Sida Studies in Evaluation 99/2

Managing and Conducting Evaluations

Design study for a Sida evaluation manual

Lennart Peck
Stefan Engström

Department for Evaluation
and Internal Audit

Managing and Conducting Evaluations

Design study for a Sida evaluation manual

Lennart Peck
Stefan Engström

Sida Studies in Evaluation 99/2

Department for Evaluation and Internal Audit

Sida Studies in Evaluation is a series concerned with conceptual and methodological issues in the evaluation of development cooperation. It is published by Sida's Department for Evaluation and Internal Audit. Reports may be ordered from:

Biståndsforum, Sida
S-105 25 Stockholm
Phone: (+46) 8 698 57 22
Fax: (+46) 8 698 56 38
E-mail: info@sida.se, Homepage: http://www.sida.se

Authors: Lennart Peck, Stefan Engström.

Sida Studies in Evaluation 99/2
Commissioned by Sida, Department for Evaluation and Internal Audit.

Registration No.: UTV-1998-
Date of Final Report: May 1999
Printed in Stockholm, Sweden 1999
ISSN 1402—215X

SWEDISH INTERNATIONAL DEVELOPMENT COOPERATION AGENCY
Address: S-105 25 Stockholm, Sweden. Office: Sveavägen 20, Stockholm
Telephone: +46 (0)8-698 50 00. Telefax: +46 (8)-20 88 64
Telegram: sida stockholm. Postgiro: 1 56 34–90
E-mail: info@sida.se. Homepage: http://www.sida.se

EXECUTIVE SUMMARY

This *"Design Study for a Sida Evaluation Manual"* has been written to give Sida a basis for discussion and decision-making regarding the production of a new evaluation manual. According to the terms of reference the study should provide "an empirically grounded analysis of the needs and demand for a new evaluation manual among Sida staff and other stakeholders in development co-operation activities supported by Sida". It should also propose "an outline of a manual that would satisfy these needs and this demand."

Chapter 2 initially tries to broaden perspectives by discussing what an evaluation manual *may* look like, in terms of function, content and format. Examples are drawn from evaluation manuals of other organisations. Possible *functions* of a manual include operationalisation of evaluation policy, enhancement of the competence of the participants in evaluations and establishment of procedures and routines. In relation to *content* the report discusses topics, scope, delimitations and structures of the manuals studied. The review of different *formats* considers, among other things, the length of the documents, layout and paper vs. computer.

Chapter 3 deals with *needs and demand* for a new evaluation manual. It starts by reviewing Sida's programme documents which, indirectly, set requirements for evaluation activities and therefore also for the evaluation manual. The statements on "partnership" and different development goals are examples of such requirements. In addition, the document *Sida at work* sets an overall framework for evaluations, as do international policies to which Sida is committed. However, these documents give limited practical guidance. The report then compares the (Draft) evaluation policy of Sida with findings in relation to Sida's evaluation activities to determine needs to change and improve. Such needs are noted in many areas, particularly in relation to the use of evaluations, participation and the treatment of overall development goals, such as poverty reduction and gender equality.

The report also looks at competence requirements. These are defined in terms of tasks faced by evaluators, evaluation managers and other participants in the evaluation process. Even though evaluations are different, the basic tasks involved are largely the same. However, specific evaluation problems related e.g. to the sector, purpose and context of the evaluation may require specific competence. Evaluations are often managed by persons handling other tasks as well, and who are not evaluation "specialists". Some evaluators are primarily hired for their sector competence and have limited experience of evaluations, while others have limited experience of working for Sida or with the many issues to be included in a Sida evaluation. The manual therefore needs to be kept to a relatively basic level covering a broad range of issues. Although participants in an evaluation will have different roles and functions, the knowledge that they need about evaluation work is largely the same.

Procedures and routines are needed for evaluation work. However, as no two evaluations are quite the same, there can be no standard routines.

Chapter 4 sketches a proposed design for a new Sida manual. According to this proposal - which should be seen as a basis for further discussions – the overall function of the manual should be to improve evaluations and their use, bringing them in line with Sida's evaluation policy. More specifically, the manual should:

- enhance the competence of persons dealing with evaluations by providing relevant knowledge;
- provide practical instruments and tools for evaluation work;
- establish routines, procedures and norms;
- promote reflection and innovation.

In practical terms, the document should be possible to use as:

- a "beginner's manual", presenting an overview over relevant aspects of evaluation work;
- a "reference book", for looking up information on particular issues;
- an "idea book", for new impulses and reflections;
- an "information booklet" for anyone with an interest in Sida evaluations.

A key issue for Sida to take a stand on concerns the intended readership of the manual. Sida has expressed a demand for a manual directed to Sida staff and managers. However, other participants in evaluations could also need a manual.

Another question for Sida to take a stand on concerns the manual's official status.

As the manual should be capable of serving both as an introduction to evaluation work and as a reference book for non-specialists dealing with different types of evaluations, a broad scope is recommended. The manual should address the major tasks and related topics in the evaluation management process, the process of conducting an evaluation and the use of evaluations. What exactly is to be said on the different topics requires further discussion and consultation, particularly topics with a strong policy dimension, such as roles and responsibilities and the analysis of crosscutting issues.

A broad scope means that all topics cannot be dealt with in depth. Considering the demand expressed in policies and the needs for improvement noted in this study, it is suggested that attention be made to focus on:

- The use of evaluations, such as learning and control, different users etc.
- Roles and responsibilities in different evaluation contexts;
- Sida's development objectives, including the various crosscutting issues.

It is proposed that only topics likely to be relevant to any evaluation, i.e. not special evaluation problems related to sector, types of co-operation etc., should be included. It is also proposed that the manual be made to focus on evaluation and evaluation use only, in other words that it should not go into detail on the other phases of the project cycle, notwithstanding their relevance to evaluation work.

A tentative outline proposed for the new evaluation manual begins with "Main guiding principles", such as the definition of evaluation, Sida's evaluation policy and quality criteria. The reason for starting with main guiding principles is that these should permeate all evaluation work and the continuing discussion in the manual.

This is followed by a chapter on the "Use of evaluations", including learning and control, which people evaluations are intended for, management response etc. The reason for bringing up the issue of evaluation use early in the manual is that this use should guide the evaluation and evaluation management process.

The third chapter, "Main components of the evaluation", covers different evaluation objects, LFA and evaluation criteria. This should give the reader the necessary understanding of the main elements of an evaluation and the assessments to be made, as well as indicating differences between different types of evaluations.

The proposed Chapter 4, "Conducting an evaluation", gives the reader a basic introduction to the main tasks, including the choice of different approaches, data collection and analysis, formulation of recommendations and report writing.

Next comes a chapter on "Managing evaluations". This covers the different steps in the evaluation management process: the decision to evaluate, preparation of the terms of reference etc. It gives the reader practical guidance, making reference to what has been stated earlier in the manual, e.g. on guiding principles and different types of evaluations.

Having identified the tasks to be performed, the manual turns to "Roles and responsibilities in evaluations". Different forms of participation are discussed with reference to central principles, such as objectivity and partnership, the various objectives of evaluations, e.g. learning and control, and different evaluation contexts.

It is proposed that practical instructions for how to address poverty reduction, gender equality, environmental impact, democracy and human rights in evaluations be attached as appendixes. A similar appendix on the assessment of institution building is also proposed.

In terms of format, it is proposed that the manual should be produced as a soft-cover book, similar to the existing manual.

This manual can be supplemented by "Issue Papers" on topics relevant to special target groups and reflecting the on-going discussion and development work in the field of evaluation. The printed manual can also be supplemented, though not substituted, by a digital version. Making the manual and issue papers available on the Internet will require little effort. Computer-based training programmes, reference books, linkages to word documents etc. can be prepared in the mid-term perspective with assistance from outside expertise.

A layout should be chosen which facilitates reading and the search for information. As the manual will be read by people with different backgrounds and different ways of absorbing information, it should use both text and illustrations. The need to give the reader both a deeper understanding of evaluation work and practical guidelines calls for a mix of explaining texts and checklists. Standard formats of Sida publications should be taken into consideration.

Table of Contents

1 Introduction

Sida's organisation for evaluation activities is decentralised, and involves a large number of people, both internally and externally. Many of them are engaged in evaluation work only occasionally. This calls for clear guidelines on how evaluations[1] should be managed and carried out.

The latest edition of Sida's evaluation manual dates from 1994[2]. Since then, major organisational changes have occurred within Swedish development co-operation and a number of new programmes and policies have been adopted, including Sida's evaluation policy.[3] New issues have been emphasised in the discussion on evaluation and new lessons have been learned concerning Sida's own evaluation activities. The Department for Evaluation and Internal Audit (UTV) has therefore decided to produce an updated manual, and UTV has commissioned this "Design Study for a Sida Evaluation Manual" so as to ensure that it matches both needs and demand

1.1 Purpose of the assignment

The study has been written to provide Sida with a basis for discussion and decision-making regarding the new evaluation manual. The terms of reference mention the following two main purposes:

> *"to produce an empirically founded analysis of the needs and demand for a new evaluation manual among Sida staff and other stakeholders in development co-operation activities supported by Sida;*
>
> *to propose an outline of a manual that would satisfy these needs and this demand."*

The outline should cover the format of the manual as well as its contents, and be clearly and explicitly motivated in terms of the prior analysis of needs and demand.

The terms of reference (Annex 1) also include a number of issues to be given special consideration. These can be summarised as follows:

Factors determining the content of the manual:

- Sida policies, evaluation policy included;
- Common international understandings of concepts and principles of evaluations, including those of DAC;
- The level of knowledge possessed by the intended reader;

Important issues and perspectives to be dealt with in the manual:

- The project and programme cycle management perspective;
- Practical evaluation management in the different steps of an evaluation;
- The roles of different stakeholders in evaluations;
- Joint evaluations and participatory evaluation techniques;
- Different types of evaluations;
- Evaluation problems specific to one or several sectors of development co-operation;
- Crosscutting issues, such as poverty and environment;
- Reference to other sources of information;

[1] The word "evaluation" will here be used to describe the entire evaluation exercise, not only the evaluation report.
[2] Lewin, E., Evaluation manual for Sida, 1994.
[3] Department for Evaluation and Internal Audit, Sida's Evaluation Policy, 1995. This document is currently under revision.

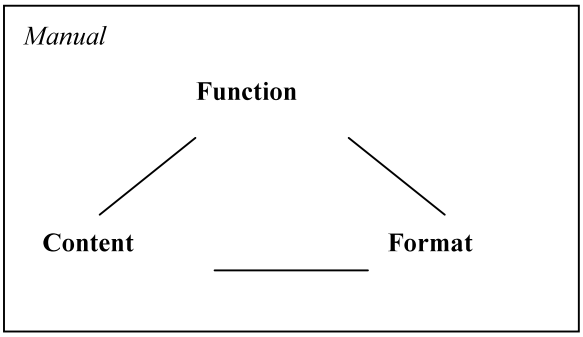
Properties of the manual as such:

- The possible usefulness of the manual to stakeholders outside Sida;
- Presentation in a workable format;
- The balance between coverage and detail, and the manual being "digestible";
- The life-span of the manual.

1.2 Model for analysis

A manual can be defined in terms of *function*, *content* and *format*. It should be motivated in terms of *needs and demand,* as illustrated here:

NEEDS AND DEMAND
- *policy requirements*
- *required improvement of evaluation activities*
- *necessary competence and procedures*
- *practical requirements*



By the *function* of a manual we mean how it can be used and what benefits that it can lead to. At an overall level, it will be assumed that the function of a manual is to *improve evaluations* and hence development co-operation as a whole. At the next level, it is assumed that the manual can improve evaluations primarily in two ways: by enhancing the *competence and capacity* of persons managing and conducting evaluations, and by *establishing systems and routines* for evaluation work. Finally, there is a third level related to the *practical* use of the manual, e.g. should it serve as a reference book or as teaching material? Where and when? Another important aspect of function concerns which people the manual is directed at, i.e. its *intended users*.

Content refers to particular topics, what is actually *said* about them and how they are structured. This is also a question of focus, scope and delimitations.

Format refers to how the content is presented, including the volume of the document/s, shape, layout and media (printed or computer).

Both content and format must be determined according to how the manual is meant to be used, i.e. its function. Furthermore, the content has implications for format and vice versa. The three must therefore be seen in conjunction.

Turning to the question of needs and demand, a short statement of the distinction made between the two will perhaps not be out of place. *Demand* refers to a person's interest or desire for something (e.g. what an evaluation manual should contain). To ensure that this interest is genuine we will, just as when using the term *demand* in market contexts, assume that the person is also is prepared to pay the price involved. For example, is Sida really willing to pay the price of the type of evaluations described in the policy, or will the person requesting an evaluation manual actually take the time to read it? By *needs,* we will mean what is considered *necessary,* given a certain context and certain objectives, regardless of the existence of otherwise of an explicit demand. For instance, if poverty should be assessed in evaluations and evaluators are incapable of doing so, we would say that there is a *need* to enhance the competence in poverty assessment – a need which must be reflected in the design of the manual.

Needs and demand can be analysed at different levels, and they vary from one group of people to another. As the manual should serve to improve evaluations, we begin by considering what evaluation activities and reports *ought to* be like, so that this can be reflected in the manual. To this end we make a review of central *policy documents and guidelines* and discuss some of their implications. Then, if there is a gap between *policy and practice* in some area, it may be interpreted as a *need for improvement* or *change.* Special attention should be paid to this in the manual. Assuming that the main benefits of a manual will be the enhancement of the *competence* of those dealing with evaluations, and the provision of suitable *procedures and routines,* we then turn to consider needs and demand in these two respects. Finally, we discuss some *practical needs and demand* in relation to a manual, i.e. how it is supposed to be used in daily work.

1.3 Report structure

Before discussing the design of a Sida manual, it may be useful to broaden perspectives and consider what an evaluation manual *may* look like. What are the possible functions, content and formats of a manual? In Chapter 2, we have looked at a number of evaluation manuals and gathered ideas from persons within and outside Sida, for the purpose of showing some of the many possibilities. This should make it easier at a later stage to define what the Sida manual should look like.

Chapter 3 deals with the needs and demand for a new manual as outlined above. It presents central policy documents and compares policy with practice by making reference to different studies made on Sida's evaluation activities. It also discusses the competence needs of different groups and practical requirements.

Having first considered what a manual *may* look like and then the *needs and demand,* it is possible to start sketching the *new Sida manual.* This is done in Chapter 4. Inevitably, there are a number of, sometimes contradictory, requirements and the discussion therefore partly deals with how a balance can be struck between the different needs and requirements. It should be recalled that the purpose is to provide Sida with a basis for discussion and decision making and not to come up with a final solution. To this end, a first rough "sketch" of an evaluation manual is provided in Annex 7.

1.4 Sources

Printed sources used for this study can broadly be grouped into:

- Sida policies and guidelines, including Sida's different action programmes, *Sida at Work* and the Sida's evaluation policy;
- Sida studies in evaluations, in particular *Using the Evaluation Tool, Are Evaluations Useful?* and the different studies on cross-cutting issues in evaluations;
- Selected evaluation manuals of different development co-operation agencies (Denmark EU, Finland, IDB, Norway, UNDP, USA as well as Sweden);
- General literature on evaluation methodology and management;

For a compete list of printed sources, see Annex 2. Please note that no assessment has been made of the quality or accuracy of the documents referred to.

Interviews were made with a number of Sida officers who have recently been in charge of evaluations. The interviewees were selected so as to achieve a reasonable spread of different Sida departments and types of evaluation. The interviews were semi-structured, following a simple interview guide, in order to assure that the most important issues were covered, but without preventing the interviewees from raising other issues that we may have overlooked. Interviews were also conducted with Sida managers representing different departments and sectors, as well as UTV staff members. A few consultants and Swedish NGOs were interviewed in order to get an outside perspective. For a list of persons interviewed, see Annex 3.

No statistical analysis has been attempted of the answers received. For the purposes of this study, quantification has been judged less important than arriving at the broadest possible picture of needs, problems and potential solutions.

1.5 Implementation

The study has been carried out by Lennart Peck, Boman & Peck Konsult AB, assisted by Stefan Engström, Stockholm University.

During the assignment, the consultants have had a continuous dialogue with UTV. A working draft was presented to UTV 22 March and discussed with one of its staff members, and UTV participated in the selection of literature, as well as suggesting persons to interview.

A draft report was submitted during April and discussed at two meetings, one with Sida's evaluation co-ordinators and one with UTV and invited ersons from other departments. The comments made at these meetings have been fed into the preparation of this final report.

We take this opportunity of expressing our gratitude for all the useful information and ideas received in the course of our work.

2 What can a manual look like?

As a point of departure for the discussion on what a new Sida evaluation manual should look like, this chapter takes a broader view on what such manuals *may* look like. For this purpose we have studied the manuals of eight different organisations, selected to be different enough to give new impulses, and still similar enough to permit some form of comparison. These are:

- Denmark, Danida, Evaluation Guidelines, 1999;
- Finland, Ministry of Foreign Affairs: Guidelines for Programme Design, Monitoring and Evaluation, 1997[4];
- EU, Evaluating EU Expenditure Programmes: A Guide, 1997;
- Inter-American Development Bank (IDB): Evaluation: A Management Tool for Improving Project Performance, 1997;
- Norway, Ministry of Foreign Affairs: Evaluation of Development Assistance, Handbook for Evaluators and Managers, 1993;
- Sweden, SIDA, Evaluation Manual for Sida, 1994;
- UNDP, Handbook for Programme Managers called Resluts-Oriented Monitoring and Evaluation, 1997;
- USA, USAID, Office of Evaluation Procedures Guidebook for Conduction CDIE (Center for Development Information and Evaluation) Evaluations, 1992.

The discussion that follows is also based on interviews and the ideas provided to us in connection with them.

2.1 Functions

What are the possible uses of an evaluation manual, and what benefits can they be expected to lead to?

2.1.1 Operationalisation of policy

The overall function of an evaluation manual can be assumed to be that of improving evaluation activities and reports. What these *ought to* look like is usually described in policy documents, and several of the manuals reviewed for this study begin by making reference to overall policies and objectives. The Danish manual, for example, states that "the policy is the basis of the present guidelines"(p. 6.) and the policy in itself is attached as an annex. In this perspective the manual can be seen as an instrument for operationalisation of the policy, i.e. from going from policy to practice. It interprets the practical implications of the policy, stating not only what should be done but also *how to* do it.

When introducing *new* policies, a manual can thus serve as an instrument of change and organisational development. The UNDP manual, for example, seems to have this aim, referring as it does to "reinforcing the linkages between monitoring and evaluation" and "featur[ing] new dimensions of (...) monitoring and evaluation within the context of participatory development, the country co-operation framework etc."

---

[4] This publication is not an evaluation manual only but provides guidelines for the whole project cycle. It is also more of a policy document than a practical handbook. In this sense it resembles *Sida at Work*.

***Implications for manual design:** The new evaluation manual should reflect Sida's evaluation policy. Hence, it is important that there is such a policy. The manual can make reference to this policy or even include it as an annex. The evaluation manual may supplement the policy by stating not only what should be done, but also* how *to do it.*

2.1.2 Enhancing competence and capacity

In what ways, then, can a manual improve evaluations?

A "manual" is usually understood to be a publication for enhancing a person's capacity to perform a task. "Capacity to perform a certain task" is also a common definition of *competence.* Hence, the function of a manual can be said to be to be that of *enhancing someone's competence* in a particular area, e.g. aid evaluation.

Looking closer at the notion of competence, it is often broken down into *knowledge, skills, experience and attitudes.* What a manual can provide is primarily *knowledge,* e.g. *what* to do, as well as *why, when* and *how* to do it. It can of course give practical instructions and include exercises, but these are no substitute for actual skills. It may present general experiences gained by others and state what attitudes and values will be expected by the organisation (e.g. on issues such as partnership and integrity), but this is no substitute for the reader's own experiences and attitudes. Hence, an evaluation manual may contribute to competence but is not sufficient to ensure it.

A manual may also enhance the capacity of the reader by providing instruments for practical use. For example, the second part of the UNDP-manual, entitled *Developing Selected Instruments for Monitoring and Evaluation,* contains a number of forms and checklists on how to select indicators, how to make an evaluation plan, what a report outline should look like etc. The IDB manual has a special annex which "provides a set of tools to help project teams plan from monitoring and evaluation (M&E) throughout the project cycle, and to incorporate this plan into project design. The "Logical Framework Approach" (LFA), presented in some of the manuals, can also be seen as such a practical instrument.

***Implications for manual design:** A Sida manual can be used to enhance the capacity of the reader, primarily by providing knowledge and practical instruments. However, for practical skills, personal experiences and attitudes, other measures are required.*

2.1.3 Provide systems and routines

Manuals may provide instructions for the different tasks and processes in an evaluation, e.g. how to prepare the terms of reference and how to process an evaluation report. There are two sides to this. First, the existence of systems and routines facilitates the work of the persons actually engaged in the evaluation. Second, it can be a way of establishing norms[5] for the way in which evaluations *ought to be* conducted and managed. The existence of such norms is important for quality control and management.

The extent of a manual's normative function partly depends on its official status. In the preface to the Finnish document it is stated that it should "be applied in all future Finnish development interventions". Swedish evaluation manual states that it "has no 'official' status in the sense of

[5] A *norm* can be defined as a standard of proper behaviour, or principle of right and wrong. The fundamental principles to guide evaluations are generally established in the policy. We refer here to behaviour e.g. how certain tasks *should be* carried out and how vocabulary *should be* used. Naturally, this should reflect what is stated in the policy.

laying down compulsory instructions for evaluation activities". It has most probably had an important normative function all the same.

The extent to which a manual is perceived as normative also depends on how things are explained and presented in it. Some manuals tend more to present *the* way in which evaluations should be carried out, while others open up for more alternatives and give suggestions. Including "good" or "best" practices is something of a middle way, as it sets norms for what is suitable but also indicates that there are many ways of doing things.

If checklists, standard formats etc., may facilitate work, they may also be an effective way of making people stop and think for themselves. Hence, there must be a balance between, on the one hand, standard procedures and instruments and, on the other, knowledge and understanding.[6] A fundamental question appears to be to what extent a manual should define fixed standards or to what extent it should challenge the reader to reflect and encourage innovation.

Implications for manual design: *A Sida evaluation manual may have the function of providing different procedures and routines. A choice can be made between a manual which emphasises certain standard routines, facilitating quality control, or leaves room for more flexibility, which may promote reflection and innovation. The manual must be clear about what are to be taken as mandatory instructions and what are merely wise suggestions. A decision must also be made regarding the official status of the document.*

2.1.4 Intended readers of the manual

There are many people involved in an evaluation: those managing the evaluation process, the evaluators, those who are being evaluated, etc. Roles and responsibilities differ; this will not be entered into here. Suffice it to note that a manual has different potential users with different perspectives and different requirements.

Who, then, have the manuals reviewed actually been written for? The IDB manual is directed to "Headquarters and Field Office Bank staff". The EU document is "aimed at programme managers within the commission services, as well as others who require a general introduction to the theory and practice of evaluation." The Norwegian manual states in its title that it is a "Handbook for Evaluators and Managers". The Swedish manual also opens up to a wider audience by stating that the "manual's primary aim is (...) SIDA management and personnel (...). A secondary aim is to explain SIDA's evaluation policy and quality requirements to consultants. Finally, it is also hoped that NGOs involved in development assistance and students interested in evaluation will be able to find information of value in this book". UNDP has the broadest targeting, stating that the document has been written for, "*Internally,* various levels of management within UNDP (...) and *externally,* other UN agencies and development partners, Governments and NGOs of programme countries, independent evaluators engaged by UNDP and members of the academic community".

The picture that emerges is that the manuals, first and foremost, have been written for the respective organisation's own staff, secondarily for consultants/evaluators and finally for everybody else who may have an interest in the subject.

There is nothing strange about an organisation preparing manuals and guidelines for its own staff. Nonetheless, the contrast between, on the one hand, manuals' talk of "participation", "stakeholder involvement" and "user orientation", and on the other hand, the absence of guidance or manuals for large groups of users and stakeholders, is somewhat remarkable.

[6] To quote an example from Sida: Interviews reveal that Sida's standard outline for writing terms of reference is highly appreciated, but voices are also raised that these are making work too mechanical.

Implications for manual design: *Sida must decide what groups the manual should be directed to, e.g. i) managers of evaluations, ii) evaluators iii) project owners iv) NGOs and other potential readers. Hence, Sida must also consider what the specific needs and demand of these groups are.*

2.2 Content

What may be the content of an evaluation manual? For the reader's own comparison, Annex 4 contains the tables of content of the evaluation manuals reviewed.[7] Below are some observations.

2.2.1 Topics

There are a number of topics that are virtually standard:

- What is an "evaluation"?
- Different types of evaluations;
- Reasons for undertaking evaluations and evaluation use;
- The evaluation in the project cycle;
- Goal hierarchies/LFA;
- Evaluation criteria and different bases for assessment;
- The different steps of an evaluation (the "evaluation cycle);
- Stakeholders, roles and responsibilities, and "participation";
- Issues (policy, management etc.) related to the specific organisation.

It is also quite common that the evaluation manuals reviewed deal with:

- Evaluation ethics;
- Development theory and policy;
- Development factors, including crosscutting issues;
- Quality standards;
- Evaluation methods;
- Evaluability.

Even though the headings largely are identical from one manual to another, what is actually *stated* under each of these headings may vary. For example, discussions of "different types of evaluations" may deal with the various objects of the evaluation (e.g. project, thematic, organisational evaluations), the timing of evaluations (mid-term, ex post evaluations etc.), roles and responsibilities (internal, external, collaborative, participatory evaluations etc.), what is measured (process and impact evaluations), quantitative and qualitative evaluations, formal and informal, summative and formative etc. The descriptions of evaluation use frequently evolve around words such as "decision-making", "learning", "management", "capacity building", "documentation", "accountability" and "control", but the descriptions vary. Similarly, the chapters on roles and responsibilities show variations, perhaps due to partly different policies. Hence, dealing with the same topics, different things can be said.[8]

Implications for manual design: *Sida must decide which topics are to be included in the manual. The list on the previous page and Annex 4 may serve as a first point of departure. Sida must further decide what is actually to be said in respect to these different topics. This is partly a policy issue.*

---

[7] Some manuals only included a list of the main headings. We have then prepared a list of content by using the headings actually found in the manuals.

[8] A distinction may here be made between topics which basically refer to technicalities, e.g. how to conduct an interview, and topics with an important policy dimension, e.g. who should participate in evaluations.

2.2.2 Focus, scope and delimitations

Each manual pursues different themes and ideas and focuses on some particular issues. Thus the foreword to the Norwegian manual explains that "it is structured around three dimensions: the goal hierarchy, general evaluation components and a set of development or crosscutting issues." It is considerably LFA-influenced. The theme of the UNDP manual is "monitoring and evaluation" and the linkages between the two. The IDB manual places a relatively heavy emphasis on evaluation aspects in the project preparation and implementation stages while the EU manual focuses on the management and actual implementation of evaluations. The current Swedish manual places relatively heavy emphasis on evaluation criteria and methods.

Manuals are limited in scope in different ways, including the following[9]:

Stages in the project cycle considered: The Finnish and IDB manuals are good examples of documents dealing with the entire project cycle, in which evaluation issues enter along the way. The EU and the Danida manuals represent the opposite, dealing exclusively with the actual evaluation phase, paying limited attention to what comes before and after. The UNDP manual on "Monitoring and Evaluation" lies somewhere in between. With a broader scope, "evaluation" can be explained in its context and important issues such as how to plan for evaluations and the use of evaluation will have a natural place. However, with a fixed number of pages it also means less depth. Furthermore, the project or programme cycle perspective may not always be the most relevant to, say, thematic evaluations.

Objects of evaluation: UNDP manual states that it is only concerned with *project* performance, making no reference to evaluation at programme, policy and strategy levels. The IDB manual has a similar scope. Others, particularly the Nordic ones, make a point of the fact that there are many different types of evaluations, including thematic and organisational evaluations.

Managing vs. conducting evaluations: The topics relevant to evaluation management and the conduct of evaluations are not quite the same. (See also chapter 3, on competence needs). The Danish and IDB manuals deal extensively with issues relevant to the evaluation manager. The EU and Swedish manuals, entering deeper into analytical issues, evaluation methods etc., are likely to be more useful to an evaluator.

Practical and theoretical: A manual may be more or less practical, stating what to do and how to do it, or theoretical, i.e. explaining the underlying principles and the reasons for doing things in one or the other way. Here again, all manuals present a mixture, which does not preclude a tendency in one or the other direction.

Furthermore, the organisations have a number of guidelines, manuals, and tools that should complement each other. The UNDP document, for example, is one in a series which also includes, among other things, a document about participatory evaluations. The organisation also has separate "guidelines for evaluators". Apart from its Procedures Guidebook, USAID has a series of *"Evaluation Tips"* on different topics, addressing a broader audience. As mentioned, the Finnish document is something of a combination between policy document for the whole project cycle and a manual. The choice of scope and delimitation must be made with consideration to what is written in other documents.

---

[9] Delimitation is not only a question of including or excluding a particular topic. They are made just as much in terms of the *depth of discussion* on the respective topics. A reflection here can be made regarding how the so-called crosscutting issues (gender equality, environment etc.) are presented. Although they are usually included, the treatment of them is often meagre and most manuals give limited practical guidance.

Implications for manual design: *Sida should consider which main themes and ideas to pursue in the evaluation manual. It must also define scope and limitations in relation to, inter alia, i) stages in the project cycle, ii) different objects of evaluation, iii) management vs. conducting evaluations, and iv) practical vs. theoretical aspects.*

Consideration has to be taken to the intended function of the manual as well as what is stated in other Sida documents.

2.2.3 Structure

By structure we mean the order in which the different issues are presented, how they are linked together and the main lines of thought in the manual. The structure may guide the reader in different directions and serve to emphasise different issues, aspects and linkages.

What many manuals do is to start with a conceptual discussion on what evaluations are, main guiding principles etc., then go on to a description of the different steps of an evaluation and finally dig deeper into areas of particular concern, such as evaluation methods and the use of evaluations.

Manuals tend to have more or less of a *"process"* structure, as opposed to a *"thematic"* structure. The EU manual starts with a relatively short discussion on "Key concepts and definitions" followed by "Preparing and managing evaluations", "Conducting evaluations" and "Reporting and dissemination". The IDB manual follows the steps in the project cycle and USAID follows the different steps in the actual evaluation. The presentations are based on *processes*. However, in the Danish and Swedish manual, "The evaluation process" is merely one chapter among others and the different themes are presented under their own headings. Also the UNDP manual is basically thematic. An advantage of the process structure is that it may link to the actual tasks of the reader, giving "hands-on" instructions. However, to highlight fundamental principles relevant throughout the evaluation process a thematic structure is probably more useful. This may explain the mix found in most reports.

The Norwegian manual is partly structured around a model which links the hierarchy of objectives with the evaluation criteria (impact, efficiency etc.) and a number of development factors explaining results. The Finnish manual makes reference to the same model. This makes the presentation logical and easy to follow.

Are there other possible structures than those observed seen in the manuals reviewed? Instead of making a chronological description of the different steps in the evaluation an alternative would be to depart from different *uses* of evaluation, e.g. learning and control, and then move *backwards,* asking what additional information is needed for this, how the evaluation should be designed to meet such information needs etc. This would give a more "needs-based" structure.

Other alternatives could be to focus on different types of evaluations, discussing how they are implemented, how they can be used, or to begin with different stakeholders and their functions in the evaluation process. Possibly the manual could also be structured around LFA.

Implications for manual design: *The choice of structure will affect how the reader perceives the content. Sida may chose between primarily a "process" and a "thematic" structure. Different models, including LFA may also provide a basis for a structure. The choice depends on how the manual should be used and what issues and linkages Sida wants to emphasise.*

2.3 Format

By format we mean the way the content of the manual is presented. Consideration has to be given to the function of the manual: should it primarily serve as a reference book, as teaching material or as what? Consideration must also be given to the content. How many issues should be covered and at what level of complexity? Below we discuss some of the options in terms of format.

2.3.1 Length and number of documents

How long is an evaluation manual? The EU manual runs to 113 A4 pages, annexes and illustrations included. The Danish is 128 pages long and the Norwegian 123 pages, but these have a different format, with less text per page. The Swedish, US and IDB manuals are all between 80-90 pages long. So is the Finnish, albeit that the part actually dealing with evaluation runs to no more than about 20 pages.

This may give an indication of what has been considered necessary in order to present the subject. However, it must be recalled that the manuals differ considerably in scope, content and structure.[10] The advantage of a longer presentation is that it can be more detailed. The advantage of a shorter presentation is often said to be that the document will be read. However, if too short, there is also a risk that the document will contain little new to the reader and thus become uninteresting. Also, a document being long does not necessarily mean that finding the relevant information in it is difficult.

Furthermore, it is not necessarily so that everything has to be presented as one document. It can be divided to meet the needs of different target readers (e.g. separate manuals for administrators and consultants/evaluators) or for different purposes (e.g. one part explaining the general framework and one part that providing the instruments, like the UNDP manual). There are various alternatives. The USAID Evaluation Tips, which complement the official guidelines, consists of a series of separate papers, each covering a particular issue. The UNDP manual, though divided into two parts, is kept together in the same binding. The document is in itself part of a series of publications.

One way of making documents short and still ensuring that necessary information is available is by using appendixes. The US document goes furthest in this respect, presenting only a 20-page main document and putting everything else in annexes. There is also the loose-leaf option. The advantage of this is that it can be updated gradually and be tailored to different target groups. However, it requires someone responsible for updating and distribution and it is not always easy for the reader to know if he or she has access to the latest edition.

Implications for manual design: *A choice has to be made in respect to the number of pages of the manual, recalling that it must be long enough to be interesting but not so long that it discourages reading. The manual may be published as one or several volumes, with or without appendixes.*

[10] One may wonder if authors have considered, say, a hundred pages as the maximum for any document to be read, and then adapted the content and scope accordingly.

2.3.2 Printed and digital

The development of information technology offers a number of new opportunities, both in terms of accessibility and presentation. On its home page, USAID has a long menu with different evaluation documents, both evaluation reports, policy documents and the "Evaluation Tips". Also the World Bank offers a large amount of interesting information through Internet. Many organisations, Sida included[11], are introducing "Intranet". A computer-based manual at inter/ intranet or CD-ROM may:

- include a larger amount of information without the problem of the manual becoming physically too bulky;
- be continuously updated and thus be more of a "living" document;
- include information tailored for specific groups of readers;
- give flexibility in terms of structure, e.g. direct links between different functions and particular themes;
- be designed as inter-active training programme[12];
- guide the reader to directly to what is relevant for her or him, include question-and-answer functions etc;
- be linked to Word documents, and include the formats for e.g. terms of reference, protocols for evaluation of tenders, correspondence etc.
- be accessible to anyone who has access to the web site or CD-ROM without any need of physical distribution.

However, there are also limitations or disadvantages:

- Longer texts are usually easier to read on paper;
- When working with other documents in the computer, it may be more convenient to look for information in a book;
- A physical document on the bookshelf may serve as a constant visible reminder of the existence of a manual;
- Not everybody searches actively for information on the computer;
- Users' computers have to be equipped with the right programs and functions;
- It is impossible to bring the manual to the "field".

A computer-based manual is somewhat more expensive to produce than a printed one but may be cheaper in the long run, as there are no distribution costs, and updating is less expensive. Otherwise, the production process is much the same for both: careful selection of topics, establishment of a structure, preparation of text etc. Differences in cost should therefore not be exaggerated.

Implications for manual design: *given the advantages and limitations of both a printed and digital manual, the two should be seen as complements rather than as substitutes. Most of the work to prepare the manual will be the same irrespective of media.*

[11] For an overview, see Projektbeskrivning, Sidas Intranet, 22 December 1998.
[12] Such a programme, on project work, has recently been produced for Sida and may serve as an example of what is possible.

2.3.3 Shape and layout

The broad variety of available options in terms of shapes and layouts can be demonstrated simply by placing Sida's different policy documents, guidelines and handbooks side by side. The Swedish, Norwegian and Finnish manuals are all produced as soft cover books, while other manuals, such as the UNDP's, have steel bindings and more of a "manual" look. Both send different signals to the user.

The choice of *layout* can make the manual more or less attractive, easy to find information in and easy to read and to learn from. Among other things, the following may be considered:

- Pictures that visualise what is said, e.g. processes and linkages (Pictures having no connection with the message should be avoided);
- Use of colours, indicating e.g. different topics or different chapters;
- Use the symbols (e.g. "!" for important and "?" for points to reflect over);
- Text in the margin for quick reading;
- Headings explaining what the main message of the text e.g. "there are many who have an interest in the evaluation" instead of just "stakeholders".
- Presentation of check-lists and similar;
- Use of examples ("good and bad");
- Boxes for the presentation of particular issues;

The manuals reviewed differ considerably in terms of layout. The Norwegian one uses its evaluation model, showing it in each chapter, and shading the parts of the model to be discussed. Mostly it presents a text on the left page and a checklist or similar on the right side. This makes the presentation logical and clear to the reader. However, most of the manuals are rather conservative in terms of layout. The EU manual basically looks like any report. There are some listings with bullet-points and illustrations but mostly it is massive text. It is well-written but for many readers it is likely to be heavy to penetrate. It is not known to what extent manuals employ standard formats of their respective institutions.

Implications for manual design: *For obvious reasons, layout is difficult to write about. For further discussion, it is suggested that a number of documents actually be put on the table. Sida should consider whether to adapt the format to that of other publications or choose one that makes the evaluation manual conspicuous.*

3 What are the needs and demand for a Sida Manual?

So far, the report has only presented different *options* for the design of a manual. We now turn to the *needs and demand,* so as to determine what are the suitable options for Sida.

This chapter starts by looking at what *overall policies* state with respect to evaluation. We then turn to the *evaluation policy,* which is more specific and therefore also serves as a good starting point for a discussion of actual practice and *needs to improve or change.* Finally, we consider needs and demand in terms of *competence, procedures* and *routines* as well as *practical requirements* on the manual.

3.1 Overriding policy documents and guidelines

3.1.1 Sida's programme documents

Sida looks forward, Sida's programme for global development, outlines Sida's policy at a high level. It describes the objectives and types of co-operation that will have to be evaluated as well as main guiding principles for Swedish development co-operation. These have indirect implications for evaluation work and the evaluation manual.

It points to the need for *learning*, by stating that knowledge is the key to development and that Sida will carry out a programme which systematically raises levels of knowledge and skills in partner countries, of Swedish partners and at Sida (p. 29). "In order to be able to meet changes and new demands, Sida will keep a continuous *learning process* in order to utilise and benefit from the experience gained by all members of staff" (p. 36-37). "Sida will take the initiative to have independent evaluations made of projects to obtain information about the strengths and weaknesses of operations and to learn lessons from this information. Evaluations are an important instrument for both *securing quality* and for *transformation"* (p. 37). These statements are reflections of Sida's commitment to learning through evaluation.

Another main theme is *"partnership".* Co-operation partners should take more responsibility, particularly for needs assessment, planning and implementation, and Sida stay more in the background, participating in the dialogue, providing finance and follow-up on results (p. 36). This has implications for who evaluations are for, who should participate in them and hence also for who a manual may be used by.

Sida's program to fight poverty (Minskad fattigdom) states that support will be given for research and for elaboration of *new methods* to analyse poverty, measure the results of action to fight poverty and evaluate the effectiveness in different types of poverty strategies (p.30). The document makes reference to Sida's evaluation policy, which states that *the effects on poverty should always be considered,* and it outlines the responsibilities of different Sida departments (p. 31-32). The action plan states that the evaluation activities should be given *clear directives* regarding the analysis of the effects of poverty of Sida's work. The responsibility for evaluation rests on all departments while UTV is said to have an advisory function and carry out strategic evaluations. Evaluation of poverty should be done according to the evaluation plan with annual reporting. The document finally states that UTV prepared guidelines for poverty focus during 1997[13] (p. 73).

[13] A statement which according to interviews with UTV seems to rest on some misunderstandings.

Sida's programme for sustainable development (Omsorg om miljön) states that UTV systematically assesses to what extent environment aspects and guidelines on environment are adhered to, internally and in the different programmes (p. 18). It is of strategic importance that environment aspects be *integrated in all documents* by which work is guided. The *methods* for evaluation of environment aspects in development co-operation should be further developed (p. 32-33). The recommendations of the study "Environment and Sida's Evaluations" (Miljö och Sidas utvärderingar) should be followed up[14] and a standard phrase regarding evaluation of environment aspects/consequences should be introduced in the terms of reference (p. 57). Regarding development of methods, the document defines the objectives *as increased integration/analysis* of environmental aspects in connection with project and programme evaluation as well as *increased internal control* to see to it that such aspects are integrated.

Since the programme was written, Sida has issued *guidelines for environmental impact assessments.*[15] These explain the function of such assessment, the process and different levels and types of environmental consequences. It also presents "good practice" and a standard outline for terms of reference regarding environmental impact assessment. The document can be used in all stages of the project cycle, including evaluation.

Sida's programme for democracy and human rights (Rättvisa och fred) is rather vague with regard to follow-up and evaluation. It states that where circumstances so permit, Sida should give *more attention to result reporting and evaluations,* in order to assure that the support reaches the desired and planned results. Evaluation of results and effects in these areas (democracy and human rights) are said to be, for natural reasons, a *complex problem* and Sida will, in co-ordination with its partners try to develop *better methods* for this purpose (p. 64). The experience analysis notes that the effects of development co-operation on democratisation and peace processes are *difficult to measure* in a way which make them *comparable* over time and between countries.

Sida's programme for gender equality (Kvinnor och män) notes that gender equality is not dealt with adequately in sector reviews and evaluations due to a lack of concrete and measurable objectives, and that few efforts are made to analyse the situation of men and women. *Impact indicators* must be developed (p.42-43). On the positive side, the document notes that it is generally accepted within Sida that the promotion of gender equality should be *included in the terms of reference* for sector reviews and evaluations. However, there is a *need for more detailed guidelines* on how to apply the terms of reference more effectively. The *sector-specific handbooks* give some guidance on to how gender equality should be *mainstreamed* into reviews and evaluations (p. 49). However, the changing character of development co-operation requires new approaches. *New methods and instruments* will have to be developed (p. 50). Guiding principles include the imperative necessity of including a *gender equality perspective* in analyses made in the course of preparation, implementation and follow-up, in particular within LFA (p. 79). Evaluation should determine whether there *are suitable goals, strategies and indicators.* UTV, according to this document, should ensure that the gender equality objective receives proper attention[16] in studies and advice to the Sida departments (p. 84-85). Methods must be developed to ensure that consultants have the appropriate competence to work with equality between women and men. This includes i) requiring documented competence and

---

[14] Tom Alberts and Jessica Andersson, 1996. The recommendations of this report included improvement in the primary phases in the project cycle, baseline studies, better monitoring, guidelines on how to carry out environmental impact assessment with insufficient data and information, methodological discussion and strengthening of competence.

[15] Riktlinjer för miljökonsekvensbedömningar i utvecklingssamarbetet, Juli 1998.

[16] According to UTV (Ann-Marie Fallenius) this is incorrect. UTV cannot take responsibility for the quality of the departments' evaluations.

experience of gender equality work, ii) emphasising gender equality in tenders, iii) evaluation of tenders with a gender equality perspective, iv) mentioning gender equality in terms of reference and work descriptions, v) demanding that gender equality be mainstreamed into the reporting (p. 96-97). The action plan itself should be evaluated when it comes to an end in 2001.

Implications for manual design: *The programmes on poverty, environment, democracy and human rights and gender equality all outline fundamental objectives of Sida's co-operation. Hence, evaluations and the evaluation manual should also contribute to the fulfilment of these objectives.*

The programmes state that the respective objectives should be addressed and integrated in all evaluations. Hence, they should be addressed, and preferably integrated, also into the manual.

The programmes give no practical guidance on how to deal with these objectives in evaluations. On the contrary, they indicate that there is a need to develop methods and strengthen competence. Some areas have made more headway than others: in the field of environment, guidelines for environmental impact assessment have been issued. Such practical instructions are urgently needed to supplement the programme documents.

3.1.2 Sida at work

This document, describing Sida's methods for development co-operation, may be seen both as a policy document and as a handbook.

In respect to roles and responsibilities, it notes that Sida has *a "responsibility to exercise control."* This responsibility must be exercised "in such a way that the co-operation partners retain their *ownership* and Sida does not take command of projects". The document points to the need for a balance between the two (p. 18). Co-operation partners are "to be responsible for the progress of the project by, e.g. (...) evaluating and checking results and costs (...). Part of Sida's role is "to evaluate and approve changes in the planning of the project's organisation and its plan of operations" and "to propose, contribute to and take due note of evaluations of the project (p. 72-73).

Sida at Work mentions that Sida has decided to use the *Logical Framework Approach* (LFA) as an instrument to provide a structure for its own analysis of project proposals and follow-up of projects (p.20). Evaluation is described as part of the *project cycle* and the difference between evaluation, monitoring and audit is explained. The purpose of evaluation is said to be to get "more detailed information on results and impact than can be obtained from the progress reports provided by the project" (p. 50). Reference is also made to the *country strategies* and the *results analyses* which should be based on, among other things, evaluations (p.36).

One section deals exclusively with evaluation (p. 88-90). Evaluation is defined as *"a systematic and thorough assessment of the organisation, implementation and results of a project."* It involves a *deep* and *broad* analysis of the project, its *context* and long-term *impact.* The ambition is "to establish a *relationship* between activities, results and impact." The *uses* of evaluations are said to be, "on the one hand the *further organisation of the project"* and on the other hand, *"to provide experience and lessons for future co-operation*."

Regarding *quality,* Sida is said to make strict requirements, meaning, among other things, *scientifically acceptable procedures* and the participation of *external evaluators*, with the aim of ensuring that the assessments and conclusions are *reliable* and *impartial.* Evaluations should *"not be used* to postpone decision nor to justify decisions which have, in practice already been made. Neither should they be used to assess facts and information which the project should obtain through its regular follow-up of activities and results. The reason for an evaluation is to seek new information and new relationships, above all regards the impacts and consequences (...).” Specific reference is made to *Sida’s Policy on Evaluations.*

Implications for manual design: *Sida at Work provides an overall framework, for evaluation activities, as described above, without going into detail on e.g. the different roles and responsibilities in evaluations, practical uses of evaluations and ownership. Just as Sida’s programme document, it has to be supplemented with operative guidelines for persons dealing with evaluations.*

Sida’s evaluation policy, discussed in the following chapter, is a first step in this direction. It largely reflects[17] what is stated in Sida at work but is considerably more detailed and therefore a more useful point of departure for what should be stated in the manual.

3.1.3 Instructions for project cycle management and LFA

It is assumed that the reader is familiar with the 1996 instructions for LFA, describing a framework for project analysis, follow-up and evaluation. Sida’s experiences of LFA have been mixed. A review of the introduction of the LFA method in Sida “Two Years After”[18] noted that it *can* be a useful tool, but that it has been applied primarily in the project preparation phase while having been under-utilised in the implementation and evaluation phases. The report therefore recommended a *project cycle perspective* “where the application of the LFA method at the various steps of the cycle is given equal attention” (p. 30).

The LFA instructions are now under reformulation and we have studied a *draft* for the new ones. Please note that *this is not yet an official document*. However, it may give an indication guidelines to come.

This document describes LFA as an aid for *logical analysis* and *structured thinking* to be used in a *flexible way*, depending on the project context. It has the aim of developing a more consistent use of LFA during the entire project cycle. It highlights the basic principles of LFA, giving less attention to vocabulary and technicalities. The LFA matrix is gone.

The draft document notes the importance of an initial *problem analysis* and *establishment of objectives* (p. 17-19). In respect to *participation,* it states that just as the LFA method builds on participation, so should evaluations. Participation must mean something more than providing information to the evaluator. There must also be a possibility for stakeholders to influence the analysis, conclusions and recommendations of the evaluation.

The paper gives some direct suggestions on points to bear in mind when conducting an evaluation with participation and ownership in focus:

- The evaluation should be user-oriented, i.e. focus on the interests and information needs of identifiable people;
- Intended users of the evaluation should also have a possibility of participating in the decision-making regarding the evaluation;
- High quality rather than quantity; it is better to select a smaller number of stakeholders for close participation than to involve many just for the sake of it;

[17] There are certain differences between the documents, e.g. in their descriptions of evaluation use and their definitions of “evaluation”. However, please note that this report deals with the implications of different policy documents on a manual, not with the harmonisation of different policy documents.

[18] Carlson, J., Eriksson-Baaz, M., 1998.

- The stakeholders must understand the method and approach of the evaluation; this knowledge should not be kept to the evaluator alone;
- Sensitivity for the project environment, including consideration for different stakeholders' possibility to participate and use the evaluation;
- The timing has to be right, not only with regard to the decision making process of the project but also to those of the other stakeholders;
- The recommendations have to be realistic i.e. it must be possible for the different stakeholders to act on them.

The document ends with a chapter on feedback of knowledge and experiences, focusing on three questions: *who* should learn, *what* shall they learn and *how* shall they learn.

Implications for manual design: *As Sida has decided to use LFA as a tool for analysis in the entire project cycle, including in evaluations, it has a natural place also in the evaluation manual. LFA may be taken as a point of departure for discussions on e.g. objectives, relevance and efficiency, as well as stakeholder and target group analysis. However, it may be less relevant in connection with thematic and organisational evaluations.*

As the LFA guidelines and evaluation manual partly will overlap each other it is important that they are consistent. Depending on when the LFA guidelines will be finalised, there might be room for adjustments in both directions.

Unnecessary overlapping between the two documents should be avoided. Given that the LFA instructions state e.g. how projects should be planned and monitored to permit evaluation it may be questioned whether an evaluation manual should do the same. If the evaluation manual focuses on the evaluation phase and makes reference to the LFA instructions for planning and monitoring, the two documents may complement each other.

3.1.4 International Policies

There are also a number of international policies to which Sida is committed. Those established by DAC deserve particular mentioning. As these are quite extensive, no particular review of them will be made here. Reference is made to the *DAC Manual for Development Assistance*.

Implications for manual design: *The manual should be consistent with the policy and terminology of DAC.*

3.2 Evaluation policy and practice

Sida's current evaluation policy from 1995 is now under review. We have studied the latest *draft version* of a new evaluation policy. *This is not an official document,* but it may be assumed to be what comes closest to the future policy. It probably also reflects what is already largely the *de facto* policy of UTV.

We shall contrast the policy with the findings of some of the reports that have been made on Sida's evaluation activities as well as observations from interviews. For this purpose, quotations will be made of what has been considered as key sentences in the draft evaluation policy.[19]

[19] As special attention is given to *differences* between policy and practice, the description may have a certain negative bias. However, it should be recalled that the review is done with the very purpose of identifying areas which need to be improved.

General

> *"Evaluations of development assistance through Sida are implemented by the sector and regional departments and by the Department for Evaluation and Internal Audit (UVT). Each operative department carries out evaluations within its own area of activity as part of its operational responsibility."*

The report *Using the evaluation tool*[20] starts by noting this "combination of a widespread – even diffused – operational responsibility at the level of departments, and a centralised, semi-independent evaluation function" (p.11). Evaluations are mainly initiated and managed by the programme officers.

Implications for manual design: *The decentralised organisation implies that the number of people involved in evaluation activities will be relatively large and that these people will be handling evaluations only now and then. Therefore, they are probably not evaluation "specialists". This indicates a need to include both very practical instructions and the underlying principles. The manual must not be too advanced. Rather is should give the reader a general orientation of a number of different issues.*

Objectives

> *"Sida's evaluation activities have two main objectives:*
> *- learning (promotion)*
> *- control (accountability)"*

Most evaluations, one way or another, fit in with these two rather broad objectives. Looking through Sida's evaluation plan, it can be seen that an overwhelming part of evaluations are carried out to decide upon future financing. The report *Using the evaluation tool*, applying a somewhat different categorisation of uses, notes that the purposes are i) management, ii) learning and iii) audit, in this order of importance.

However, the same report also indicates that evaluations are often loosely initiated. "Evaluation efforts are not defined on the basis of actual needs. There are no specified goals, nor is there any clear notion of what questions the evaluations are to answer. Also, in many evaluations it is difficult to distinguish the needs of different actors" (p. 66). "Indeed, rather than elaborating on the initial reason, thinking through what information is needed and setting priorities among conflicting purposes, the terms of reference are almost copied from the manual. There is a lack of practical focus (...) and instead we find shopping lists that are called purpose". This problem was recognised by some of the persons interviewed and can be seen with a quick glance at some terms of reference.

Then, how are evaluations actually used? The follow-up of Using the evaluation tool, *Are evaluations useful?,* notes a "complex pattern" of uses. Direct or instrumental use[21] is said to be surprisingly rare, primarily for the local stakeholders. "Other types of use are more common, e.g. conceptual use. Using evaluations for legitimising decisions already taken is common, like the use of evaluations as ritual flag waving" (p. ii)[22]. A follow-up study of the SIDA evaluation Promoting

[20] Carlsson et. al., 1997

[21] According to Leviton & Brach, 1983, *instrumental use* refers to "documented and specific use of evaluation findings by decision makers and other stake holders" as opposed to *Conceptual use* which refers to "the impact of evaluations on policies, programs and procedures through sensitising persons and groups (...)" and *Persuasive use*, i.e. "enlisting evaluation results in efforts either to support or to refute a political position already taken (...). The typology used in *Are evaluations useful* was *conceptual* use, *legitimising* use, *ritual* use and *no* use.

[22] The two reports quoted here have had a rather strong focus on *learning*. Documents reviewed for this study contains surprisingly little information about the control purpose of evaluation.

Development by Proxy[23], seemed to indicate that use of this evaluation had been primarily conceptual.

Sometimes there are several objectives. *Are evaluations useful?* raised the potential problem of having several, conflicting objectives. Evaluating with the purpose of establishing accountability and with the purpose of learning and management of projects may for example require quite different approaches (p. 58).

The same report also discusses *to whom* evaluations may be useful, and concludes that they are useful to a very limited group of stakeholders and primarily to Sida.

Evaluations should also serve as an input for Sida's results analyses. However, according to studies carried out by UTV, evaluations often do not present sufficient lessons learned for this purpose.[24]

Implications for manual design: *Findings indicate that there is a need for a manual to go into depth on the issue of what evaluations can contribute to and how they can be used. For example, what can be learned from evaluations, how does learning take place, what structures can be created to enhance learning and how can lessons learned be fed back into operations? How and why is control exercised? A manual can also point at other dimensions of use than those mentioned in the Sida policy e.g. instrumental and conceptual use. The issue of whom evaluations are for should be addressed as well as potential conflicts between different objectives.*

The concept of evaluation

> *"An evaluation is a careful and systematic ex-post assessment of the design, implementation and results of an activity. An evaluation may be concerned with one or several aspects of ongoing or competed projects, programmes, action plans, or policies. An evaluation may also take the form of an assessment of one or several aspects of the capacity of Sida, or of organizations supported by Sida (...)"*

Using the evaluation tool notes that the Sida evaluation reports basically fit within this definition. However, interviews show that there is some uncertainty about where to draw the line between evaluations and "studies" in general.

Evaluation is defined somewhat differently in Sida at Work and Sida's evaluation policy. Other definitions are found in the evaluation manuals reviewed.

Despite the fact that the definition permits a considerable diversity in terms of evaluation objects there is a strong focus on project evaluations. This has implications for what is actually learned and at what level control is exercised.

Implications for manual design: *The manual should serve to explain Sida's definition of evaluation but recognise that there are other definitions as well. It must go beyond semantics and give the reader an actual understanding of what an evaluation is. It may also discuss what an evaluation is not.*

The manual should emphasise that there are other objects of evaluation than projects, e.g. sector programmes, organisations and policies. This may be a starting point for a discussion on the need for different approach and methods.

---

[23] MPI, Managing the NGO Partnership, A study of the impact of the SIDA evaluation Promoting Development by Proxy, First draft report, November 1998.

[24] "Huvudproblemet är att avdelningarnas projektorienterade utvärderingar som regel inte innehåller den typ av generaliserbar information om effekter och orsakssamband som förutsätts i resultatanalysens riktlinjer", Sida memorandum of 980520. For a discussion on the subject, also see Schill, G., Granskning av resultatanalyserna i Sidas landstrategiarbete.

> *"Evaluation cannot be sharply distinguished from monitoring, the continuous assessment of project progress through written project reports, periodic reviews, and informal communication, but in practice the distinction is usually clear enough."*

Two problems may be noted. First, there is a certain, perhaps unnecessary, confusion in terms of terminology on where to draw the line between evaluation and monitoring. Second, there is often an actual lack of monitoring, which makes evaluation work more difficult and time consuming.

Implications for manual design: *The manual should explain the difference between the concepts of monitoring and evaluation. It may also point at the linkages between the two, e.g. the importance of monitoring for evaluability, ability to point at casual relationships and gaining "lessons learned". The linkage between the time requirements (and thus also the cost of an evaluation) and the existence of monitoring information should be highlighted.*

> *"(...) evaluations must be carefully conducted and based on systematic methods of observation and analysis that are consistent with those of science. As a rule, studies that violate this principle should not be accepted as evaluations."*

Using the evaluation tool noted considerable methodological weaknesses in the evaluation reports analysed. It also noted that methods were poorly accounted for in the reports. It is not known if this is a sign of inadequate competence or, considering evaluations have been approved by Sida, a lack of actual demand for higher methodological standards.

Not only evaluators but also the person commissioning the evaluation must have basic orientation of methodological problems and evaluation methods in order to set requirements and to assess the quality of evaluation reports.

Implications for manual design: *The manual may underline the importance that Sida places on the use of proper methods. It should discuss some of the more common methodological problems (e.g. fungibility and different evaluability problems) and give a basic orientation of the suitability of different approaches and methods in different evaluation contexts.*

Evaluation criteria

> *"When planning an evaluation of a project, programme, action plan or policy, and writing the Terms of Reference for any such evaluation, the following aspects of the activity to be evaluated shall always be taken into consideration. (Relevance, Effectiveness, Impact, Efficiency, Sustainability, Replicability)"*

The draft policy is not completely clear whether this should be interpreted as if analysis *should* be done with all these criteria or whether they are criteria that may be *considered* for an analysis.[25]

Most terms of reference include all of the evaluation criteria mentioned above, apart from "replicability" which has not been one of Sida's "traditional" criteria. This is often the case also for small evaluations. There seems to be a tendency to include all of the evaluation criteria for fear that something important will be missed.

However, the treatment of evaluation criteria in the report leaves much to be desired. According to *Using the evaluation tool*, adequate treatment of fulfilment of objectives was found in 57% of the reports reviewed, relevance in 30%, development impact in 23%, cost effectiveness in 19% and sustainability in 25%. It is not clear whether this is due to lack of competence, over-ambitious

[25] This issue has not yet been fully resolved; UTV, Stefan Molund.

terms of reference in relation to the time for analysis or that there has not been a genuine demand for more thorough analyses.

Implications for manual design: *Apart from explaining, at a conceptual level, what the evaluation criteria refer to, a manual may discuss when to use which criterion and conditions for evaluability. The manual may also give practical suggestions for how analyses should be made, particularly when data is limited.*

Development goals and cross-cutting issues

> *"Sida's operative departments are required to review all its projects and programmes from the point of view of Swedish development goals and policies regarding poverty, economic growth, democracy and human rights, the environment and gender. If a particular goal or policy is not regarded as relevant to a project or programme under review, this should be explicitly stated in the evaluation report."*
>
> *"This part of the policy cannot be properly implemented, unless Sida and its partner organizations make sure that evaluation teams possess the expertise needed, for the analysis of poverty impact, gender analysis, environmental effects, or whatever the task at hand may be. Obviously, the terms of reference for the evaluation must include precise instructions regarding the issues at stake."*

The evaluation policy reflects what is stated in Sida's different action programmes. It may be noted that "economic growth", for which there is no particular action programme, is also mentioned, but not "economic and political independence", which is another of Sida's official development goals.

The studies that have been made concerning the treatment of overall Sida objectives in the evaluations all note considerable weaknesses. A 1998 follow-up on how *gender* equality is dealt with in evaluation[26] concluded that about 65% of the reports, and somewhat fewer terms of reference, mention gender in one way or another. The main problem, however, was found to be the quality of analysis. The report also noted a lack of concrete instructions for how gender equality *should be* dealt with and elaborated on some possible measures to improve evaluations with respect to the assessment of gender equality.

The study basically confirmed the picture painted in another report written in 1997, which also included *poverty* issues.[27] In its summary, this report noted that "by and large, scant attention is paid to poverty (...) issues" and that the analysis with few exceptions is "superficial" and "usually limited to a few comments on poverty based on simple income criteria". It was found that the terms of reference "almost never mention poverty reduction (...) as a major objective of the project and rarely ask the consultants to pay special attention to these issues". "The lack of established methodology, and quite often, data, for measuring the impact on poverty reduction" was said to be "evident."

A similar study was carried out in 1996 on *environment* issues.[28] Almost 60 % of the evaluations assessed or discussed the long-term impacts of the project but only 13% of the evaluations included the environment in the discussion. A lack of reference to methodological literature was noted. The report stated: "In sum, it can be argued that evaluations of Swedish international development co-operation generally has ignored environmental effects"(p. ii). It suggested improvements in the formulation of terms of reference, the Environmental Impact Assessment

---

[26] Peck, L., Evaluating Gender Equality – Policy and Practice, Sida Studies in Evaluation 98/3.
[27] Tobisson, E., de Vylder, S., Poverty Reduction and Gender Equality, An Assessment of Sida's Country Reports and Evaluations in 1995-96, Sida Studies in Evaluation, 97/2.
[28] Alberts, T., Andersson, J., The Environment and Sida's Evaluations, Sida Studies in Evaluations 96/4.

system, the monitoring of the environmental impact of projects and the quality of baseline studies.

No equivalent study has been made of the treatment of democracy and human rights. It has often been stated that the evaluation in this area is inadequate and that methods need to be developed. For this purpose, Sida has initiated a number of evaluations in the field of democracy and human rights.

One common problem for the areas mentioned above is that they are often not given priority. It may be questioned if there is always a genuine demand for analysis of all of Sida's development goals and crosscutting issues.

Implications for manual design: *A manual should re-emphasise the need to consider all of the crosscutting issues and overall development goals of Sida in all evaluations. Reference should be made to the use of evaluations which, inter alia, is to learn and exercise control in relation to gender equality, poverty, environment etc. Reference could also be made to Sida's planning process which requires data and "lessons learned" in these areas.*

The manual should give practical instructions on how to analyse the different issues. Given that new methods are developed, as stated in the Action Programmes, there will be a need to spread these, within Sida as well as externally to co-operating partners and consultants.

The manual must also address the questions of how to find a suitable balance between different crosscutting issues in an evaluation, and how to "mainstream" five different aspects without loosing focus. The evaluability aspect and the question of what it takes in practical terms (including time) to make an adequate analysis of the crosscutting issues in an evaluation must also be addressed.

Objectivity and impartiality

> *"Evaluations should be done with due concern for objectivity and impartiality. This concern should permeate evaluations from design and planning to reporting and feedback. By objectivity is meant that value judgements are clearly distinguished from statements of facts, and that statements of fact are based on trustworthy methods of observation and inference. The meaning of impartiality is that differences in perspectives between stakeholders involved in the activities under review are properly taken into account."*

We do not have sufficient information to state whether or not Sida evaluations are objective and impartial. What can be said is that it is subject to discussion. *Are evaluations useful?* argues that evaluations never can be seen as neutral or objective or as representing some outside neutral perspective. However, this is said to become a problem only when the evaluation arena is characterised by dominance arising out of dependency (p. ii). *Using the evaluation tool* questions impartiality of the evaluation reports by noting that consultants often are procured informally and that Sida sets the agenda for the evaluations.

Only these discussions show that objectivity and impartiality are intricate matters, that there may be different interpretations and that it is important to have an awareness of potential problems.

Implications for manual design: *The evaluation policy's definitions of objectivity and impartiality should be reflected and explained in the manual. Furthermore, the practical consequences of the policy should be discussed; how, in practical terms, are the objectivity and impartiality of an evaluation assured? The manual should recognise the possibility of conflicting interests in an evaluation and discuss how these can be dealt with. As evaluation means assigning a value of determining the merits of something, the question of how to distinguish between value judgements and statements of facts should be given attention, as well as how to deal with different perspectives of different stakeholders.*

Transparency and co-operation

> *"Evaluations should be transparent and carried out in a spirit of co-operation. Interested parties at Sida and in the recipient countries should be informed of evaluation plans at as early a stage as possible, and, whenever feasible, be provided with the opportunity to participate in the evaluation process. Interested parties in the recipient countries should be consulted at every major step in the process, including the writing of the Terms of Reference and the recruitment of the evaluation team."*

Studies indicate that there often is a huge gap between policy and practice. When summarising the evaluation co-ordinators views, *Using the evaluation tool* describe the evaluation process as something "very decentralized, highly personal and informal, relevant only for those most directly concerned with the particular project or programme being evaluated (...)". "According to the co-ordinators, most of the work in preparing an evaluation occurs within the division." "The real influence of the recipient is assumed to be minimal"(p. 14-15). "The picture that emerges is one of a limited discussion between stakeholders about the evaluation"(p. 28). "Participation of the recipient country actors is weak. Sida sets the terms for the evaluations, and co-operation partners are not allowed or encouraged to participate" (p 66).

The follow-up study *Are evaluations useful?* reaffirms this description. Stakeholders are frequently unaware of Sida's evaluation agenda and have limited possibilities of affecting it. It describes evaluation as a "rather undemocratic exercise with little concern for any other needs than Sida's own"(p. 59). "A creation of ownership requires a change in the decision making and influence over the evaluation process"(p. 58).

However, we have also seen exceptions to the negative picture presented above. For example, some of the evaluations that have been carried out in the education sector have been characterised by a high degree of co-operation.[29] The interviews with Sida officers made for this study give an impression that the level of co-operation varies a great deal. Furthermore, there are many different forms of co-operation.

One may wonder whether problems of transparency and co-operation are due to lack of time, lack of knowledge of how to go about to achieve transparency and co-operation or simply that it is not given priority.

Implications for manual design: *The evaluation policy's position on transparency and co-operation should be reaffirmed in the manual. The manual may further elaborate on questions such as co-operation by whom, in what forms and at what stages of the evaluation? How is transparency achieved in practice? It may, for example, provide standard routines for consultation, dissemination of documents etc.*

The policy statement on objectivity and impartiality also justifies a discussion on what co-operation is feasible and advisable in different contexts.

The principles of transparency and co-operation also have implications for who the manual should be directed to. To achieve actual co-operation and dialogue, also the co-operating partners must have sufficient evaluation competence. This indicates that a manual should be written not only for Sida but for all of those co-operating in the evaluation process.

[29] E.g. Agneta Lind, "Free to Speak Up" 1995.

Participation, credibility and commitment

> *"To avoid conflicts of interest and safeguard the credibility of evaluations, individuals should normally not be engaged as evaluators in activities in which they have a personal stake. (...) The requirement that evaluators should have no stake in evaluated activities is particularly important where evaluations are undertaken for the purpose of control."*

As noted by the policy, there may be conflicts of interests in an evaluation. Typically, Sida must assure that the Swedish taxpayers' money is put to good use and the representatives of the co-operating country want to ensure continued funding.

Implications for manual design: *To create an awareness of the problems mentioned above, the manual should discuss what conflicts of interest may arise in evaluations. There is also a need to discuss the suitability of different divisions of roles and different forms of co-operation, depending on the purpose of the evaluation and the evaluation contexts. The discussion may be linked too issues such as ownership, transparency, credibility, partnership, commitment, and evaluation use (learning and control). It must emphasise what Sida's policy actually states and give practical guidance on how to live up to the policy.*

> *"Participation of evaluators or stakeholders from the recipient country is frequently a prerequisite for the credibility of an evaluation (...) In evaluations as well as in other areas of development co-operation participation is often a necessary condition for commitment."*

Above, it was noted that the direct participation of local stakeholders often was limited. Regarding the use of local consultants, we have different information. *Using the evaluation tool* noted that only 6 of the thirty evaluation teams analysed included local consultants. In many cases, the use of local consultants had not even been considered and there was a considerable scepticism towards it (p 24). However, the 1998 annual report for evaluation activities of Sida[30] states that local consultants were engaged in 85% of the evaluations.

Implications for manual design: *The manual should also discuss benefits and problems in connection with the use of local consultants and give practical guidance for how to find such, contract them and give them a suitable role in the team.*

Evaluation planning and monitoring/Selection of objects for evaluation

> *"The following are the criteria to be applied by the departments in the selection of objects for evaluation and establishing an evaluation agenda: (Utility, operational importance and risk, filling gaps in knowledge, evaluation feasibility and cost-effectiveness, balance/coverage)"*[31]

Using the evaluation tool indicates that evaluations undertaken by the operative departments are initiated rather arbitrarily. "The (evaluation) co-ordinators all state that there are no formal rules or procedures governing the choice of evaluation objects" (p.14). Possibly, there has been a change in this respect with the introduction of a system of mandatory reporting of evaluation plans to UTV in recent years.

---

[30] Årsredovisning för utvärderingsverksamheten vid Sida, Verskamhetståret 1998 p. 7

[31] This list may, according to UTV be subject to change. A recent document related to the evaluation activities of UTV (Underlag för val av utvärderingsprojekt för budgetåret 1999, UTV, 1998-11-02) states that what should be considered "strategically important" evaluations primarily is determined by the evaluation objects' policy relevance, financial importance and qualities in terms of innovation and replicability. It is also important to consider evaluability, demand/ usefulness among stakeholders, cost and labour input for the evaluation.

However, the criteria for selection of objects for evaluations may give rise to new questions. How are they to be interpreted and applied in practice? How should they be balanced against each other etc?

Implications for manual design: *The manual should describe Sida's yearly evaluation planning process, including preparation of the evaluation plan and the division of responsibilities. It should also specify the criteria for selection of evaluation objects and how these are to be applied in practice.*

Evaluation and the project cycle

> *"Evaluation is an integral part of development assistance. Projects and programmes should be designed with a view to future evaluations. Provision for monitoring and evaluation should be clearly spelled out in project proposals."*

"Evaluation activities are not integrated in a natural and systematic way with the development co-operation efforts" states *Using the evaluation tool* (p. 66). Observations that support this statement is that projects often are not planned in a way to facilitate evaluation and that evaluations are not always utilised to their full potential. The fact that LFA has been used in the earlier stages of the project cycle but rarely in the later is also an indication of a need for better integration.

Implications for manual design: *Findings justify the discussion in the manual of evaluation in a project cycle perspective, paying consideration also to planning (initial problem analysis, base line studies, target group definition, establishment of objectives and indicators) and monitoring (information required to carry out an evaluation). The use of evaluations, which has already been discussed, can be explained in terms of how evaluations enter in the project cycle. The consistent use of LFA throughout the project cycle should be mentioned.*

However, it should be recalled that there are also thematic and other types of evaluations in which the project cycle perspective is perhaps not the most relevant. Structuring the different topics of the manual around the project cycle only is therefore not suggested.

Reporting and dissemination of evaluation results

> *"The result of evaluations shall be presented in a report, using clear and accessible format and language, which shall be disseminated to all major stakeholders and other interested parties."*

Using the evaluation tool noted that in 27 cases out of 30, the persons who had commissioned the evaluation report were satisfied with its "language, clarity and organisation" (p.34). However, with regard to dissemination, the same report stated that "it is clear that the evaluations are not intended for a broader public. Draft reports in particular have limited circulation", (p. 30) especially to local stakeholders. Hence, the policy requirement that reports be disseminated to "all major stakeholders and interested parties" is not being fulfilled. In its summary, *Are evaluations useful?* state that "The further away you are from the center of the project in terms of power and involvement, the less information you receive on evaluation recommendations and findings" and the report concludes that "Generally there is very little thought given to the important issue of how an evaluation shall be disseminated in an organisation and how organisation learning shall be brought about. There seems to be a rather naive belief in the value of the report as such" (p. 56).

The observations should be considered in relation to what was said about co-operation, participation and transparency. Ultimately, it can be seen as a problem of under-utilisation of the evaluations.

Implications for manual design: *Even if the negative picture painted above is not necessarily true for all evaluations, it shows that reporting and dissemination may constitute a problem and can be improved. The manual may provide concrete instructions concerning, for example, whom to disseminate evaluation results to (and why), different ways of doing so (apart from sending out reports) and how the report should be written so as to be accessible to target readers.*

Feedback and lessons learned

> *"All departmental directors at Sida bear the responsibility of ensuring that the results of Sida's evaluation activities are taken into consideration when forming policy, processing ongoing development assistance activities and taking decisions as to new activities."*

Are evaluations useful?, quoted above, goes on to say that evaluations are of limited value to many stakeholders as they "1) never received any feed-back on the findings and recommendations of the evaluation 2) they never had the power to act on the findings even if they wanted to" (p. i-ii).

A *Sida memorandum*[32] describes more efficient feedback of results and lessons from evaluations as a major challenge. Among other things it is said to be a question of finding suitable channels for dissemination of evaluation reports, establishment of better systems to for action on conclusions and recommendations, and making evaluations available to a broader range of interested parties.

There is no comprehensive analysis of feedback and lessons learned within Sida. However, there appears to be a widespread feeling that lessons from evaluations could be utilised more effectively than today and that the "corporate memory" could be improved.

Implications for manual design: *The problems mentioned above are again linked to the use of evaluations. The manual should point out ways of using lessons learned from evaluation both for immediate and for long-term purposes. The manual should point out the difference between learning in connection with specific projects or programmes, and the institutional learning of Sida as a whole. For example, how can lessons learned in one department be spread to other departments? How can lessons learned by Sida be spread to other stakeholders and how can Sida use lessons learned by other organisations?*

Feedback of lessons learned to activities also has to do with how recommendations are formulated. Therefore, the manual should provide instructions for how to formulate recommendations to make proper use of lessons learned.

Procedures for management response

> *"To ensure that the findings and recommendations of evaluations are properly taken into account by the Sida departments concerned, and that the actions taken as a result of evaluations are adequately recorded, Sida's response to evaluations should always be formally documented in a plan of action."*

This procedure was only established on January 1, 1999 and experiences are so far limited. It is usually described in positive terms when discussed during interviews. However, as it is a new Sida instruction there is probably some uncertainty about how to apply it.

Implications for manual design: *Being a new procedure, the manual should give Sida staff practical guidance on what management response requires from them and inform how it can benefit their work.*

[32] UTV, 1998-05-20.

Quality and quality assurance

This section remains to be written in the *draft policy* and we thus have no policy statement to refer to. However, there is no doubt about Sida's commitment to quality and quality assurance.

The *Sida memorandum* of 1998-05-20 states that evaluations are of a "shifting quality"(p. 9) and that "there appears to be considerable room for improvement." "The descriptions of theory and method are usually very brief and the reports are rarely written in a way that encourages critical reflection" (p.12). *Using the evaluation tool* notes the difference between the Sida staff's generally positive assessment and its own rather critical assessment of the reports as well as the actual evaluation process.

Interviews indicate that there is little or no systematic follow-up on quality. An even more fundamental problem is that there are no generally established quality *criteria*.[33]

Quality control measures can focus either on the finished product or on the process. A main reason for assessing quality already in the process is that when the product is ready it is usually too late to act anyway. Furthermore, in the case of evaluation, it may be argued that the "process" in itself also is an important part of the product; it is during the process that much of the learning and dialogue takes place. ISO 9000 and other quality systems are based on the identification of different functions (e.g. procurement of consulting, dissemination of a report etc.) in processes, guidelines for how they should be performed and follow-up that this is done. Hence, a manual can be seen as a tool contributing to better quality control.

Implications for manual design: *It may be assumed that one of the objectives of the manual is to improve the quality of evaluations. It should therefore state what quality criteria to be applied, such as utility, feasibility, propriety and accuracy.[34] It should also state i) who has the responsibility of assuring quality, ii) in what stages of the evaluation process it should be done and iii) in practical terms, how it is done. The manual may also give information on how to take in "second opinions" and use external competence to assess evaluation reports when needed.*

3.3 Competence

One of the potential functions of a manual was said to be to enhance the *competence* of persons dealing with evaluations. We should therefore look at needs and demand in this particular respect. What competence is required in the different stages of evaluations and what competence is there today?

3.3.1 Competence requirements

If "competence" be someone's *capacity to perform a task,* "competence needs" can be defined in terms of tasks that an evaluator or evaluation manager will be facing.

Tasks can be seen as linked together into different processes, covering different time spans and involving different people. There is the process of actually *conducting the evaluation,* including planning, data collection and report writing. There is the process of *managing the evaluation,* including such tasks as defining evaluation purpose, procure consultants and disseminate the evaluation report. In a wider perspective, they are both part of the *project cycle process* in which e.g. establishment of objectives and indictors, monitoring and reporting are important for a later evaluation.

[33] Let be that the existing evaluation manual provides a useful Checklist for Assessment of Evaluation Reports (p. 73).
[34] These four criteria have been developed by American Evaluation Association and have been adopted by some development co-operation agencies. They are e.g. mentioned in the Danida manual.

Annex 5 includes a tentative list of *main tasks* related to evaluation work. It has been based on interviews as well as on the various evaluation manuals reviewed. The list could of course be broken down into sub-tasks etc. but the purpose at this point has been to get an *overview* of competence requirements.

Linked to the processes and tasks are a number of *questions* or *topics.* For example, the task of assigning evaluators must be performed with consideration for procurement regulations, objectivity and stakeholder participation. Some topics are fairly easily relatable to specific tasks, but not all. Thus main guiding principles such as transparency, co-operation and impartiality, and Sida's organisation for evaluation cannot be discussed only in connection with a specific task. They are relevant throughout the process.

Annex 6 includes a tentative list of *major topics* in evaluation work. The list has been based the literature review, interviews and the previous analysis of needs and demand in this report. Again, each of these topics could be further enlarged on, but the purpose now is to get a first idea of relevant topics for the manual.

To what extent do *different evaluations* require *different competence?*[35] Evaluations may differ in many respects, e.g. evaluation object, purpose, timing, size, methods and divisions of roles. However, despite these differences, there is nothing that indicates that basic tasks involved should be substantially different. Even though the purpose of an evaluation, and the methods to be used, will differ, the actual tasks of defining the purpose and choosing suitable methods will still be there. Hence, also the basic topics for evaluation work, such as evaluation criteria, evaluability, and quality requirements, are likely to be similar for all evaluations.

However, there may be different problems in relation to the evaluation as such, depending e.g. on sector, the type of assistance provided, the evaluation context or the general evaluation approach. The evaluation of an alphabetisation programme and that of support to democracy will of course give rise to completely different questions that will have to be answered by different methods. The list of such *specific evaluation problems* can be made long. Some of them are included in Annex 6.

Evaluations should always be made with consideration for the overall objectives of Sida. Hence, there must be basic competence in the areas of poverty, democracy and human rights, gender equality and environment.

The different tasks involved in an evaluation can be carried out in a number of ways. In order to conduct or manage an evaluation *for Sida,* it is necessary to know how *Sida* wants things to be done, e.g. how to analyse the crosscutting issues, and how to use certain terminology. A distinction must therefore be made between general competence and the *Sida specific competence.*

Finally, *new* issues and policies keep cropping up, as well as new methods and approaches. The development taking place in fields such as gender equality and environment is a case in point. Perspectives change, and so do competence requirements.

[35] Please recall that we are now talking about evaluation competence only. Needless to say, the required sector competence will differ.

Implications for manual design: *Competence requirements can be defined in terms of the tasks to be performed by evaluators, evaluation managers and other participants in the evaluation. A tentative list is attached in Annex 5. There are also topics related to the tasks or the evaluation process as a whole. The question of how to link tasks and related topics is important for the structure of the manual.*

The main tasks and topics are likely to be similar for all evaluations. However, competence needed to solve specific evaluation problems may vary. This indicates that there are some topics that ought to be mandatory in an evaluation while other could possibly be dealt with in separate publications targeted to certain readers.

Competence requirements also vary over time. A manual must therefore not be too rigid but be written in a way that it actually promotes innovation.

3.3.2 Competence requirements of different participants

Defining competence as a person's capacity to perform a task, it is necessary to consider *who* is to perform this task.

In the previous chapter it was noted that roles and responsibilities may differ and, if Sida policy is implemented properly, they will do so even more in the future. Hence, it *cannot be taken for granted* that the Sida officer, project owner or consultant must possess one or the other competencies. It is therefore better to talk in functional terms, e.g. of the *manager of the evaluation, evaluators* and *project owners.*

The competence required by the *evaluation manager* – typically, but not necessarily the Sida officer – is largely of administrative nature. Tasks include preparing for the evaluation, procure consultants, publish and distribute the report. However, the evaluation manager should also identify information needs, set a framework for the evaluation and critically review the conclusions from an evaluation. This requires analytical capacity.

The different tasks to be handled by the evaluation manager were briefly outlined in Annex 5. It is not necessary for the evaluation manager actually to be able to carry out an evaluation. However, in order to draft terms of reference, procure consultants, determine the time for the evaluation etc. he or she must *know* what an evaluation is and what it takes to carry out one: application of different evaluation criteria, common evaluation problems, the suitability of different approaches methods, time requirements, conditions for evaluability etc. Basically, we are talking about the difference between competence and knowledge.

The evaluators should have competence to actually handle the tasks in the evaluation process, such as collection of data, analysis of effects on poverty, gender equality etc. and report writing. As any provider of services, the evaluator must understand the needs of the client and it facilitates if she or he *knows* how Sida manages evaluations.

The previous chapter noted that the role of *project owners* and local stakeholders in an evaluation might vary. Under some circumstances they may actually be the ones managing or conducting the evaluation. Assuming that they are not, they still have an important role in the evaluation, e.g. that of providing information (giving interviews, preparing documentation) and commenting on terms of reference and draft reports. Most likely, they should be users of the evaluation or will at least be affected by its results. Often project owners are expected to sign a contract which lays down that an evaluation should take place.

Project owners must therefore have at least have a basic understanding of why the evaluation is taking place, whether it is for learning or control purposes, what methods are used, and what is expected from them during the evaluation. In order, for example, to comment on terms of reference, the project owners must understand the concepts and terminology used.

Implications for manual design: *As the roles and responsibilities vary from evaluation to evaluation, so will the competence requirements of different stakeholders. It is therefore necessary to talk in functional terms, e.g. "evaluators", "evaluation managers" and "project owners", rather than Sida staff, consultants etc.*

There are differences between the tasks that evaluators, evaluation managers and project owners should be able to perform in an evaluation, i.e. their respective competence requirements (see Annex 5). However, the knowledge that they require is largely the same. In order to co-operate each must have an understanding of the other's perspective.

A manual, outlining the major steps to take in an evaluation and fundamental principles would be relevant to all participants in the evaluation. However, more detailed information, e.g. on evaluation management or evaluation methods, should be targeted to specific groups.

3.3.3 Existing competence, needs and demand

Having outlined what competence is needed, we turn to the question of what competence there is today and where it must be supplemented. Obviously, it is difficult to generalise, but a few reflections can be made.[36]

As mentioned earlier, the decentralised organisation of *Sida* implies that evaluations are handled by persons who also have a number of other tasks to deal with. Evaluation is seldom a main activity. There are programme officers who have extensive experience of managing evaluations, but also a constant inflow of new people who will have to learn things from scratch. There are also those who have handled evaluations their own way for a long time but perhaps are not fully aware of other ways of doing it or of the implications of Sida's evaluation policy. Interviews seem to indicate that evaluations are handled in many different ways at Sida, but that usually each person is set in his or her particular way of handling them.

The report *Using the evaluation tool* noted that "even though most respondents work with evaluations, few have much experience of conducting them themselves" (p.18) and concluded that "The programme officers' competence in administrating evaluations must be improved" (p. 66).

The interviews made for this study gave some indications of what is perceived as difficult, and thus, indirectly, where there is a need to strengthen competence. Preparing terms of reference is mentioned most frequently, perhaps because it actually includes a number of problem areas: How to establish focus and delimitations, how to assess time requirements and cost. Consultant procurement and the planning and timing of evaluations are also mentioned. Consultants interviewed have indicated that Sida officers must become better at designing realistic terms of reference, at defining the evaluation assignment with respect to purpose, scope and limitation, and at reading draft reports.

Competence requirements are often described in administrative terms. However, the problems observed (e.g. defining purpose and scope of an evaluation, estimate the time required given the use of certain methods etc.) indicate that it is largely a question of analytical capacity.

[36] On the subject of *actual* competence, we must here refer to actual groups of people such as Sida staff, consultants and co-operation partners.

In respect to *consultants*, *Are evaluations useful?* states: "(...) although experienced in terms of quantity and possessing sector competence, the evaluators are less qualified in evaluation theory and method. To change and improve the evaluation process no doubt requires more skilled evaluators" (p. 59).

Nuancing this statement somewhat, it is probably possible to distinguish between, on the one hand, "evaluation generalists" with limited sector competence and, on the other hand, "sector experts" with limited experience of evaluation.

In 3.3.1 a distinction was made between general and Sida-specific competence. Some consultants work on a regular basis for Sida, participate in Sida courses etc. However, local and other international consultants are participating in many Sida evaluations and many Swedish consultants work for Sida only occasionally. It cannot be assumed that that these are continuously up to date on Sida policies, preferred methods, Sida terminology etc.

One of the reasons why there is a tendency to contract some consultants over and over again, is that they know how Sida wants the job to be done. By *informing* other consultants of this – in a manual, for example – Sida could broaden its resource base, in particular the international one.

The observations of how crosscutting issues are dealt with indicate a need to strengthen competence. Partly, this is a question of team composition. "An evaluation team composed of, for example, one engineer and one expert in financial economics cannot reasonably be expected to handle issues related to poverty reduction and gender equality in a satisfactory way", states *Poverty Reduction and Gender Equality* (p. 22), and it should then be recalled that poverty reduction and gender equality only are two out of at least four crosscutting issues to be addressed in evaluations. However, if teams should not become excessively large, these issues will have to be handled by non-specialists even in the future.

Regarding project owners and local stakeholders it is virtually impossible to make any assumptions regarding evaluation knowledge and competence. However, *Are evaluations useful?* noted that in many cases, the local project staff had rather vague ideas of what the evaluation was about and often looked at the exercise in a completely different way than Sida. As interviews have only been carried out in Sweden we cannot comment on this. However, it ought to be a reasonable assumption that dialogue, consultation and co-operation could be improved if co-operation partners had access to more knowledge and information on evaluation and Sida's evaluation policy.[37]

Implications for manual design: *Needs to strengthen competence vary considerably within Sida. There are newcomers who need an introduction to evaluation work and there are experienced staff members who still need to refresh and broaden their knowledge. This indicates a need for the manual to include both the basics of evaluation and information on how to deal with more complex issues.*

Turning to the competence needs of consultants, a manual may be particularly useful for "sector experts" with limited evaluation experience. Practical guidelines for how to deal with issues beyond the consultant's main field of expertise,

[37] A completely different issue from needs and demand is what *responsibility* Sida has to enhance the competence of different actors in an evaluation. Naturally, it must cater for the competence of its own staff. It could also be argued that Sida, as part of its overall commitments, has a responsibility to strengthen evaluation competence of its co-operating partners. Generally speaking, the competence of consultants is not the responsibility of Sida; Sida should procure the evaluation services of competent consultants. However, if Sida expects consultants to know its policies, routines, and preferred methods (what was referred to as "Sida-specific competence") it must communicate and explain these to consultants. One possibility could be to make it a standard procedure to refer to Sida's evaluation manual in terms of reference. Furthermore, Sida may have an *interest* in enhancing the competence of some categories of consultants to broaden its resource base.

including the different crosscutting issues, are needed. Considering the large number of international evaluators and Swedish evaluators not working regularly for Sida, a manual may also be used to communicate and explain what Sida will be expecting from them. Apart from improving evaluations, this may be seen as a way of broadening the resource base.

Evaluation competence among co-operation partners is likely to vary. This means that a manual for this group should assume no, or limited, previous knowledge of evaluation, implying a basic level.

3.4 Procedures and routines

Apart from providing competence, one of the possible functions of a manual was said to be to provide systems, procedures and routines.

What guides work today? *Using the evaluation tool* describes evaluation activities as "a very decentralised process, left to the initiative and professional competence of programme officers (...). In the absence of routines, policies and procedures at their level of work, that is, within the department, most of the work seems to be of an arbitrary and haphazard nature" (p. 16).

The same report noted that the Sida officers interviewed were aware of, but had not consulted, the policy. "None of the respondents suggest that the policy has led them to do other things than they otherwise would have done, nor does anybody claim that they have developed or changed their "normal" way of conducting the evaluation process." (p. 44). It also found that "The actual procedures appear to have developed out of practice, and to some extent the programme officers do as they always have done"(p.13).

No comprehensive research on what guides evaluation has been done for this study, but the answers received during interviews are basically in line with the reports quoted. What many state to be doing is try to recall how things were done the last time, look at the terms of reference of the last evaluation etc. Some ask colleagues for help, which is easier than looking through a manual when quick information is required. The standard outline for terms of reference on computer and in the evaluation manual is also used and appreciated. Some guidance is found in other Sida documents, such as the manual for procurement and LFA instructions.

The overall picture that emerges is one of a rather conservative and standardised way of dealing with evaluations. Evaluations are dealt with in a number of different ways. However, differences are not always based on the particular context of the evaluation, but rather on whom is managing it.

Our conclusion is that there is a need to provide routines and procedures for the different tasks in the evaluation. Clearly established norms and procedures are important for quality control and transparency. Chapter 3.2 indicates that a more stringent handling of e.g. procurement of consultants, reporting procedures and dissemination would be necessary live up to the policy. Procedures for how to estimate cost of an evaluation have been requested.[38]

However, as evaluations are not a standard product, there can be no standard routines. It may be necessary to propose alternative routines for different situations, e.g. different processes for elaborating terms of reference depending of whether the evaluation purpose is learning or control. Establishment of such alternative routines could actually contribute to a greater variety than exists today, and more important: that differences in how evaluations are handled are based on needs and not the individual Sida officer or department.

---

[38] "There are no rules and no explicit recommendation on how the costs (for an evaluation) are to be calculated, or how much an evaluations may cost", *Using the evaluation tool,* p. 15.

Implications for manual design: *The evaluation manual should outline different standards and routines for the evaluation process, e.g. consultation, dissemination of reports, costing etc. as indicated in previous chapters. However, it should avoid turning evaluations into standardised exercises. Alternative routines for different contexts and needs may therefore be suggested.*

3.5 Practical requirements

We have seen that a manual may enhance competence and help to improve routines. How, then, should the manual be *used in daily work?* To answer this question, we have looked at how the existing manual has been used.

3.5.1 How has the evaluation manual been used?

Most of the people interviewed for this study were aware of the existence of the evaluation manual but few had recently looked at it and even fewer could refer to what was said in it.[39] Answers received on *how* they had used the manual included:

- I read it when I first came to Sida;
- I have used it to find information on particular issues;
- I have used the checklists for terms of reference, quality assessment etc.;
- I have distributed it to counterparts;
- I read it to get ideas and then continued on my own;
- I have used it to check on vocabulary.

The manual seems to have been most important for the more recently employed.

All of the four Swedish NGOs contacted had the manual in their offices. At least one indicated that the manual had been used frequently. Other were more sceptical about its usefulness, indicating that the Sida manual put too much emphasis on *control* and *external* consultants to be relevant to evaluation activities of their own carried out as *internal* processes for the purpose of *learning*.

It is not known to what extent the manual may have been used in countries of co-operation and to little information to draw any conclusions regarding consultants' use of the manual.

Implications for manual design: *The different uses of the existing evaluation manual indicates that there are many different needs, including i) general introduction to evaluation work ii) quick reference on particular topics iii) ideas and inspiration and iv) communication with third parties.*

3.5.2 Accessibility

A condition for an evaluation manual to be used is also that it is accessible. What are the needs in this respect?

A manual must be accessible not only at head offices in Stockholm but also at the embassies or other places where evaluation activities are planned and take place. It should be possible to bring along on evaluation missions and be handed over to consultants and local stakeholders. It must be translated to the working languages of the countries where it is to be used.

Another consideration is *timing*. It may be assumed that the manual is not for everyday reading but should be available whenever an evaluation is coming up.

[39] *Using the evaluation tool* noted that a majority of the Sida officers interviewed had read the Sida evaluation manual (p. 22)

A third consideration is that people *learn* and absorb information in different ways. For example, here are "readers" and those who prefer getting things visualised. Therefore, there is the need to present the content of a manual in a variety of different ways (see the discussion on format). Furthermore, experience has usually shown that people learn most effectively if the learning is linked to their own particular problems and tasks. Hence, there is a need to link the presentation to concrete problems and tasks faced by the users of the manual.

A manual will have to compete with a large quantity of other documents for the readers' attention. The challenge for people nowadays is usually not to find information but to select the information relevant to them. Persons interviewed express a certain fatigue over all the new manuals and guidelines they are faced with, and make it clear that the time available for reading them is limited. Care must therefore be taken to ensure that the manual is made attractive and perceived as something that that can actually help them in their work.

A frequent request is that the manual be kept short. However, some find the current manual of a reasonable length and point to the risk of it becoming uninteresting if too brief. It should be recalled that a manual should not necessarily be read from the beginning to the end. Therefore, keeping it short is probably less important than ensuring that it is given a structure and layout which make it easier for the reader to find the information required.

People at Sida have been presented with a number of new manuals and guidelines over the last couple of years. The evaluation manual is one document which is relatively well known and established within Sida. It may therefore be subject to discussion whether the coming evaluation manual should be presented as a "new" document (in order to create attention) or merely as an "up-date" (to signal continuity).

Some, but not all, search for information on the computer. In time, and particularly following the introduction of the Intranet, this can be expected to become increasingly common. However, even though most persons interviewed are basically in favour of a computer-based manual, they do not see it as an alternative to a printed one.

Implications for manual design: *The manual should have a format making it accessible wherever evaluation problems may arise. A digital manual on the Internet may be seen as a complement to the printed one, not as a substitute. It does not have to be used continuously, only when evaluations are coming up. It should consider that different people learn in different ways. It must be attractive and perceived by the reader as something that can actually benefit his or her work. Sida may choose either to present the new manual as a brand new document or as an up-date of the old one.*

3.6 Strengths and weaknesses of the existing manual

What can be said about the existing manual in the light of the needs and demands outlined in this study?

In terms of *functions,* the current manual can be described as rather "all-round". It meets various needs and has been used in various ways by Sida, NGOs, consultants and possibly also by co-operating partners. It provides a mix of explanatory text and checklists and contains information useful both to evaluators and to evaluation managers. However, a manual specially *targeted* for evaluation managers or evaluators, or some specific use (e.g. teaching material, practical handbook when carrying out evaluations) would probably have looked different. Generally speaking, the manual appears to have been appreciated by those who have read it. The most common point of criticism is that it is not practical enough.

In terms of *content,* the current manual touches on all the major topics for which this report has indicated a need or demand. However, the *weight* given to different topics and what is *said* about them must be looked over. Inevitably, the manual does not reflect the changes that have taken place in Sida since 1994. For example, the manual deals with "Methods for Integration of *Gender Aspects* in Evaluation" but the current objective of *gender equality* is not reflected. Poverty, democracy and human rights, underlined in Sida's current programmes are only mentioned briefly. The manual discusses the hierarchy of objectives but only makes quick reference to LFA, which nowadays is Sida's preferred tool for project analysis and planning.

As the manual makes limited reference to particular types of support, the merger of the different Swedish development co-operation agencies with the "new" Sida has not greatly affected the relevance of the contents. However, organisational and administrative changes, such as the creation of UTV, the yearly evaluation planning, "management response" and the use of country analyses, ought to be reflected in a new manual.

On the whole, Sida's old evaluation manual is in line with the existing evaluation policy. However, it must be revised to correspond more closely to what is said on e.g. the objectives of evaluation, objectivity and impartiality, evaluation criteria, participation and quality requirements. Terminology must also be reviewed.

The *structure* of the manual is rather conventional, starting by defining evaluation and then going through the evaluation process step-by-step, assessments to be made and different evaluation approaches and data collection methods. It is built around different questions and topic more than tasks and processes. There is a certain repetition, e.g. evaluation criteria are treated both in the introductory chapter and in a separate chapter. In comparison to some other manuals, the Swedish ones tend to use more words to explain the same things.

The fact that many people interviewed at least knew where they had kept this five-year-old manual may have to do with it being produced as a book. It is about the same length as other evaluation manuals reviewed for this assignment. It uses a number of different fonts, underlinings, bullet-points, numbers and shadings to highlight the structure of the text. However, a more consistent way of doing so could have facilitated reading and information search considerably.

Summarising: *The manual corresponds to a broad range of needs and demands and appears to have had a variety of uses. On the other hand, it is not tailored for any specific target group or use. It includes the most relevant topics for evaluation work but it must be revised to reflect the changes that have taken place since 1994 and existing needs to improve evaluation activities. The soft cover book format has been suitable but layout may be improved to underline the structure and make it easier to find information.*

4 Proposals for the design of a new Sida manual

Considering needs and demands identified in this report, how should the new Sida manual look?

4.1 Function

The analysis shows that the new evaluation manual can, and ought to, have several functions. Ultimately, it should contribute to better evaluations and complement the evaluation policy, *Sida at Work* and other central policy documents.

More specifically, the manual should:

- enhance the *competence* of persons dealing with evaluations by providing relevant knowledge;
- provide practical *instruments and tools* for evaluation work;
- provide *systems and routines;*
- promote *reflection and innovation.*

In practical terms, there are a variety of needs that a manual should fulfil. Someone with limited previous knowledge of evaluation work should be able to use it to gain an *introduction* to evaluation work. Persons requiring information on some particular issue should be able to use it as a *reference book.* It should also serve as an *"idea book"* for reflection and innovation. Finally, it ought to be used to *communicate* Sida's policy and procedures to third parties. It is not believed that these different functions are in contradiction with one another.

A key issue for Sida to take a stand on is what the *intended readers* of the manual should be. The terms of reference for this assignment underline the evaluation manager perspective but it also mention the possible use of the manual outside of Sida. Three things may be noted.

First, there is a *demand* within Sida[40] for a manual directed primarily to the Sida staff and management, or possibly to evaluation managers in Sweden and abroad. The main arguments presented for this is that Sida has a responsibility to support its own staff, not consultants or other external parties. Furthermore, a manual directed to Sida staff and management could still be made available other readers so that transparency is assured.

Second, this report notes that all participants in evaluations – evaluation managers, evaluators as well as local stakeholders – could find a manual useful and that this could improve evaluations. Dialogue and co-operation, as called for in Sida's evaluation policy, require all participants in an evaluation to have similar access to information and knowledge. A manual directed to a broader audience could also be used to achieve a higher degree of stakeholder participation (when suitable) and to broaden Sida's resource base of consultants. Finally, it is not always possible to state who will participate in what ways in an evaluation. This indicates a *need* to produce a manual that can be used both within and outside of Sida.

Third, the knowledge required by the different target groups is largely the same. We are talking about a basic orientation of the main principles to guide the evaluation work and of the tasks involved in an evaluation. Hence, there would be no greater differences between a manual aimed at Sida staff and one aimed at a broader target group.

[40] The demand of other groups have not been analysed.

Ultimately, the choice of target group is a question for Sida to decide on. For now, it is assumed that the manual will be directed to all parties that actively participate in the Sida evaluations.

Another question to discuss and decide upon is what the official status of the manual should be.

The life-span of the manual will largely depend on how long Sida's evaluation policy remains unchanged. Unless there are any unforeseen events, a reasonable ambition would be that the manual could be used for, say, five years.

4.2 Content

The content, scope and limitations of the manual should be based on the functions described above. As the manual should give non-specialists dealing with different types of evaluations a general introduction of evaluation work, a *broad scope* and a relatively *basic* level is recommended. The manual should address all of the major tasks in relation to:

- the evaluation management process;
- the process of conducting an evaluation;
- the use of evaluations

It should deal with the central topics related to these tasks, tentatively outlined in Annex 6.

What exactly to be *said* on the different topics requires some discussion and consultation, particularly topics with a strong policy dimension, e.g. roles and responsibilities and analysis of the crosscutting issues.

The manual should provide both *practical instructions* and the underlying *theory and principles* as both were noted to be important for evaluation work.

It should not limit itself to some particular forms of evaluation but rather emphasise that there are many *different types of evaluations* (in terms of evaluation object, purpose, methods etc.) to be handled in partly different ways.

A broad scope means that not all topics can be dealt with in depth. It is suggested that *focus* be placed on:

- The *use* of evaluations, learning, control, the use of different stakeholders etc.
- *Roles and responsibilities* in different evaluation contexts;
- Sida's *development objectives*, including the various *crosscutting issues.*

This is because needs for an improvement of evaluation activities in these three areas were seen to be particularly great, added to which, these are important aspects of Sida's evaluation policy.

It is proposed that the manual be delimited to include topics likely to arise in *any* evaluation (see above). The question is then how to deal with topics related to specific evaluation problems, determined by e.g. sector, type of support etc. We will return to this in relation to format.

In relation to the *project cycle,* it is recommended that the manual focus on the *evaluation phase* and the *use of evaluations* only. Reference will have to be made, for example, to baselines, monitoring, the phasing of projects, final reports etc., but it should not go into detail on these issues. The reason for this is to avoid overlapping with the new guidelines for "project cycle management". Furthermore, the reader is more likely to look in project cycle document rather than in the evaluation manual when establishing objectives, monitoring projects etc.

The evaluation manual should be limited so that it *supplements* other Sida documents. The manual may e.g. deal with special problems in relation to the *procurement of evaluation services,* but it should not deal with formal procurement regulations and tendering procedures as these are described in other guidelines. Duplication is unnecessary and may also lead to confusion.

In relation to *structure,* it was noted to be an advantage for the manual to reflect the evaluation process and thus link into the different tasks of the reader. However, it was also seen that some important issues could not be dealt with in this way. Therefore, a combination of a "process" and a "thematic" structure is proposed. Furthermore, a structure should be chosen that make it easy for the reader to find information on a particular subject. At the same time it should be logical for someone who reads the manual from the first to the last page.

Annex 7 presents a *tentative* structure for the new evaluation manual, outlining the headings and some of the main topics to be included under each heading.

It starts with *"Main guiding principles",* including the definition of evaluation, Sida's evaluation policy and quality criteria. The reason for starting with main guiding principles is that these should permeate all evaluation work and the continued discussion in the manual.

This is followed by a chapter on the *"Use of evaluations",* including the learning and control purposes, who evaluations are for, management response etc. The reason for bringing up evaluation use early in the manual, is that it is this that should guide all evaluation efforts (beside "main guiding principles").

The third chapter is *"Main components of the evaluation",* including different evaluation objects, LFA and evaluation criteria. It should give the reader the necessary understanding of the main elements of an evaluation, the assessments to be made and also point at differences between different types of evaluations.

We then turn to the actual evaluation process. The proposed Chapter 4 goes through the different tasks faced by someone *"Conducting an evaluation",* such as selecting an evaluation approach, the use of methods for data collection and analysis, formulating recommendations and report writing. The chapter ought to be relevant both for evaluators and evaluation managers needing an orientation of how evaluations are carried out.

It is followed by chapter on *"Managing evaluations".* The reason for having this relatively late in the manual is that a clear notion of *what* is to be managed is required before there can be a discussion on *how to* manage it. The chapter follows the different steps in the evaluation management process, including the decision to evaluate, preparation of the terms of reference, dialogue with evaluators and dissemination of the final report. The chapter makes reference to what has been stated earlier in the manual, e.g. guiding principles, different types of evaluations and how evaluations are carried out.

Having identified the different tasks in an evaluation the manual discusses *"Roles and responsibilities in evaluation".* Alternative forms of participation are discussed with consideration to different purposes of evaluations and key principles such as objectivity and partnership.

Poverty reduction, gender equality, environment, democracy and human rights will be treated in the report like the overall development goals they are. As these goals should be addressed in all evaluations, and as an urgent need for practical guidance was noted, it is proposed that special guidelines for assessment of the respective goals be attached as appendixes. As most Sida interventions have an institution-building component, guidelines for assessments to be made in this area are also proposed (in appendix). Other appendixes include a glossary, explaining Sida's use of evaluation terminology and a bibliography where the reader may search for additional information, both general literature and Sida documents such as the guidelines for environmental impact assessment.

4.3 Format

Considering the function and the content of the manual, it is proposed that it be published as a soft cover book, similar to the existing evaluation manual. 75-100 pages would appear adequate for the content suggested above. A book can be kept in the bookshelf until the day it is to be used, and it is not lost among day-to-day working papers in A 4 format. It can also be brought along wherever it is to be used, and handed over to potential readers.

With regard to layout, reference is made to chapter 2. Apart from being attractive, the layout should facilitate the search for information and the reader's understanding of the text. As the manual should be read by people with different backgrounds and different ways of absorbing the information it is recommended that the report use both text and illustrations to forward its message. The need to give the reader both a deeper understanding of evaluation work and practical guidelines call for a mix of explaining texts and checklists. The Norwegian manual may serve as an example in this respect. Professional expertise should be used to assure a suitable layout.

It was noted that many things cannot be included in an evaluation manual, e.g. specific evaluation problems, information relevant only to specific target groups and sector specific issues. Furthermore, there is development work going on and new issues coming up continuously. Depending on the resources and ambitions of Sida, supplementing the evaluation manual with separate "Issue papers" may be considered. These could include more in-depth discussions on particular topics, be tailored for different readers and be supplemented continuously to reflect the development work that takes place within the field of evaluation. The issue of monitoring and assessment of different types of support could e.g. be addressed in this way. One paper could also provide a summary of the evaluation manual to be distributed to persons who are not likely to read the entire document. With regular publishing, the Issue papers could form a "series" similar to the USAID Evaluation Tips.

The report showed that the printed manual can be supplemented, though not substituted, by a digital version. A first step for Sida should be to consider how to make use of Internet and Intranet in general, not only in connection with the manual. Simply publishing different documents (including an evaluation manual or issue papers) can be done without any greater efforts. Computer based training programmes, reference books, linkages to word documents etc. can be prepared the mid-term perspective with the help of external expertise.

4.4 Note on continued work

Although not explicitly mentioned in the terms of reference, a few comments can be made regarding the future continued preparation of a new evaluation manual.

Preparation of a manual must be co-ordinated with the establishment of the new evaluation policy. As the manual should reflect the policy, the policy naturally comes first.

Furthermore, work is now going on to prepare new guidelines for LFA, to concretise the action programmes and to prepare an "Absolute Sida". It is important that the preparation of an evaluation manual is co-ordinated with this work, that there are clear links between the different documents and that the terminology is harmonised.

The purpose of this report has been to provide a first "sketch" for a future evaluation manual, not to go into depth on each of the topics to be included in the manual. Many of the issues to be dealt with in a manual are – or should be – subject to discussion. Development work is taking place in a number of areas. The preparation of the manual may be linked to this development work.

The dialogue with intended users of the manual, initiated in connection with this study, seem to have been appreciated. In order, both to improve the quality of the new manual and to increase the acceptance for it, it is suggested that this dialogue is continued and extended to parties not yet contacted, particularly in co-operation countries.

It is sometimes stated that it is during the evaluation process that learning takes place. Similarly, much learning regarding evaluation work could take place in the process of preparing an evaluation manual. It can, for example, be used to engage people in a discussion and exchange of information. It could also be a way of activating the "evaluation co-ordinators".

Sida could, for example, decide on the overall content and structure of the manual but let working groups prepare different chapters or determine what should be stated on specific topics. Of course, all of this will require time and staff resources. If Sida is in a hurry to produce a manual and wants to use its resources in other ways, it must find other, simpler, ways of preparing the manual.

Annex 1: Terms of Reference

Sida/UTV
Stefan Molund
99-01-19

Terms of Reference for a Design Study for a Sida Evaluation Manual

1 Background

In consultation with other departments within Sida, the Department for Evaluation and Internal Audit (UTV) has decided to produce a manual for evaluation that can be used by Sida staff and others in the planning, implementation and assessment of evaluations of development interventions supported by Sida. This manual will replace *Evaluation Manual for SIDA* (1994), a handbook that was written by Elisabeth Lewin for SIDA before the merger of this agency with other Swedish development agencies in 1995.

A demand for an updated evaluation manual emerged with the formation of the new agency. As long as Sida was still in the process of establishing a groundwork of governing policies, action plans, and operational guidelines, however, the project of producing such a manual was deemed premature. Within UTV there was also a feeling that not enough was known about the practices of evaluation existing within the agency and that an investigation of existing ways of doing evaluation should be made before the project of writing new evaluation guidelines was embarked upon.

Today, the situation is different. At the agency level, the work of defining policies and procedures has advanced a great deal, and within UTV the gaps in knowledge regarding evaluation practice within Sida have to some extent been filled. As a result, UTV has decided that the production of the new manual should no longer be postponed.

2 Purpose

The production of the new evaluation manual for Sida will begin with a background study analysing the need for new guidelines and sketching in outline a manual that might satisfy that need. More formally, the study has the following main purposes:

- to produce an empirically grounded analysis of the needs and demand for a new evaluation manual among Sida staff and other stakeholders in development co-operation activities supported by Sida;
- to propose an outline of a manual that would satisfy these needs and this demand.

The outline should cover the format of the manual as well as its contents. It should be clearly and explicitly motivated in terms of the prior analysis of needs and demand.

3 The Assignment: Scope and Limitations

The assignment can only be described in fairly general terms. What it involves in analytical and practical terms will to some extent have to be decided step by step as the work proceeds. Thus, it requires a close and open dialogue between UTV and the consultant. The following considerations and requirements are guidelines for that dialogue:

1. A Sida manual for evaluation must obviously be consistent with *Sida's Evaluation Policy*; this document may even be incorporated into the manual as an appendix. But the consultant should also consider other important Sida policies and operational instructions, not least *How Sida Works*, Sida's operational primer published in 1997, and Sida's instruction for the use of Logical Framework Analysis. It is part of the consultant's task to consider what policies and instructions are most relevant for the manual and how they should be taken into account. The consultant should also consider the implications of the fact that Sida has recently adopted a *system of management response* to evaluations.

2. Wherever possible, the manual should be consistent with *common international understandings of concepts and principles of evaluation*. It is especially important that Sida's guidelines for evaluation do not deviate unnecessarily from the framework for evaluation of development co-operation promulgated by the DAC Working Party on Aid Evaluation of which Sida is a member. In the manual, explicit reference should be made to the principles for evaluation laid down in the *DAC Manual for Development Assistance*.

3. We must not lose sight of the fact that insofar as Sida is concerned, evaluation is a practical tool to be used along with other such tools in a results-oriented management of programmes and projects. Thus, although the manual will be mainly devoted to problems of evaluation in a more narrow sense, its point of departure must be a wider one of *project and programme cycle management*. The following are basic questions to be addressed: How should provisions for the *evaluability* of projects and programmes be made? To what extent are evaluators dependent on information produced through *monitoring*? In what situations are evaluations required; when can information from monitoring be exclusively relied upon?

4. The primary users of Sida's new evaluation manual will be Sida staff, especially programme officers managing project evaluations. It may also be used by project administrators in the partner countries, project consultants, and NGOs. This means that it should give a great deal of importance to problems of *evaluation management*, including those concerned with negotiations with partners and stakeholders, writing of terms of reference, recruitment of evaluators, reporting, dissemination and utilisation of evaluation results.

What are the issues that a Sida programme officer must or should deal with at the different stages of an evaluation process? How should a manual be constructed in order to be practically useful for programme officers and others involved in the management of Sida evaluations? These are two basic questions for the consultant to answer.

5. A question of great importance concerns *the roles of stakeholders in evaluations* of activities supported by Sida. In *Sida's Evaluation Policy* it is clearly stated that in evaluations initiated by Sida external stakeholders should be invited to participate to the fullest extent feasible, and the policy also says that such evaluations should be guided by a concern for objectivity, impartiality and transparency. The manual should provide clear guidelines and advice for the practical implementation of these directives.

6. In a similar vein, the manual should deal with the topic of *joint evaluations*, the pros and cons of such evaluations as well as the practical problems of organising and managing them. The manual should not be written on the assumption that Sida always is or always should be in charge of the management of evaluations of activities supported by Sida. With regard to questions of division of labour and management, it should rather encourage open-mindedness and flexibility. The consultant should discuss how questions concerning alternative stakeholder arrangements might be

incorporated into the manual. In this context, the notion of *participatory evaluations* should be addressed.

7. While the primary target group for the manual consists of Sida staff, it might also be *useful to stakeholders outside Sida*. For hired evaluators as well as for partner organisations it may serve as an introduction to Sida's approach to evaluation, and, depending on how it is designed, some stakeholders may even be able to use it for their own evaluation purposes. The question of how interests and needs of other stakeholders could be accommodated in the manual is one that the consultant should consider.

8. The manual should not give the misleading impression that every evaluation is a *project evaluation*. Since project evaluations are far more common than any other type of evaluation of Sida-supported activities, however, such evaluations are bound to figure prominently in the manual. Still, it would probably be useful if the manual also presented some other *types of evaluation*. The consultant should discuss whether and to what extent it would be both possible and useful to include discussions of *non-project evaluations* (evaluations of multi-project programmes, evaluations of budget support to countries, sectors, or organisations, evaluations of policies and strategies, thematic evaluations, etc.).

9. Another preliminary question concerns the extent to which the manual may usefully deal with *evaluation problems that are specific to one or several sectors of development co-operation*, such as education, research co-operation, democracy and human rights, or agricultural development. Should all differences between sectors or types of sectors be ignored? If not, which are the ones that should be discussed?

10. As shown by studies of evaluation practice within Sida, the evaluation of Sida projects and programmes in the perspective of the so-called cross-cutting issues of poverty reduction, gender and the environment is not systematically pursued. One reason for this seems to be that the programme officers managing evaluations do not know how to deal with these issues. The consultant should consider how this problem should be dealt with in the planned manual.

11. There is also an important question concerning the level of knowledge and pre-understanding that should be attributed to the intended reader. Do we expect the reader to be someone who already knows a great deal about evaluation, or is he or she not very knowledgeable in this field? This should obviously be discussed at the levels of topics and subtopics as well as with regard to the manual as a whole. As a rule, it is probably better too assume too little knowledge than too much.

12. The consultant should consider what would be a workable format for the manual. In doing this the consultant should keep in mind that the manual should be available on Sida's internal computer network and through Internet as well as in print. One idea is that it should be designed in a way similar to that of a help-system of a computer program. The consultant should consider this and other possibilities for organising and packaging information.

13. With regard to design, the basic problem is that of striking an optimal balance between the requirements of coverage and detail, on the one hand, and the requirement that the manual should be as easily and rapidly digestible as possible, on the other. The consultant is asked to suggest a workable solution to this difficult problem. Since there will not be room for everything, the consultant should be clear about priorities.

14. In many cases Sida's evaluation manual will not provide the reader with all the information and knowledge that he or she needs. By explaining how other sources of information can be

accessed and utilised, however, the manual can be of some use even in such cases. This is a point for the consultant to keep in mind while thinking about possible contents for the manual. The consultant might also think about the design of an effective system of referral.

15. For how long do we expect the new manual to last? What is its intended life-span? These are difficult questions to answer. Yet we certainly do not want it to be outdated very soon. One way of making it last is to exclude information of passing validity or importance, such as information concerning Sida's changing administrative practices. As information of the latter sort may be functionally very important, however, we should consider the possibility of putting it in appendices or sections that can be easily updated.

4 Methodology

As described above, the assignment consists of two parts, one an analysis of existing needs and demand for evaluation guidelines in Sida and among other stakeholders in development activities supported by Sida, the other the design of an outline for an evaluation manual that might satisfy identified needs and demands.

For the first part of the study there is an important database consisting of studies from UTV and others of evaluation practices in Swedish development agencies (primarily *Sida Studies in Evaluation* 96/1, 96/2, 96/4, 97/1, 97/2, 98/1). These studies should be used by the consultant to identify areas or aspects of evaluation that might be given special attention in a manual. For further information on topics that may need special emphasis the consultant should interview UTV staff and representatives of other departments and functions within Sida. Interviewees should be identified in consultation with UTV. It would probably also be useful for the consultant to talk to a handful of evaluators who have previously been engaged to evaluate Sida projects. Another possible source of information consists of studies of evaluation practices in development agencies outside Sweden.

In discussions with Sida personnel the consultant should try to elicit information about how and to what extent Elisabeth Lewin's *Evaluation Manual for SIDA* has been used. A careful analysis of the strengths and weaknesses of this manual should be a major point of departure for the design of the new manual. In large part, the writing of the new manual may be regarded as a work of rewriting and updating this manual.

In connection with the second part of the study, the production of an outline for a new evaluation manual, the consultant should also scan evaluation manuals produced by other agencies. The selection of reading materials should be made in consultation with UTV.

5 Time Plan and Reporting

Starting in January 1999 and ending with the delivery of a final report in mid-April 1999 the assignment covers seven man-weeks in total. A first draft report containing a summary of deskwork and consultations with stakeholders as well as a closely argued presentation of the proposed outline should be given to UTV not later than March 19, a revised draft should be presented not later than March 26, and the final report should be delivered April 22 at the latest.

The report shall be written in English. The contents should be presented in such a way that it can be used in further discussions about the planned manual. It should include a brief summary of not more than 2 pages.

Annex 2: List of References

Sida's policies and guidelines

Carlsson, J., Rylander, B., *Råd och anvisningar för tillämpning av LFA metoden under Projektcykeln*, Second Draft for Comments, Sida, 1999
Sida, Department for Policy and Legal Services, *Guidelines for the Application of LFA in Project Cycle Management*, 1996

Sida, *Sida at Work, Sida's methods for development co-operation,* 1997

Sida, *Sida Looks Forward, Sida's programme for global development*

Sida, *Minskad fattigdom, Sidas program för att bekämpa fattigdomen*

Sida, *Rättvisa och fred, Sidas program för fred, demokrati och mänskliga rättigheter*

Sida, *Kvinnor och män, Sidas program för jämställdhet*

Sida, *Omsorg om miljön; Sidas program för hållbar utveckling*

Sida, Department for Evaluation and Internal Audit, Sida's Evaluation Policy, 1995

Sida, Department for Evaluation and Internal Audit, Sida's Evaluation Policy, Draft 1997

Sida, Department for Natural Resources and the Environment, *Riktlinjer för miljökonsekvensbedömningar i utvecklingssamarbetet*, 1998

Studies on evaluation

Alberts, T., Andersson, J., *The Environment and Sida´s Evaluations*, Sida Studies in Evaluation 96/4, 1996

Carlsson, J., Eriksson-Baaz, M., Fallenius, A-M., Lövgren, E., *Are Evaluations Useful? Cases from Swedish Development Co-operation,* Draft for Comments, 1998

Carlsson, J., Eriksson-Baaz, M., *Two Years After: A review of the introduction of the LFA method in Sida,* Sida 1998

Carlsson, J., Eriksson-Baaz, M., *Two Years After: A review of the introduction of the LFA method in Sida,* Summary for Sida Evaluation Newsletter, 1998

Carlsson, J., Jörnmark J., *Evaluation and Learning in BITS,* SASDA, 1994

Carlsson, J. et. al. *Using the Evaluation Tool, A survey of conventional wisdom and common practice at Sida*, Sida Studies in Evaluation 97/1, 1997

Forss, K., *Några aspekter på kvaliteten hos utvärderingar av svenskt biståndssamarbete, Genomgång av databas upprättad vid Sekretariatet för Analys and utvecklingssamarbete,* SASDA, Working Paper Nr. 40, 1994

Lindahl, Claes, *The Management of Disaster Relief Evaluations, Lessons from a Sida evaluation of the complex emergency in Cambodia,* Sida Studies in Evaluation 98/1, 1998

Molund, S., Schill, G., *En kort granskning av: Tre utvärderingar av svenskt stöd för arbetsmarknadsreformer i Polen, Ryssland och Lettland,* Sida UTV, 1998

Peck, L., *Evaluating Gender Equality, Policy and Practice,* Sida UTV, 1999

Schill, G., *Granskning av resultatanalyserna i Sidas landstrategiarbete,* Sida UTV, 1996

Schill, G., *Uppföljande studie av Sidas resultatanalyser,* Sida UTV, 1998

Stokke, Olav, *A Perspective on SAREC's Evaluation System: Policy and Performance,* Working Paper Nr. 10, SASDA,1994

Tobisson, E., Vylder de, S., *Poverty Reduction and Gender Equality: An assessment of Sida´s Country Reports and Evaluations in 1995-96*, Sida Studies in Evaluation 97/2, 1997

Evaluation manuals of various development co-operation agencies

European Commission, *Evaluating EU Expenditure Programmes: Ex post and Intermediate evaluation, First Edition*, January 1997

Inter-American Development Bank, Evaluation Office (EVO), *Evaluation: A Management Tool for Improving Project Performance,* Washington, 1997

Lewin, E., *Evaluation Manual for Sida,* SIDA, 1994

Ministry for Foreign Affairs, Finland, *Guidelines for Programme Design, Monitoring and Evaluation,* Helsinki, 1997

Royal Ministry of Foreign Affairs, Norway, *Evaluation of Development Assistance, Handbook for Evaluators and Managers,* Oslo, 1993

UNDP, Office of Evaluation and Strategic Planning, *Results-oriented Monitoring and Evaluation -A Handbook for Programme Managers*, 1997

Center for Development Information and Evaluation, U.S. Agency for International Development, Office of Evaluation, *Procedures Guidebook For Conducting CDIE Evaluations,* 1992

General literature on evaluation methodology and management

Flint, M., Austin, C., *Learning Lessons, A Review of ODA Evaluation Impact,* Department for International Development, London, 1997

Forss, K. and Carlsson, J. *The Quest for Quality - Or Can Evaluation Findings Be Trusted?,* Evaluation, Vol. 3 (4): 481-501, 1997

Freudenthal, S., Narrow, J., *Baseline Study Handbook, Focus on the Field,* Sida, 1993

International Labour Office, Design, *Monitoring and evaluation of technical cooperation programmes and projects, a training manual,* Geneva, 1996

JEFF, *The Joint Evaluation of Emergency Assistance to Rwanda: A Review of Follow-up and Impact One Year After Publication,* Joint Evaluation Follow-Up Monitoring and Facilitation Network (JEFF) to the Joint Evaluation Steering Committee for its Meeting on 14 February 1997 in Copenhagen, 1997

Leviton, L. et. al. *Teaching Evaluation Using Evaluability Assessment,* Evaluation, volume 4(4): 389-409, 1998

OECD, DAC Expert Group on Aid and Evaluation, *Evaluation of Programs Promoting Participatory Development and Good Governance, Synthesis Report,* Paris, 1998

OECD, *DAC Principles for Effective Aid, Development Assistance Manual,* Paris, 1992

OECD, DAC Working Party on Aid Evaluation, *Good Practice in Evaluating Humanitarian Aid, Room Document No. 3,* 1998

OECD, DAC Working Party on Aid Evaluation, *Effective Practices in Conducting Joint Multi-donor Evaluation*, Room Document No.6, Draft, 1999

OECD, *Methods and Procedures in Aid Evaluation,* Paris, 1986

Poate, D., *Measuring & Managing Results: Lessons for Development Cooperation*, Sida and UNDP, 1997

Riddell, et. al., *Searching for Impact and Methods: NGO Evaluation Synthesis Study,* OECD/DAC Expert Group on Evaluation and Ministry for Foreign Affairs of Finland, 1997

Rudqvist, A., Woodford-Berger, P., *Evaluation and Participation - Some Lessons,* A report prepared for the DAC Expert Group on Aid Evaluation, Sida, 1996

Kommittén för Analys av utvecklingsamarbete, *Analys och utvärdering av bistånd,* SOU 1994:102

The Joint Committee on Standards for Educational Evaluation, James R. Sanders, Chair, *The Program Evaluation Standards, 2nd edition, How to Assess Evaluations of Educational Programs,* Sage Publications, 1994

UNDP Office of Evaluation and Strategic Planning, *Who Are the Questionmakers?: A Participatory Evaluation Handbook*, 1997

USAID, *Selecting Data Collection Methods and Preparing Contractor Scopes of Work*, 1985
Wholey, J., *Assessing the feasibility and Likely Usefulness of Evaluation,* Handbook of Practical Programme Evaluation, Chapter 2, Undated copy
Wholey, J., *Evaluation: Promise and Performance,* The Urban Institute, Undated
World Bank Operations Evaluation Department, *Assessing Development Effectiveness, Evaluation in the World Bank and the International Finance Corporation,* 1994

Sida Evaluations

Biggs, F., *Centre for Legal Education and Aid Networks in Kenya*, Department for Africa, Sida Evaluation 98/33
Cameron, I., Flodman, K., Lindblom, A-K., Åhlström, E., *Evaluation of the Swedish NGO Foundation for Human Rights*, Department for Democracy and Social Development, Sida Evaluation 98/37
Forss, K., Bjern, L., Milano, B., *The Impact of Sida Financed International Training Programmes -A case study of the Philippines*, Department for Infrastructure and Economic Cooperation, Sida Evaluation 98/31
ITAD, Ltd, Logframe-related evaluation of Sida's democracy and human rights support: Phase 1 Evaluability assessment, March 1999.
Lind, A., Free to Speak Up, Sida and UNDP, 1996
Rylander, L., *Programme for Total Quality Management in Russia*, Department for Central and Eastern Europe, Sida Evaluation 98/23
Wang, C., Poluha, E., Thorvaldsson, J., Mutono, S., *The Water and Environmental Sanitation Programme , WES in Uganda*, Department for Natural Resources and the Environment, Sida Evaluation 98/28
Wohlgemuth, L., Ewald, J., Yates, B., *Three Human Rights Organisations Based in Banjul, Africa*, Department for Democracy and Social Development, Sida Evaluation 98/35

Other publications

Johnston A, *Ways of measuring an organisation´s or institution´s condition and development*, Draft 2, Sida, 1999
Sida, Beslut GD 158/98, *Införande av systematiska ställningstaganden och åtgärder avseende utvärderingar och revisioner i Sidas verksamhet*, 1998
Sida, Department for Evaluation and Internal Audit, *Sidas utvärderingsverksamhet samt återföring och användning av utvärderingsresultat,* Promemoria, 1998-05-20
Sida, Department for Democracy and Social Development, *Metodik för utvärdering och lärande i utvecklingsamarbete inom demokrati och mänskliga rättigheter*, 1999
Sida, Department for Evaluation and Internal Audit, *Direktiv för uppföljning av 1998 års utvärderingsplan samt planering av Sidas utvärderingsverksamhet för 1999*, 1998
Sida, Department for Evaluation and Internal Audit, *Sida Evaluations FY 1999: Departmental Evaluation Plans*, 1999
Sida, Department for Evaluation and Internal Audit, *Summary of UTV´s Activity plan: Appendix to Sida´s Evaluation Plan FY 1999*, 1999
Sida, Department for Evaluations and Internal Audit, *Underlag för val av utvärderingsprojekt för budgetåret 1999*, 1999
Sida, Department for Evaluations and Internal Audit, *Årsredovisning för utvärderingsverksamheten vid Sida : Verksamhetsåret 1998*, 1999
Sida, Department for Natural Resources and the Environment, Environment Policy Division, *Environmental Analysis- Layout and Content*, 1997
Sida, Department for Natural Resources and the Environment, *Miljöanalys -några synpunkter till hjälp i bl.a. landstrategiarbetet*, 1997

Annex 3: List of persons interviewed

Persons at Sida recently in charge of evaluations:

Eriksson, Ina	RELA
Elfving, Rikard	DESO/UND
Siwertz, Annika	INEC/IKTS
Borrman, Therese	SEKA/EO
Johansson, Jan	ÖST/ERO
Folkesson, Rolf	DESO/DESA
Runnquist, Jan	NATUR/AFRIKA
Lind, Agneta	DESO/UND
Schildt, Lena	AFRA/AFRA ÖST/VÄST
Gerremo, Inge	NATUR/MILJÖ
Hellsten, Elisabet	RELA

Other persons at Sida:

Molund, Stefan	UTV
Berlin, Anders	UTV
Granlund, Anders	SAREC/UFO-P
Lövgren, Eva	UTV
Fallenius, Ann-Marie	UTV
Segnestam, Mats	NATUR/MILJÖ
Johnston, Anton	DESO, DESA
Ekman, Bengt	GD-staben
Gustafsson, Ingemar	DESO AL/DESO
Lagus, Maria	Sida ÖST/EVA
Löfström-Berg, Ingrid	POL
Rask, Greta	PEO/EOL
Andersson, Mikael	PEO/EOL
Gunilla Cederqvist	INFO
Dahlgren, Stefan	POL

Persons outside Sida:

Carlsson, Jerker	Andante
Lundgren, Carl-Erik	Swedish Mission Council
Neuman, Leif	Diakonia
Wiklund, Berit	Afrikagrupperna
Johansson, Göran	SCC/Utan Gränser
Moberg, Carin	Ord och vetande
de Vylder, Stefan	AB SVEK
Rylander, Lars	SPM
Karlstedt, Cecilia	ÅF International

Annex 4: The lists of contents of selected evaluation manuals

Evaluation Guidelines

Danida, 1999

Evaluating EU Expenditure Programmes
Ex post and Intermediate evaluation
First edition, January 1997
European Commission

TABLE OF CONTENTS

Guidelines for Programme Designs, Monitoring and Evaluation
Ministry for Foreign Affairs, Finland
Department for International Development Cooperation, 1997

TABLE OF CONTENTS

Evaluation: A Management Tool for Improving Project Performance
Inter-American Development Bank
Evaluation Office (EVO) 1997

TABLE OF CONTENTS

Evaluation of Development Assistance
Handbook for Evaluators and Managers
Royal Ministry of Foreign Affairs, Norway, 1993

TABLE OF CONTENTS

Evaluation manual for SIDA
Elisabeth Lewin, 1994
Evaluation Unit, Planning Secretariat, SIDA, 1994

Results-oriented Monitoring and Evaluation
A Handbook for Programme Managers
UNDP, 1997
CONTENTS

PART II. DEVELOPING SELECTED INSTRUMENTS FOR MONITORING AND EVALUATION

Office of Evaluation
Procedures Guidebook for Conducting CDIE Evaluations
Center for Development Information and Evaluation U.S Agency for International Development, 1992

TABLE OF CONTENTS (does not include TIPS published on the internet)

Appendices

A: The Process for Setting the Administrator's Evaluation Agenda

1. Targeting Appropriate Topics for the Agenda
2. Setting the Agenda
3. Defining the Role of CDIE Staff
4. Selecting the Topics

B: Evaluation Focus And Study Approaches

1. Introduction
2. Evaluation Focus
   - 2.1 At the Intervention Level
   - 2.2 At the Program and Policy Levels
   - 2.3 At the Operational and Management Systems Levels
3. Evaluation Approaches
   - 3.1 Evaluation Desk Studies
   - 3.2 Limited Field-Based Evaluations
   - 3.3 Field Assessments
   - 3.4 Longer Term Intensive Evaluations

C. External Review Panels

1. Introduction
2. Stages of the Evaluation Process and Use of External Reviews
3. Review Questions
   - 3.1 Evaluation Design
   - 3.2 Draft Reports
4. Responsibility for Reviews
5. Composition of Review Panels and Selection Criteria
6. Review Process
7. Action on the Recommendations

D: CDIE Team Planning Workshops on Field-based Program Assessments

1. Introduction
2. The Need for Team Planning Workshops
3. Workshop Objectives
4. Intended Workshop Outcomes
5. Workshop Facilitation and Participants
6. Workshop Duration
7. Suggested Agenda

E: CDIE Evaluations Publications: Types of Reports and the Review Process

1. Background
2. Audience
3. Evaluation Publications both for Technical and Experts and Generalists
4. The Review Process
5. Editing Tracks

F: Guidelines for Preparing Short Periodic Reports and Oral Presentations

1. CDIE Evaluation Highlights
2. CDIE Evaluation Bulletin for Senior Managers
3. Guidelines for Effective Oral Briefings

G: New Principles for CDIE Evaluations Under the Administrator's Evaluation Initiative

1. Background
2. New Emphasis on Greater Rigor
3. New Emphasis on Objectivity
4. More Comprehensive Coverage of A.I.D.'S Program Portfolio
5. New Emphasis on Ensuring Management Relevance and Use of CDIE Evaluations

Annex 5: Competence requirements - Tasks

The "project cycle" process

(How to...)

Project planning and preparation

... make a problem analysis
... analyse the target group
... establish base lines
... formulate objectives
... establish indicators

Implementation

... prepare monitoring information
... report and document

Evaluation

(See below)

Feed-back

... learn
... exercise control
... make decisions

Evaluation management process

(How to...)

... decide to evaluate
... define the purpose of the evaluation
... select the object of evaluation
... take stock of available information
... identify additional information needs and key questions
... determine criteria for assessment
... define scope and limitations
... define necessary competence
... indicate suitable methods
... determine time required for the assessment
... assure financing of the evaluation
... formulate terms of reference
... assign and procure evaluators
... inform stakeholders about the evaluation
... provide relevant documentation
... maintain a dialogue with relevant parties during work
... process of the draft report
... assure quality of the report
... publish and disseminate the final report
... take action on the evaluation

Evaluation process

(How to...)

... plan the evaluation
... carry out consultation and dialogue
... collect data
... analyse data and draw conclusions
... formulate recommendations
... write the report
...communicate results

Annex 6: Tentative list of issues for an evaluation manual

What is "evaluation"?

Different definitions
Sida's definition
What an evaluation is not
  Evaluation and auditing
  Evaluation and monitoring
  Evaluation and "studies"

Objectives/The use of evaluations

Accountability
  Who is accountable to whom and for what?
  Exercising control
Learning
  What can be learned from evaluations?
  Who is learning?
  How does learning take place?
  Learning for management purposes vs. broader purposes
Other ways of describing evaluation use
    Instrumental use
    Conceptual use
    Persuasive use
Evaluation as part of the dialogue with stakeholders
Management response

Different evaluation objects

Project, programmes, sector programme, country
Aid forms (e.g. grants, contract financed Technical Co-operation, balance of payment support)
Sectors (health, education, industry etc.)
Organisational evaluations
Thematic evaluations
Criteria for selection of evaluation objects
  Strategic importance (policy relevance, financing, innovation and replicability)
  Evaluability
  Usefulness
  Cost

Timing of evaluations

Mid-term, end or project, ex-post etc.
Implications for evaluability

LFA

The hierarchy of objectives
Making use of indicators
Reviewing assumptions

Assessment criteria

What is an evaluation criterion?
Sida's main evaluation criteria
- Relevance
- Goal fulfilment
- Impact
- Cost-effectiveness/efficiency
- Sustainability
- Replicability
- Lessons learned

Development objectives and crosscutting issues (poverty, environment, gender equality etc.)

Development factors

Policy support measures
Institutional aspects
Financial and economic aspects
Technological factors
Socio-cultural aspects

Evaluability

Availability of data (baselines, monitoring data etc.)
Existence of objectives and indicators
Timing of the evaluation
Resources for the evaluation
The attribution problem
The counterfactual problem
Fungibility and additionality

Evaluation methods and design

General approaches
- Before and after project
- With and without project
- Cost-benefit analysis
- Surveys and case studies
- Quantitative vs. qualitative analysis
- Formal methods vs. informal (low cost) methods
- Participatory evaluation techniques

Data collection
- Desk studies
- Direct measurement
- Interviews (individual and group)
- Questionnaires
- Primary and secondary sources

Data analysis methods
- Statistical and non-statistical analysis
- The use of models

Specific evaluation problems

- Development goals and crosscutting issues
  - Poverty reduction
  - Environment
  - Gender equality
  - Democracy and human rights
- Sector specific issues (e.g.)
  - Evaluation of relief and humanitarian aid
  - Evaluation of research programme
  - Evaluation of health, education projects etc.
- Other
  - Evaluation of institution building
  - Evaluation of budget support etc.

Roles and responsibilities

- Main parties in the evaluation (Sida, project owners, local stakeholders, consultants, NGOs other donors etc.)
- Whom are evaluations for?
- Different forms of participation during the evaluation
- Factors determining suitable forms of participation
- Joint evaluations with other donors
- NGO evaluations

Procurement

- Alternative ways of procurement
- Sida regulations
- Where to find consultants

Sida's organisation for evaluation

- Sida's evaluation policy
- The roles of UTV and the departments
- Evaluation planning procedures
- Use of the Plus system for evaluations

Evaluation quality

- Quality in the process and of evaluation reports
- Quality criteria
  - Utility
  - Feasibility
  - Propriety (including evaluation ethics)
  - Accuracy
- Other guiding principles (meaning and how to assure them)
  - Credibility
  - Commitment
  - Objectivity
  - Impartiality
  - Partnership

Evaluation terminology

Annex 7: The new evaluation manual – a first sketch

Preface

The preface explains the *purpose* of the manual, *why* it has been written as well as *when, where* and *how* it is supposed to be *used.* The *intended readers* of the manual are indicated. It is further stated what *status* the document has and how it is related to Sida's *evaluation policy* and *other Sida documents.*

1. Main guiding principles

Making reference to Sida's evaluation policy, this chapter presents the fundamental guiding principles for evaluation activities. It explains the very *concept of evaluation,* what it is and what it is not (monitoring, audit, "studies"). Reference is made to the project cycle, in which evaluation is positioned. The chapter points at *quality criteria,* such as utility, feasibility, propriety and accuracy, or other quality criteria to be determined by Sida. It also highlights central issues such as *objectivity, impartiality, transparency, partnership, credibility* and *commitment.* The difference between statements of facts and value judgements is discussed. Reference is also made to Sida's *action programmes* and overall development objectives.

2. Using evaluations

Chapter 2 should give the reader knowledge of how evaluations are *used* to improve development co-operation. One objective of evaluation is *learning.* The manual brings up questions such as *who* are to learn from evaluations, *what* can be learned and *how* does learning take place? It also discusses how learning from evaluations can be used and fed back into *decision making* and *policy formulation.* Reference is made to Sida's planning systems as well as the procedures for *"management response".* Another objective of evaluation is *accountability and control.* The manual explains what this actually means and how evaluations can be used to exercise control. Evaluation use may also be described along other dimensions such as *instrumental, conceptual* and *persuasive* use. The use of evaluation as part of the *dialogue* between Sida and co-operation partners is discussed. Together with the main guiding principles, it is the *use* of evaluations that should *guide* the evaluation process described further ahead in the manual.

3. Main components of the evaluation

Having discussed general principles and what purpose evaluations may serve, this chapter focuses on the evaluation in itself and gives the reader a general orientation of different types of evaluations. It notes that there may be a number of different *evaluation objects* (such as projects, sector programmes, organisations and different forms of co-operation) and that *timing* may differ (e.g. mid-term, end of project and ex-post, evaluations) which has implications for how the evaluation should be designed and what information can be produced. Making reference to LFA, the chapter discusses the *hierarchy of objectives, indicators* and *assumptions.* It presents the different *evaluation criteria* (relevance, goal fulfilment, impact, sustainability, replicability etc.) and explains how and when they can be used. The chapter also discusses *evaluability,* from a theoretical as well as a practical viewpoint (existence of indicators, proper monitoring etc.). Common *evaluation problems* such as attribution and fungibility are also raised.

4. Conducting an evaluation

The chapter gives a practical orientation of *how* evaluations are conducted. It goes through the different steps of an evaluation, including initial *consultation*, organisation of the *team,* selection of evaluation *approach,* different methods for *data collection* and *analysis,* formulation of *recommendations,* report *writing* and *communication of results.* Having read this chapter, the reader should have an understanding of what it takes to conduct an evaluation, the suitability of different approaches and methods in different contexts and the implications, in terms of output, time requirements etc.

5. Managing an evaluation

Using a mix of explaining text and check-lists, this chapter goes through the different tasks of the *evaluation management process,* starting with the *initiation* of the evaluation, definition of evaluation *purpose,* selection of *evaluation object* (based on criteria such as strategic importance, evaluability, usefulness and cost) and *timing* of the evaluation. Reference is made to *Sida's system for evaluation planning.* The chapter goes on to the tasks of defining *information needs* and *key questions*, determining the *evaluation criteria* as well as the *scope* and *limitations of the evaluation,* identifying the necessary evaluation *competence* and determining *roles and responsibilities* of different stakeholders in the evaluation, all to be reflected in the *terms of reference.* Tasks in the preparation process also include assessing the *time* required for the evaluation (for evaluators, evaluation managers and other participants) and *budgeting,* two things that are closely inter-linked. This is followed by the tasks of *procuring evaluators* (in Sweden or abroad), *informing* concerned parties of the evaluation, *making data available* to evaluators, assuring necessary *co-ordination* and maintaining a *dialogue* with the concerned parties during the course of the evaluation. Then, there is the *processing of the draft report,* ensuring *quality* of the report, *publication* and *dissemination* of the final report and *taking action* on the evaluation. How these tasks are to be performed depends on the particular evaluation context and type of evaluation. Hence, the chapter will highlight the different *questions to ask* and provide *alternative suggestions* with consideration the specific evaluation context.

6. Roles and responsibilities in evaluations

The chapter notes that there are many *stakeholders* in an evaluation: the party commissioning the evaluation, project owners and other stakeholders, including the target group. Sometimes evaluations are carried out in co-operation with other donors. The *interests* and *perspectives* of different stakeholders may vary, as well as their participation in the different phases of an evaluation. Referring to main guiding principles (e.g. objectivity, credibility, ownership and partnership) and different types of evaluations (e.g. control and learning), the chapter discusses alternative and suitable *forms of participation* for different groups of stakeholders in the various tasks described above. Practical suggestions for how to achieve such participation are provided. Sida's internal organisation for evaluation activities is also described in this chapter.

Appendixes:

- Glossary, giving Sida's definition of central evaluation terminology (which should be in line with DAC terminology) but also explaining the actual meaning behind the words and recognising that other terminology may be encountered.
- Dictionary, with the most central evaluation terminology in Swedish, English, Spanish and Portuguese.
- Sida's evaluation policy
- Bibliography with suggested reading for anyone who needs more information in some area.
- Guidelines for the assessment of poverty reduction.
- Guidelines for the assessment of gender equality.
- Guidelines for the assessment of democracy and human rights.
- Guidelines for the assessment of environmental impact (in line with the existing ones).
- Guidelines for the assessment of institution building.

Sida Studies in Evaluation - 1995/96 - 1999

96/1	Evaluation and Participation - some lessons. Anders Rudqvist, Prudence Woodford-Berger Department for Evaluations and Internal Audit
96/2	Granskning av resultatanalyserna i Sidas landstrategiarbete. Göran Schill Department for Evaluations and Internal Audit
96/3	Developmental Relief? An Issues Paper and an Annotated Bibliography on Linking Relief and Development. Claes Lindahl Department for Evaluations and Internal Audit
96/4	The Environment and Sida's Evaluations. Tom Alberts, Jessica Andersson Department for Evaluations and Internal Audit
97/1	Using the Evaluation Tool. A survey of conventional wisdom and common practice at Sida. Jerker Carlsson, Kim Forss, Karin Metell, Lisa Segnestam, Tove Strömberg Department for Evaluations and Internal Audit
97/2	Poverty Reduction and Gender Equality. An Assessment of Sida's Country Reports and Evaluations in 1995-96. Eva Tobisson, Stefan de Vylder Secretariat for Policy and Corporate Development.
98/1	The Management of Disaster Relief Evaluations. Lessons from a Sida evaluation of the complex emergency in Cambodia. Claes Lindahl Department for Evaluation and Internal Audit
98/2	Uppföljande studie av Sidas resultatanalyser. Göran Schill Department for Evaluation and Internal Audit
98/3	Evaluating Gender Equality - Policy and Practice. An assessment of Sida's evaluations in 1997 - 1998. Lennart Peck Department for Evaluation and Internal Audit
99/1	Are Evaluations Useful? Cases from Swedish Development Co-operation. Jerker Carlsson, Maria Eriksson-Baaz, Ann Marie Fallenius, Eva Lövgren Department for Evaluation and Internal Audit

SWEDISH INTERNATIONAL DEVELOPMENT COOPERATION AGENCY
S-105 25 Stockholm, Sweden
Tel: +46 (0)8-698 50 00. Fax: +46 (0)8-20 88 64
Telegram: sida stockholm. Postgiro: 1 56 34–9
E-mail: info@sida.se. Homepage: http://www.sida.se