors are discussed briefly below.

Policy Support Measures
The recipient's commitment is one of the most commonly identified factors affecting success of aid 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 pre-dominantly 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 perspec-tive. Evaluations will frequently have to look specifically at environmental policy, incentives and regulatory measures, the interests of different stakeholders, and the effects of aid 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.

Aid 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 aid interventions, particularly their access to means of production and support services, as well as their rights and benefits.

Sustainability Factors:

1. *Policy support measures:* Policies, priorities, and specific commitments of the recipient supporting the chances of success.
2. *Choice of technology:* Choice and adaptation of technology appropriate to existing conditions.
3. *Environmental matters:* Exploitation, management and development of resources. Protection of the environment.
4. *Socio-cultural aspects:* Socio-cultural integration. Impact on various groups (gender, ethnic, religious, etc.).
5. *Institutional aspects:* Institutional capacity and distribution of responsibilities between existing bodies.
6. *Economic and financial aspects:* Economic viability and financial sustainability.

Institutional Aspects

The strength of institutions is the single most important factor in the success of aid interventions. Current trends to change the division of roles between public and private institutions 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.

6.5 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.

Organisations of professionals such as the American Evaluation Association or Canadian Evaluation Society have developed *ethical standards* for evaluators. Taking international development assistance as a special case, some donor agencies have also issued ethical guidelines. One such standard is quoted below (CIDA):

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 burden-some. 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 donor country's values affecting minorities and particular groups, such as women. 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 Secretariat as necessary.

Evaluation of Individuals
Performance evaluation is not normally a part of project 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. In general the evaluation team should consult Danida when there are questions about what should be reported, how and to whom.

Sharing the Results
Aid partners and executing agencies are generally eager to receive results of the evaluation. Technically, the reporting relationship is between the evaluator and the donor agency. The dilemma is to weigh the social and political concerns of the donor country's programme against the evaluator's commitments to participants. In general, it is appropriate to share briefings and unofficial summaries, but sharing the evaluation report itself is the donor's prerogative.

7.1 PREPARING THE REPORT

Reporting evaluation findings starts with Terms of Reference for the evaluation. 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 would therefore normally draft a table of contents at the earliest stage of the evaluation, based on Danida's requirements for reporting, the Terms of Reference and the introductory discussions with Danida. This will also help to distribute responsibilities between the members in the evaluation team.

Evaluation Report

The evaluation team has the final responsibility for the contents of the report. Any differences of views among the evaluation team, authorities in the partner country and Danida must be included in the report, either in the form of comments in the text or as a *special section.*

Evaluation reports must be brief and concise. The normal language is English and the presentation must be clear and adjusted to the target group.

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

If an evaluation is thought to be of interest to a broader public, a Danish language version will be produced (possibly abbreviated) and edited with a view to readers that are not professional specialists.

(Danida policy)

The draft table of contents will be a convenient means to organise information as the work proceeds. This includes information from initial desk studies, special studies, etc. Normally, the team would start preparing as much as possible of the draft report before departing for the field work. This will help to focus the field work to collect missing information, verify information, and draw and discuss conclusions and recommendations, as discussed in chapter 3.

Usually, the team will split after the field work. Team members would therefore normally be obliged to complete their inputs to the preliminary draft report while in the field. This will help the team leader presenting a draft summary to be discussed with country officials and the Embassy before leaving.

The preliminary draft of the report is reviewed by the Evaluation Secretariat and Danida's technical advisors and operational staff. They will assess the quality of the report and may submit comments regarding factual information, conclusions, and recommendations to be considered by the team before the final draft version is prepared. This is usually done in a meeting or workshop where the evaluation team is debriefed. Danida will then obtain comments on the final draft report from the authorities in the partner country and make 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) as well as the first draft of the follow-up memorandum of main recommendations. The structure and content of the report is discussed in Chapters 7.2 – 7.3. The final report should be in accordance with Danida's layout for evaluation reports. The layout is available from the Evaluation Secretariat.

7.2 THE STRUCTURE OF THE REPORT

Since the focus of evaluations may vary between project, sector, country evaluations, etc., there is no common format for reporting. However, there are different types of information and different principles of presenting information that should be adopted by all evaluation teams.

Two main key principles are that:

- The main report should be relatively short. It should contain the aggregated information and highlights. All detailed information should be referred to annexes.
- Information should be organised in a simple format which is easily accessible to the reader. This implies that discussion of the evaluation criteria will be a focus in the report.

The evaluation report will consist of three levels of information. Firstly, the executive summary to be written as a selfcontained paper providing the bare essentials for decision-makers regarding the background, major conclusions, recommendations and lessons learned.

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

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

The third level in the report should contain annexes. These should 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 anne

The Main Report Should Include the Following Information:

- *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 aid interventions are implemented, in terms of geography, policy, sectors and institutions, as appropriate*
- *Description of the planning, and implementation of aid activities. This includes the presentation of objectives, target groups, components, financing, management, etc.*
- *The main conclusions based on overall assessments, including efficiency, effectiveness, impact, relevance, and sustainability*
- *Administrative issues regarding financing, management and follow-up of aid activities by Danida and partner authorities, and possible need for further studies*
- *Lessons learned*

7.3 EDITING THE REPORT

The evaluator is expected to maintain a professional standard in reporting. The 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 should meet the needs of all groups.

The language, use of terminology, illustrations, etc., are the main aids for efficient communication. The report should be written keeping target groups with the least advanced readers in mind.

Below are some general advice to make reports easier to read:

Be Brief

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

Structure

- The structure of the report should be simple with titles and subtitles precisely identifying the topics discussed.

- A “horizontal” structure should be used with not more than two levels of headings.
- The text should be broken down in relatively small thematic or sequential parts, with simple and clear subtitles.

Editing

- The first sentence of paragraphs should be used to make the main point, and the remainder to supplement, substantiate and discuss the main point.
- Illustrations, graphs, tables, etc. should be used extensively to illustrate and elaborate on points that should be emphasised.
- The details of tables and diagrams should not be repeated in written form. Instead, the reader’s attention should be focused on the important points that can be derived, and related to the problem under discussion.
- Jargon and difficult words should be avoided. As far as possible, the report should use words that the readers will understand. Essential technical terms that may be new to the reader should be defined in the text, and in a glossary at the end.
- Long and complicated sentences should be broken down in several short sentences. It may be useful to give lists of items or conditions in a step-by-step form.
- Active sentences should be preferred. The more complicated passive form should be used essentially to focus attention on a specific topic.
- Ideas should not be crammed together in complex sentences with commas and brackets. Separate main ideas should be presented in separate sentences.
- The meaning of abbreviations and colloquial words should be explained. Abbreviations should be used as little as possible.
- Simple link words should be used to split sentences and indicate the direction in which the argument is moving. Link words should be simple, such as: also; even so; on the other hand; in the same way. Long words like: moreover; nevertheless; notwithstanding, should be avoided.

Presenting Numbers

- Only data tables or diagrams should contain detailed numbers. The written text should highlight only the most important numbers and say what they mean. Percentages should in most cases be rounded up 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 must therefore have 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 should be used to make the report lively and interesting. These are often more effective than written text for communicating straightforward messages.

- Use space around the text. Ease of reading and understanding is more important than reducing the volume of writing.

8.1 THE 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 a general awareness, interest and support for development assistance. The evaluation report provides a status of strength and weaknesses of aid 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 aid 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 Evaluation Secretariat undertakes to control the implementation of the follow-up activities at regular intervals.

The Evaluation Secretariat contributes actively to the dissemination of Danida's own as well as other aid organisations' evaluation experience via workshops and seminars for staff in co-operation with the Ministry's education section. Further, the Evaluation Secretariat assists Danida's Training Centre 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 Secretariat will contribute to the incorporation of evaluation experience in policies, strategies and guidelines, etc.

(Danida policy)

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 Secretariat 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 operating department 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 Secretariat 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 Secretariat 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 aid 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 Secretariat, seminars and workshops, and guidelines issued by the Evaluation Secretariat. Brief summaries of evaluation reports are submitted to the database of the OECD/DAC Working Party on Aid Evaluation, and all Danida’s evaluations are published and made public in the form of printed reports, summaries, and electronically in the Ministry’s web-site.

8.2 DISSEMINATION OF INFORMATION

Whether evaluation reports should be made public has been a matter of debate. Evaluation activities were never meant to be a means of strengthening public support 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 at present 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 information in mass media has been more thorough with the availability of evaluation reports.

Apart from distributing the evaluation report itself, the most common ways to disseminate evaluation information is through the evaluation summaries, annual reports, bibliographies, thematic reports, seminars, press releases, and public debate. The recent initiative to make evaluation reports available electronically via the Internet opens an entirely new type of dissemination channel. The effects remain to be studied.

Whichever channel is preferred, the best way to ensure dissemination of lessons learned and knowledge gained in evaluations is probably to improve both the content of reports and the presentation of material. The main purpose of evaluations will still be to improve the quality of international development assistance. The extent to which they will also improve transparency and public insight will be a most welcome, but secondary benefit.

ANNEX II: DANIDA'S EVALUATION POLICY

1. OBJECTIVES

According to DAC's definition, evaluation is an assessment, as systematic and objective as possible, of on-going or completed aid activities, their design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, developmental efficiency, effectiveness, impact and sustainability.

Danida's evaluations have two objectives:

Accumulation of Experience (lesson learning)

- The evaluations shall contribute to the improvement of aid by the collation, analysis and dissemination of experience from current and completed aid activities. Evaluations shall seek the causes and explanations as to why activities succeed or fail to succeed and produce information that helps make future activities more relevant and effective. The target group is Danida's management and staff, interested parties in partner countries and the resource base of Danish and foreign aid professionals.

Documentation

- Further, evaluations shall provide political decision makers and the general public with professional documentation as to the use and results of aid resources and also contribute to a better understanding of development aid, its potential and limitations as an instrument for economic and social change.

Both objectives must be satisfied in Danida's evaluation activities, though individual evaluations shall not necessarily fulfil both objectives to the same degree.

2. THE TASKS OF THE EVALUATION SECRETARIAT

The tasks of the Evaluation Secretariat are:

- plan, formulate and carry out evaluations of Danida-financed aid activities;
- disseminate and contribute to the dissemination of evaluation results and experience by other departments to Danida staff, authorities in partner countries, other interested bodies and to the Danish public;
- contribute to the development of strategy and policy in Danida on the basis of own and other donors' experience;
- build and maintain an electronic database of evaluations;
- develop and improve methods of evaluation and evaluation guidelines;
- contribute to the development of evaluation capacity in Danida Programme countries;
- participate in international co-operation on evaluation.

3. PLANNING OF EVALUATION

Danida's Evaluation Secretariat is responsible for presenting multi-annual plans which, over a number of years and taking their point of departure in the overall objective of aid, shall cover:

- ongoing or completed projects and programmes;
- Danish aid to a particular country, one or more sectors in a partner country or a sector in a number of partner countries;
- aid instruments (project and programme aid, aid via NGOs, mixed credits, balance of payments support, emergency aid, research support, fellowship programmes), and
- multilateral aid, primarily in co-operation with other donors.
- The evaluation programme shall contain a brief rationale for each evaluation: the primary objective (documentation or accumulation of experience), the target group and which features in particular the evaluation shall comprise.

Documentation evaluations are planned in order to achieve a suitable coverage of geographical areas, large and small partner countries and sectors over time.

Evaluations to accumulate experience are planned on the basis of the strategic and operational needs of the aid administration. They can cover the above subjects but will, in addition, often incorporate multidisciplinary subjects (i.e. strategies, capacity building, planning and implementation of projects and programmes, consideration of the environment, gender aspects) or elements of these.

The individual evaluations should normally consider the following issues:

Relevance – Is the activity relevant to Danish aid policy, goals and strategies: 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 partner country? Is the activity relevant to the development issues it aims to solve?

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

Sustainability – What are the long term effects of the activity? Can the activity be continued and can it provide acceptable returns, when Danish aid is terminated? Is local ownership established?

Efficiency – Are the investment and recurrent costs justified? Could the same results have been achieved with fewer resources?

Process – Is aid organised in the most appropriate way possible? Are the relevant parties and authorities in the partner country involved in preparation and implementation?

Experience – Can specific or general experience be gained that is relevant for the preparation and implementation of aid?

4. IMPLEMENTATION OF EVALUATION

Evaluations must be:

Objective and transparent – Assessments must be clearly separated from factual statements; assessments must be based on reliable data or observations. Relevant parties in Denmark and the partner country must be consulted in connection with the preparation and implementation of evaluations, drafting of Terms of Reference (ToR) and discussion of the draft report.

Impartial – Evaluations must give a balanced presentation of strengths and weaknesses. In the event that interested parties have different views these should be incorporated in the evaluation.

Independent – 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 are not eligible.

Preparation - The management and staff of Danida as well as the relevant representatives of partners (the country, NGO or multilateral organisation) should be involved in the preparation of evaluations by participating, to the greatest possible degree, in the formulation of TOR with a view to focusing the evaluation on relevant subjects. Before an evaluation is set in motion, the Evaluation Secretariat prepares, in consultation with relevant interested parties, a description of objectives that clearly describes:

- the main objective of the evaluation: accumulation of experience or documentation;
- what the evaluation must cover: the entire or part of a project, a programme, a sector, an instrument of aid or a particular issue;

- which subjects or features the evaluation shall specifically cover;
- data basis for the evaluation, primary and/or secondary data, interviews, field work etc., and
- time schedule

On this basis, the actual Terms of Reference are prepared which forms the basis for the selection of the evaluation team. The terms of reference must be presented as a number of specific questions.

Choice of Evaluation Team

The criteria for the selection of the evaluation team are professional competence, experience in relation to the task and impartiality. The team must constitute a representation of the relevant professional fields. Professional expertise from the partner country shall as far as possible be represented on the team.

Implementation

Evaluations are typically carried out in two phases. In the first phase, the evaluation team drafts, on the basis of studies of documentation and interviews in Denmark and, where necessary, in the partner country, a preliminary report that puts forward hypotheses, provisional conclusions and identifies areas where further documentation must be gathered. Based on this, the evaluation team and the Evaluation Secretariat elaborate a detailed plan for the future work in the second phase that may cover further documentation studies, interviews, primary data collection, field observations etc.

In connection with accumulation of experience evaluations, Danida management and staff plus the relevant representatives of partners must be involved in the implementation of evaluations to the greatest possible degree to:

- participate in the analysis. On the basis of documentation studies and the information gathered, the evaluation team presents the problems and/or the hypotheses put forward with a view to discuss and analyse the material produced together with relevant staff, and
- participate in the formulation of lessons learned, conclusions and recommendations.
  This will often take place in a number of workshops or seminars in the course of the evaluation.

In evaluations with a documentation objective, the participation of staff will typically be less intensive: procuring documentation, participation in interviews, debriefings and discussion of draft reports.

5. REPORTING

The evaluation team has the final responsibility for the contents of the report. Any differences of views among the evaluation team, authorities in the partner country and Danida must be included in the report, either in the form of comments in the text or as a special section.

Evaluation reports must be brief and concise. The normal language is English and the presentation must be clear and adjusted to the target group.

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

If an evaluation is thought to be of interest to a broader public, a Danish language version will be produced (possibly abbreviated) and edited with a view to readers that are not professional specialists.

6. FOLLOW UP

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 Evaluation Secretariat undertakes to control the implementation of the follow up activities at regular intervals.

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

On the basis of own and others’ evaluation experiences, "best practices" will be compiled and formulated.

Further, the Evaluation Secretariat will contribute to the incorporation of evaluation experience in policies, strategies and guidelines etc.

7. PUBLICATION

All Danida evaluations are published in the form of printed reports, summaries and electronically, in the Foreign Ministry's web-side.

27. August 1997

COLOPHON

Production: Evaluation Secretariat, Ministry of Foreign Affairs

Cover lay-out: Designgrafik A/S

Print: Ministry of Foreign Affairs, Copying Center
2. edition, revised

ISBN: 87-7964-087-7

This publication can be obtained free of charge by contacting:

Bech Distribution
Hedelykken 2-4
2640 Hedehusene
Denmark
Ph: + 45 46 55 01 55
Fax: + 45 46 55 01 99
e-mail: bdm@bech-distribution.dk

This report can also be downloaded from the Ministry of Foreign Affairs' homepage on the internet address:

http://www.evaluation.dk/

The cover is printed on Finncard, a fully coated folding boxboard. Finncard is manufactured without the use of chlorine and marked with the Nordic Swan, licence-no. 404 039