Sida Studies in Evaluation 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

Are Evaluations Useful?

Cases from Swedish Development Co-operation

Jerker Carlsson
Maria Eriksson - Baaz
Ann Marie Fallenius
Eva Lövgren

Sida Studies in Evaluation 99/1

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: Jerker Carlsson, Maria Eriksson - Baaz, Ann Marie Fallenius, Eva Lövgren.

Sida Studies in Evaluation 99/1
Commissioned by Sida, Department for Evaluation and Internal Audit.

Registration No.: UTV-1997-0025
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

Table of Contents

Executive Summary i

1. Purpose of the study 1

2. Studying USE - a point of departure. 3
2.1 Outline of this report 3

3 . Using the evaluation tool in Swedish aid - a summary of the full study 5
3.1 How and why are evaluations initiated? 5
3.2 How is the evaluation process managed? 6
3.3 How are the evaluation results used? 6
3.4 Is the quality of evaluation reports good? 8

4. Method 9
4.1 A stakeholder approach to analysing use of evaluations 9
4.2 Analysing the evaluation process 11
4.3 Case studies 12

5. Use and participation - an introduction to concepts 14
5.1 Types of use 14
5.2 Participation, partnership and ownership 15
5.2.1 Partnership, ownership and control - contradictions in Sida policy 18
5.3 Participation - a necessary condition for use? 18

6. The evaluation - useful for whom? 20
6.1 What is required to ensure utilisation 20
6.2 Are the basic premises for use fulfilled? 21
6. 3 Patterns of use as found in the case studies 26
6.3.1 No Use 26
6.3.2 Instrumental use 28
6.3.3 Conceptual use 29
6.3.4 Evaluations as means to communicate and legitimise interests 31
6.3.5 Ritual Use 34
6.4 Conclusion 36

7. Understanding the current pattern of use 39
7.1 Factors internal to the evaluation process 39
7.1.1 Participation, ownership and use 39
7.1.2 Focus on written rather than verbal communication 45
7.1.3 The role of the evaluator 46
7.2 Factors external to the evaluation process 47
7.2.1 The evaluation relationship 47
7.2.2 Donor agency and recipient organisational constraints 49

8. The link between the evaluation process and use - conclusions 51

9. Recommendations for improving use 53

Bibliography 55
Annex 1. Terms of reference 57
Annex 2. Interview guide 60
Annex 3. Summary of the 9 evaluations chosen as case studies 66

Executive Summary

This report forms part of the study "Using the Evaluation Tool", initiated by Sida's Department for Evaluation and Internal Audit. The purpose of the study is to analyse the evaluation process as it currently works in Swedish development co-operation. It attempts to broaden our knowledge of the way in which evaluations are initiated, produced and, finally, distributed and put to use.

This study is concerned with three questions:

* *How are evaluations used by various stakeholders in development co-operation and how do different stakeholders perceive their usefulness?*
* *How do different stakeholders participate in the evaluation processes and how can stakeholder participation affect utilisation*?
* *How can the usefulness of evaluations be improved and what are the factors determining utilisation?*

A sample of 9 evaluations/projects was selected for this study.[1]

It is no exaggeration to say that there is still some way to go before the typical evaluation process provides a good starting point for a broad utilisation of evaluations results. There is, across the board, a significant disregard of the critical factors which determine whether an evaluation will be relevant to all stakeholders. An evaluation continues to be a concern for a very limited proportion of all those who have an interest in a project and are affected by its outcome.

The evaluations analysed in this study have been used in various ways and for various purposes. It is discouraging to find that the general pattern has been that *the further away you are from the centre of the project in terms of power and involvement, the less information you receive on evaluation recommendations and findings.* This means that officers at Sida and, in most cases, officers occupying central positions within the major co-operation partner (i.e. the ministries) do get access to the evaluation report and can act. Stakeholders further away from the project centre very seldom have a chance of acting on the evaluation. There were two reasons for this: 1) they never received any feedback on the findings and recommendations from the evaluation; 2) they never had the power to act on the evaluation findings, even if they wanted to because they did not have access to the findings.

We found that those stakeholders who do use the evaluation rarely do so instrumentally. Other types of use are much more common - conceptual use, for example. The results of evaluations are supposed to trickle down into various decision-making processes and result in changes which are concerned more with strategies of doing things than with more immediate operational matters.

---

1 [1]*Evaluation of the Small Business Development Programme in Northern Namibia*. by Sahlin, Å.
2 *Rural Village Water Supply Programme - Botswana*. by Valdelin, J. et. al Sida Evaluation 96/10
3 *Botswana Road Safety Improvement Programme* by Davey, R. Sida Evaluation 96/43
4 *Water Supply System in Dodota*, Sida Evaluation 96/23 by Olsson, B., Judith Narrowe, Negatu Asfaw, Eneye Tefera, Amsalu Negussie
5 *Building Research Capacity in Ethiopia,* Sida Evaluation 96/9 by Thulstrup, E.W., M. Fekadu, A Negewo
6 *Support to the Education Sector in Ethiopia*, Sida Evaluation 96/27. by Valdelin,J, Michael Wort, Ingrid Christensson, Gudrun Cederblad
7 *Curriculum Development in Ethiopia*, Sida Evaluation 96/40 by Palme, M. Viggo Kilborn, Christopher Stroud and Oleg Povov
8 *HESAWA, Health through Sanitation and Water - Sida supported programme in Tanzania*, Sida Evaluation 97/12 by Smet, J., Kathleen Shordt, Pauline Ikumi, Patrick Nginya
9 *National Soil and Water Conservation Programme - Kenya*, Sida Evaluation 96/25 by Tiffen,M., Raymond Purcell, Francis Gichuki, Charles Gachene, John Gatheru

Using evaluations for legitimising decisions already taken is common, as is using evaluations for ritual flag waving.

The fact that other types of use than the conventional instrumental use are so dominant suggests that evaluations are not "pieces of objective evidence". Rather the evaluations seem to form part of an already ongoing process of dialogue between the stakeholders in development co-operation.

There are few examples of evaluations actually contributing something new in terms of knowledge. Neither the issues, the questions or the answers were new to the stakeholders. They had been discussed before by other means and perhaps in other ways. The evaluation repeated/reminded them of critical issues. The evaluation became an instrument, a tool, for stakeholders to use in their regular dialogue.

Evaluations are situated in the context of various stakeholders' interests. They are by nature political. There are always different stakeholders with different interests who try to influence the evaluation in the direction of their interest. From this it follows that evaluations cannot be seen as neutral or objective or as representing some outside neutral perspective. In addition to this, the evaluators themselves have values and assumptions of their own which they bring to the evaluation. By choosing a particular evaluator you also choose a particular perspective, a preference concerning how things should be done and even concerning the general drift of the conclusions.

The fact that evaluations are influenced by various stakeholders could be seen as problematic. Our conclusions, however, is that this can never be avoided. It is in the nature of all evaluations in all kinds of contexts. However, it becomes a problem when the evaluation arena is characterised by dominance arising out of dependency. Sida possesses considerable strength as donor in relation to most recipient governments and it does not hesitate to use the evaluation function to further its interests. It is not surprising that almost all of Sida's co-operation partners saw the evaluations as a way for Sida to convey its interests and concerns. The conclusion drawn from this study is that evaluations are useful to a very limited group of stakeholders. For a majority of stakeholders the evaluation process could just as well have been left undone.

The evaluations analysed in this report are quite far removed from the key Swedish policy concepts of "partnership" and "ownership". There is not much openness and clarity about the way in which interests make themselves felt in an evaluation. Mutual trust is not what immediately comes to mind. In terms of adherence to the present policy framework of Swedish aid, the evaluation process as we know it is in dire need of reform.

1. Purpose of the study

Evaluations have been an institutionalised part of the activities of aid agencies since the 1970´s. They are standard procedure in almost all bilateral and multilateral aid agencies. Numerous project, programme, sector and thematic evaluations are conducted each year, and quite a substantial amount of money is channelled into the evaluation function in the various aid organisations. Evaluations are widely assumed to serve an important function for the work of aid agencies. According to Sida, evaluations shall serve several purposes. One is *learning*. According to this objective, evaluations are supposed to serve as an "instrument for improving development assistance through feedback of experience from ongoing or completed activities". The evaluations are supposed to produce knowledge that can be used to make development efforts more relevant, effective and efficient. Another objective of evaluation activities, according to Sida, is *development of knowledge*. This objective is formulated in more general terms, such as contributing to a better understanding of fundamental processes and events. Evaluations shall "increase our knowledge about development assistance and its possibilities and limitations as an instrument of economic and social change" (Sida's Evaluation Policy 1995). Finally, Sida evaluations are also supposed to contribute towards *accountability* of aid by regularly controlling aid projects and programmes.

The question addressed in this study is to what extent evaluations live up to these expectations. Given the large funds allocated for evaluations and the essential purposes that the evaluations are supposed to serve, it is of central importance to study evaluation processes and in particular how evaluations are put to use. One very obvious prerequisite for evaluations to be capable of "making development efforts more relevant, effective and efficient" is that they are put to use in various ways. The present study is concerned with this question – *how are evaluations used by various stakeholders in development co-operation and how do different stakeholders perceive the usefulness of evaluations*?

This report forms part of the study "Using the Evaluation Tool", initiated by Sida's Department for Evaluation and Internal Audit. The purpose of the study is to analyse the evaluation process as it currently works in Swedish development co-operation. It attempts to broaden our knowledge of how evaluations are initiated, produced and, finally, distributed and put to use. The work has been divided into two phases. The focus of the first phase, conducted during 1996/97, was on Sida´s internal organisation and it only briefly touched upon the involvement of other stakeholders in the evaluation process. Utilisation appeared as a subject, although from the limited perspective of Sida headquarters. In the present study, which can be seen as the second part of the larger study, the focus is shifted towards the field and to the stakeholders who are actively involved in the planning, implementation of a particular project, as well as all those supposed to benefit from the aid intervention.

A study of evaluation processes and utilisation cannot only include Sida, which, by its own testimony, is not the only stakeholder in the evaluations of Sida-funded activities. As regards learning, it is pointed out that the organisations and individuals in the collaborating country who are directly concerned with the activities under review are also target groups for evaluations. Moreover, development co-operation should, according to Sida, be guided by the principle of partnership. This is also stated in Sida's evaluation policy: "Evaluation work should be transparent and be carried out in a spirit of co-operation". This is the background to another question in this study: how *do different stakeholders participate in the evaluation processes and how can stakeholder participation affect utilisation*?

Finally, as indicated above, the purpose of the study is not only to describe the evaluation processes and patterns of utilisation, but also to address the question of *what are the factors affecting utilisation and how can the usefulness of evaluations be improved.*

2. Studying USE - a point of departure

By the late 60´s there was a growing realisation that the visions of government based on rational decision-making supported by scientific truth might not be realised. Non-utilisation and under-utilisation of evaluation results emerged as critical problems in evaluation research. Patton (1986). Carol Weiss (1972: 10-11) made the point that:

> Evaluation research is meant for immediate and direct use in improving the quality of social programming. Yet a review of evaluation experience suggests that evaluation results have not exerted significant influence on program decisions.

The utilisation debate among evaluation researchers in the 1970´s and 80´s assumed that it should be possible to see a direct use of evaluation results through changes in programmes and projects. When this direct impact was not very visible, it was assumed that evaluations were not used at all. This concern suggests that there is a wide gap between the instrumental evaluation model's view of the world and the world as it actually functions.

To illustrate the causes of this perceived inefficacy, it is helpful to divide the evaluation system into three different parts, or subsystems: the production system, the distribution system and the receiving system. Weaknesses in the production system often mean that evaluation models and methods can be hard to understand and/or are inappropriate. Evaluations can also simply be examples of bad craftsmanship. Observed weaknesses in the distributive system often concern technical matters such as producing and delivering the evaluation in a timely way, presenting the results in a proper format and ensuring that they are distributed to the interested parties.

The role of the receiving system can be understood from many perspectives. Lack of knowledge and ability may be important for utilisation. The individual may lack the necessary knowledge to understand an evaluation. There may also be weaknesses in the organisational context of the individual that can impede use. The necessary organisational mechanisms for ensuring that evaluation results are acted on, for example somebody responsible for dealing with the evaluation and/or proper organisational channels may not be in place. Other influential factors can be lack of time and continuous changes in the priorities of the organisation.

The point made here is that there is far more to the evaluation process than the final report. In order to understand utilisation, we have to analyse the whole evaluation process, from initiation and distribution to receiving system. A focus on how the evaluation is produced, common in many textbooks and manuals, is too narrow a perspective for understanding the use of evaluations.

2.1 Outline of this report

This report is organised as follows. In the next section we present a summary of the whole study, phases 1 and 2. This is intended for the action-oriented reader, who may find it difficult to read through the whole report. Section 4 presents the method approach we have used for the study. Essentially this is a matter of introducing the stakeholder approach and the selection of case studies. The next section (5) introduces concepts that are central for the study. First, the concept of use, or utilisation and, secondly, the concepts of participation, partnership and ownership, and how they are related. Section 6 contains the empirical material, as well as the actual analysis. It starts with a presentation of what we consider to be basic premises for use. Conditions that have to be present in an evaluation process for stakeholders to regard the evaluation as useful. Then follows a presentation of the patterns of use as found in the case studies. Three questions will guide this

presentation. Who are the evaluations used by? For what purpose? And do stakeholders find the evaluations useful? Towards the end of the section, we return to the issue of whether the basic premises for use are present in the case studies. Section 7 is an attempt to understand the current pattern of use. The analysis is divided in two parts. The first deals with causative factors internal to the evaluation process - factors that largely are under the control of the evaluator and the sponsoring agency. The second part of the analysis looks at factors external to the evaluation process and beyond the control of the evaluator, e.g. constraints in the donor agency and recipient organisations, macro economic and/or political developments in the recipient country etc. Section 8 return to the terms of reference for the study and summarises the main findings. In Section 9, finally, we provide a set of recommendations for improving the use of evaluations.

3. Using the evaluation tool in Swedish aid – a summary of the full study

This section is for the benefit of the action-oriented reader who finds it difficult to sit down and peruse not only this report, but also the report from the first phase. This summary intends to capture the main findings from phase 1 and phase 2 of the study "Using the Evaluation Tool". Hopefully it will capture the reader's interest sufficiently for him/her to find it worthwhile reading the full versions.

One of the thematic working areas of Sida's Department for Evaluation and Internal Audit (UTV) concerns organisational change and learning. The present study forms part of that theme. Its purpose is to map the nature of the evaluation process and explore in-depth the usefulness of evaluation for Sida and its collaborating partners. There are four questions guiding the full study:

1. How and why are evaluations initiated?
2. How is the evaluation process managed, from the setting of a purpose, the decision to evaluate and the commissioning of a study?
3. How are the results from this process used?
4. Do evaluations meet acceptable standards of quality?

The full study consists of two parts. The first part, analysing the evaluation process from the perspective of Sida Headquarters in Stockholm, is based on a study of 30 evaluations. The second part shifts the focus to the field and all the various stakeholders' perceptions of the evaluation process, as well as the usefulness of the evaluation. This part is based on a sample of 9 cases, selected from the previous 30 evaluations identified for the first phase of the study. This part of the study is generally much more concerned, compared to the first one, with analysing how stakeholders use an evaluation. Let us then return to the original question and present our answers to each of them.

3.1 How and why are evaluations initiated?

What this study has clearly shown is that various stakeholders influence evaluations. Naturally, they try to influence the evaluation in the direction of their interest. From this it follows that evaluations cannot be seen as neutral or objective or as representing some outside neutral perspective. In addition to this, the evaluators themselves have values and assumptions of their own which they bring to the evaluation. By choosing a particular evaluator, you also choose a particular perspective, a preference for how things should be done and even the direction of the conclusions.

The fact that various stakeholders influence evaluations could be seen as problematic. Our conclusion is, however, that this can never be avoided. It is in the nature of all evaluations in every kind of context. Stakeholder influence becomes a problem when the evaluation arena is characterised by dominance arising out of dependency. Sida possesses considerable strength, as donor in relation to most recipient governments and it does not hesitate to use the evaluation function to further its interests. The problem is thus not so much the lack of objectivity - evaluations can never be totally neutral - as Sida's domination of the evaluation processes. It is not surprising that nearly all Sida's co-operation partners saw the evaluations as a way for Sida to convey its interests and concerns. The evaluations analysed here are not good examples of partnership. There is not much openness and clarity in the way interests make their way into an evaluation. Mutual trust is not what immediately comes to mind. In terms of adherence to the present policy framework of Swedish aid, the evaluation process as we know it is in dire need of reform.

Sida is a decentralised organisation in the sense that the organisation grants the programme officer extensive possibilities to influence decision-making. It is the programme officers who initiate an evaluation. The initiative to evaluate appears to rest on rather loose grounds. The most common reason for starting an evaluation is that it has been decided in the project document, or it was part of the agreement with other stakeholders that an evaluation should be carried out after a certain period of time There is nothing wrong in this *per se* as long as there is a carefully crafted purpose, understood and agreed by all major stakeholders. This, however, is rarely the case. Consultations are few and far between. Most programme officers prefer to have informal contacts with their closest colleagues.

The immediate and direct purpose of an evaluation is defined in the terms of reference (TOR). This is a document which appears to be largely copied from the evaluation manual. There is often no practical focus in the TOR. A good example of an unclear and impractical purpose was found in the evaluation of the rural village water programme in Botswana (Valdelin, 1996). The purpose was stated to be learning. However, there was no further elaboration on who was supposed to learn, how they should learn, and most importantly, what they should learn. Experience from this study has clearly indicated that an unclear purpose reduces the usefulness of an evaluation. In order for evaluations to be useful there must be clear ideas of why an evaluation is undertaken, which issues that should be treated and what kind of utilisation is expected. This is one of the central prerequisites for use. The fact of an evaluation of a programme being included in an agreement is no guarantee of it being used.

3.2 How is the evaluation process managed?

All evaluations were commissioned from external consultants. Evaluators were largely found with the assistance of colleagues, and most were well known to Sida from earlier evaluations. Some of them could be described as in-house consultants to Sida. Only rarely - in 17% of all 30 cases - were evaluators selected on the basis of competitive bidding. This practice illustrates a complex interaction between the evaluator and his client that cannot be described in simple client-server terms. Although the TOR of the evaluation, often formulated in neutral terms, stipulates what the client wants to be done, there are sometimes messages, intentions, views etc., which in a more informal way is communicated to the evaluator. The existence of such communication between the sponsor and initiator of the evaluation and the evaluator is particularly difficult in aid evaluation. It can easily disturb an already unbalanced relationship between the donor and various local stakeholders.

The evaluations are usually completed rapidly once they have been commissioned, often within a few months. But the process from the first evaluation initiative to the dissemination of the first report is a lengthier one. Evaluation management breaks down completely when it comes to the distribution of the evaluation report. Few of Sida's collaborating partners, particularly those in the recipient country, ever see the report they have contributed to.

3.3 How are the evaluation results used?

According to the programme officers we interviewed, the draft report is circulated to the same group of people that were consulted in connection with the initiative to start the evaluation, i.e. a fairly small group of people. In about 50% of the cases (15 out of 30 evaluations) the draft was sent to recipient country authorities for comment.

How did the programme officers assess utilisation? From their perspective utilisation was not a problem. According to the findings in the first study, the recommendations were often accepted by the programme officers and rarely led to any controversies. In about 3/4 of the cases the recommendations had also, according to the programme officers, led to concrete results. But sometimes they had led to actions other than those recommended. The second phase of the study gave us a different and a more complex picture of use.

First of all, many stakeholders are not in a position to use the evaluation findings and recommendations. The further away you are from the centre (in terms of decision making), the less information you get on evaluation recommendations and findings. This means that officers at Sida and, in most cases, the main co-operation partner (i.e. the ministries) have received the report while people working directly with implementation very seldom have access to evaluation findings.

The evaluations analysed were used in various ways and for various purposes. What experience from this study tells us is that it is rare for evaluations to be put primarily to instrumental use. Other types of use are much more common. Conceptual use is very frequent. The results from evaluations are supposed to trickle down into various decision-making processes and result in changes, which are more concerned with strategies of doing things, rather than more immediate operational concerns. Using evaluations for legitimising decisions already taken is common practice, and so is the use of evaluations as ritual flag-waving.

The fact that other types of use than the conventional instrumental use are so dominant suggests that evaluations are not "pieces of objective evidence". Rather the evaluations seem to form part of already ongoing processes of organisational encounter. This is evidenced by the fact that we found few examples where the evaluations actually contributed something new in terms of knowledge. Neither the issues nor the questions were new. They had been discussed before by other means and perhaps in other ways. Evaluation is just another possibility for the stakeholders to conduct their dialogue.

Are the case studies examples of "good" use? Usefulness can be defined in the following way: *an evaluation is useful when stakeholders find something in the evaluation which they can put to use according to their own interest and needs.* The conclusion drawn in this study is that evaluations are useful to a very limited group of stakeholders. For a majority of stakeholders the evaluation process could just as well have been left undone.

There are reasons for this strong statement. The dominance of Sida constitutes a major obstacle to use. It limits the possibility of establishing concrete ownership of the evaluation on the side of the co-operating partners. The evaluations are seen as a part of the agreement with Sida. From a legal point of view this is also what they are. In the agreement with its co-operating partners, Sida reserves the right to verify the use of funds as well as the progress of implementation, in order to ensure that the money is being properly used.

Serious application of a policy which emphasises ownership should enable the co-operation partner to "own" the utilisation and implementation of evaluation findings to a much greater extent. But in order to do this they must also have greater ownership of the whole process. Expecting the co-operation partner to implement and "own" the implementation process (following an evaluation) without granting them ownership of the rest of the evaluation process is unrealistic. Thus, there is a discrepancy between, on the one hand, the emphasis on "partnership" and "ownership" in the overall aid policy, and, on the other, the actual terms of the agreements which Sida concludes with its co-operating partners.

A creation of ownership requires a change in decision-making and influence over the evaluation process. "Consulting" the co-operation partner is not enough. Concrete and real use demands: (i) that the co-operating partner has the right to suggest evaluators; (ii) that Sida's partners have a real influence on the selection of issues to be treated; (iii) that the co-operation partner takes evaluation initiatives, (and of course that they have a say in whether an evaluation should be undertaken or not); (iv) that the process becomes more transparent and democratic.

3.4 Is the quality of evaluation reports good?

The quality of evaluations is an elusive subject, and it must be recognised that what one person regards as a good and reliable report can be given a totally different interpretation by another. When assessing the quality of evaluations in the first phase of the study, we have focused on the reports. But the report itself is only a manifestation of the whole process - which may contain quite different qualities.

An interesting picture emerges when we contrast the programme officers' opinion of quality with an independent, subjective assessment based on 50 variables. Most programme officers were pleased with the outcome of the evaluation process. They thought that the evaluators had arrived at reliable conclusions, and that the project was given fair and adequate treatment. The reports were found, however, to be relatively weak on new ideas and practical usefulness.

With a different kind of assessment, based on a standardised set of quality criteria, the picture changes dramatically. In the same sample of 30 evaluations, quality was found to be much less than desired. In particular, the reports were found to be methodologically weak.

Switching the perspective from the report to the process affords new insights as to what quality is all about. One of the most significant findings in the second phase of the study was that so many stakeholders were excluded from the evaluation process. They could not influence it and they were never in a position to use the results from the evaluation. The stakeholders most frequently excluded from the process were those working with day-to-day implementation and the beneficiaries. This pattern is partly an ethical problem, provided that the stance is taken that everybody with an interest in and/or put at risk by an evaluation should also be entitled to put their concerns and issues on the table. It is, however, also a problem of quality. The quality of evaluations will most probably be enhanced if these groups are allowed to participate in the evaluation process.

There are several reasons for this. Firstly, if these groups have a hand in identifying the issues to be addressed and in the collection of data, the evaluation results tend to be more realistic and more closely attuned to the project/programme context. Secondly, if they are involved in this way, the findings tend to be more accurate since participation removes the tendency to hide and exclude problematic issues. Thirdly, if these groups are included in the evaluation process, both in the way outlined above and in final workshops, the process of "trickling down" directives can be shortened, and the evaluation can become a learning process in itself.

4. Method

4.1 A stakeholder approach to analysing use of evaluations

The stakeholder approach has been central to this study. The notion of stakeholder is based on the idea that different people are put at risk by an evaluation. Since an evaluation can affect the future of a programme or project, different people who are involved in, or affected by, a programme have a stake in the evaluation. *Stakeholders are groups of people or individuals who are somehow affected by an evaluation*. They can therefore be expected to have concerns and issues related to it. We have employed a stakeholder approach since we analyse the issue of use by taking different stakeholders' perception of the use and usefulness of the evaluations as our point of departure.

In the evaluation arena there are a number of stakeholders, each with his/her own view of what is going on and how things should be done etc. It is rare to find a situation where groups and individuals do not hold competing and sometimes combative views on the appropriateness of the evaluation and on whose interests will be affected by the outcome. The promotion of these interests also largely determines to what purpose the findings and recommendations of the evaluation are used. The starting point for any analysis of use is therefore a recognition of the full range of stakeholders. Who are they and what interests are they bringing into the aid project? In the box below we have listed typical categories of stakeholders found in development co-operation.

Box. 4.1 Typical categories of stakeholders found in development co-operation.

* *Policy makers and decision-makers*
  Persons responsible for deciding whether a program is to be instituted, continued, discontinued, expanded or curtailed.
* *Program sponsors*
  Organisations that initiate and fund the program to be evaluated.
* *Evaluation sponsors*
  Organisations that initiate and fund the evaluation (sometimes they are identical with the program sponsors)
* *Target participants*
  persons, households, or other social units who participate in the program or receive the services being provided by the intervention being evaluated.
* *Program management*
  The group being responsible for overseeing and co-ordinating the intervention process.
* *Evaluators*
  Groups or individuals responsible for the design and/or conduct of the evaluation.
* *Program competitors*
  Organisations or groups who compete with the project for available resources.
* *Contextual stakeholders*
  Organisations or groups, individuals, or other social units, in the immediate environment of a project.

These are very broad categories and, as is suggested, sometimes it is not possible to separate them. The Sida identification of collaborating partners exemplifies this very well. The *project owner* is the party that requests support for its project and which is responsible for the planning and

implementation of the project. The project owner can be a government agency, a voluntary organisation, a company or a research organisation. The *co-operation partner* is the party which concludes an agreement with Sida on the project. The co-operating partner can also be the same as the project owner. The *implementing partner* is, in principle, the owner of the project.

Within each of these categories there are stakeholders with different interests. The target group, for example, most often consists of groups who can be affected in different ways, depending for example on gender, economic status and age. The project owner or the co-operating partner is also divided into different stakeholders, depending for example on duties and responsibilities in connection with planning and implementation. Interested parties can by definition consist of several different groups. Moreover, there are often differences of interest and perception within one actor. In the aid agency it is common to find such conflicts between headquarters and the officers at the Embassies.

It is important to point out that we cannot claim to have talked to all the possible stakeholders in the nine evaluation cases included in this study. For each evaluation we interviewed, on average, 15 - 25 stakeholder representatives. We tried to identify the major stakeholders who were involved in the evaluations by using the evaluation report and the "lists of persons met". This means that we interviewed stakeholders who participated in the evaluation in question. Thus there could very well be other groups who might be affected by the evaluation in some way (and who should therefore be seen as stakeholders). Only in a few cases have we talked to these "neglected stakeholders". A telling consequence of following the "lists of persons met" is that only in a few cases did we talk to representatives of the target groups. The target group is rarely involved in the evaluations. If they are, their role is restricted to provide information and they are rarely mentioned by name in the lists of persons met, so they are not easily identified.

In most cases we interviewed several people from each stakeholder group. Except for a very few cases, the interviews were made individually. This was considered important, given that the persons interviewed often occupy different positions in the institutional hierarchy, which can make it difficult for people in subordinate positions to express their views in a group. Moreover, since the people in the organisation often have different responsibilities, the assumption was that they could be expected to have different views on the usefulness of the evaluation. All persons interviewed by us were promised anonymity. We have, furthermore, refrained from making the full case studies available, in order to fully protect our respondents.

The interview procedures were semi-structured (a description of the interview procedure will be found in Annex 2). At the beginning of the study we constructed a questionnaire for the interviewees to fill in. We then continued the interview by discussing the answers with the participant and canvassing explanations. However, during the course of the study it became evident that not all people were especially keen on filling out the form. During the last part of the study the questionnaire was instead used as an interview guide and was presented to the respondent in order to give him or her an understanding of what issues we were interested in.

It should be stressed that we are not claiming to provide an "objective" account of the evaluation processes. The study is based upon the different stakeholders' perceptions of their participation and use of the evaluations. It has been quite interesting to see how one evaluation process can be described in so many, and often conflicting, ways. These often divergent and conflicting views, which unfortunately cannot be fully accounted for in the present report, should nevertheless be kept in mind.

4.2 Analysing the evaluation process

Our method for analysing use and usefulness takes its point of departure in a stakeholder approach, but it also builds on a perception of usefulness and what is required for an evaluation to be seen as useful by stakeholders. We shall discuss this in more detail later on (section 5). But as it also form part of our method, it should be briefly presented already here.

The starting point for determining whether an evaluation is useful or not, has little to do with a particular type of use. Rather, by useful we *refer to an evaluation process that has enabled stakeholders to find something useful in the evaluation which they can put to use according to their interest and needs.* An evaluation is considered useful by us if as many stakeholders as possible have had an opportunity to access the evaluation and determine whether it contains anything of interest to them. Thus, usefulness is very much determined by the nature of the evaluation process and its openness. To structure our analysis of the evaluation process we have used a set of *basic premises for use* identified by Patton (1986). They serve as the framework for our assessment of the evaluation process and the extent to which it facilitates utilisation. These basic premises are presented in some more detail in section 6.1.

4.3 Case studies

The starting point for the selection of case studies was the 30 evaluations that had formed the basis for the first phase of this study. A sample of 9 evaluations/projects was then selected for the second phase of the study. When we selected these 9 evaluations we followed a set of criteria, which we felt were important to facilitate a good study of utilisation, as well as allowing for a practical organisation of the study. Of course, not all the criteria can be applied equally, but to a large extent we have tried to take each of them into account in an equitable manner.

The following criteria were used:

* Sweden's most important co-operating countries in southern and eastern Africa should be represented.
* The spread between sectors should be as wide as possible.
* The evaluation should have been carried out some 12-18 months ago
* There should be a possibility of comparing participatory with non-participatory evaluation methods.
* There should be a possibility of comparing different modes of project implementation, i.e. where implementation was the responsibility of a local organisation, or of a foreign-based consultancy company.

The Swedish co-operating countries were represented by: Ethiopia, Kenya, Tanzania, Botswana and Namibia. They all belong, or have belonged, to the group of programme countries and have been variously successful as recipients of aid. Five sectors are represented in our sample. *Education*: primary as well as tertiary education in Ethiopia. *Health,* through water projects in Ethiopia, Tanzania and Botswana; *agriculture,* through the soil conservation project in Kenya; *industry,* through small-scale industry support in Namibia; and, finally, *infrastructure,* through the road safety project in Botswana.[1]

1 We would have liked to include a project working with democracy issues. The portfolio included one such project, in Ethiopia, but closer scrutiny told us that it would not be very useful for our purpose since it mainly involved study trips, procurement of office equipment etc.

It is important that the timing of the evaluation and the project should be right. The evaluation should not be too "old". It should have been carried out fairly recently, so that we can track down all the parties involved while their memories are still fresh. On the other hand, it is also important for a certain time to have elapsed, in order for the stakeholders to react to the recommendations and findings from the evaluation. Ideally, this means that the evaluation should have been carried out some 12-18 months ago. All the evaluations in our sample were produced and published during 1996 and 1997.

Since the evaluation is to address the question of how usefulness of evaluations can be improved and how use is related to participation, we searched for evaluations with different patterns of participation (which could serve as a basis for comparisons). In an evaluation made for UNDP (1989), Forss and Rebien have outlined different ways in which stakeholders can be involved in an evaluation process. They distinguish between six different activities in evaluation processes: 1) planning and formulation of evaluation questions, 2) deciding on which methods to use; 3) providing data; 4) collection of data; 5) interpretation of data; and 6) using the evaluation results (in decision making). Forss and Rebien perceive participation in an evaluation as a continuos line extending from low to high participation. A fully participatory evaluation would involve all stakeholders in all phases of an evaluation process.

Evaluations which have employed a fully participatory approach, in the way described by Forss and Rebien were (not surprisingly) not to be found. These kinds of evaluations are very rare in government based aid-organisations. The first part of the study "Using the Evaluation Tool" (which was based on 30 evaluations) provided some information regarding stakeholder involvement in Sida evaluations. According to this report *only about one third of the evaluations involved recipient government officials during the preparations of the terms of references*.

The cases chosen for this study are, at one level, examples of "best practice" at Sida. In all the evaluations chosen here the collaborating partners have (according to the survey material in the first part of the study) been involved in the evaluation, either in preparation or in commenting on the draft. It seemed quite pointless to choose evaluations of (and spend money on field trips to) projects or programmes where the stakeholders in the collaborating country had not participated at all (and probably, therefore, had no knowledge of the evaluation in question).[2]

Even though the pattern of participation in our sample of evaluations is fairly stable there are some variations (which became clear to us during the study) which have allowed for a comparative approach. There are, in particular, two cases which differ more considerably from the others in terms of pattern of participation:

1) *Building Research Capacity in Ethiopia.* This evaluation differs from the others in terms of the co-operating partner's influence on the evaluation process. In contrast to the other evaluations the co-operating partner had a great deal of influence over the process and expressed a strong ownership of the evaluation.

2 However, as became clear during the study, participation (as defined in the first study) can mean a lot of things. Its extent and quality vary a great deal. Participation can be limited to being forced to participate in an evaluation that you do not want. It can also extend towards exerting strong influence on the purpose and design of the evaluation, choosing the evaluation team etc. During the process of evaluation, differences between the cases in terms of participation and influence became clearer.

2) *HESAWA, Health through Sanitation and Water in Tanzania.* This evaluation differs from the others in another respect. It is not an example of the co-operating partner's influence over the evaluation process (in initiation and formulation TOR and choosing evaluators etc.) In contrast to the other evaluations, which were based on more conventional data-collection techniques, this evaluation was based on workshops where the participants could articulate their ideas and concerns in a more free way.

These cases have therefore provided particularly interesting and important information regarding how participation is related to use.

Application of these criteria to the portfolio of 30 cases resulted in the following sample (a presentation of the main features of each evaluation will be found in Annex 3):

* *Evaluation of the Small Business Development Programme in Northern Namibia*. by Sahlin, Å.
* *Rural Village Water Supply Programme - Botswana*. by Valdelin, J. et al. Sida Evaluation 96/10
* Botswana Road Safety Improvement Programme by Davey, R. Sida Evaluation 96/43
* *Water Supply System in Dodota - Ethiopia*, Sida Evaluation 96/23 by Olsson, B., Judith Narrowe, Negatu Asfaw, Eneye Tefera, Amsalu Negussie
* *Building Research Capacity in Ethiopia,* Sida Evaluation 96/9 by Thulstrup, E.W., M. Fekadu, A Negewo
* *Support to the Education Sector in Ethiopia*, Sida Evaluation 96/27 by Valdelin,J, Michael Wort, Ingrid Christensson, Gudrun Cederblad
* *Curriculum Development in Ethiopia*, Sida Evaluation 96/40 by Palme, M. Viggo Kilborn, Christopher Stroud and Oleg Povov
* *HESAWA, Health through Sanitation and Water - Sida supported programme in Tanzania*, Sida Evaluation 97/12 by Smet, J., Kathleen Shordt, Pauline Ikumi, Patrick Nginya
* *National Soil and Water Conservation Programme - Kenya*, Sida Evaluation 96/25 by Tiffen,M., Raymond Purcell, Francis Gichuki, Charles Gachene, John Gatheru

5 Use and participation - an introduction to concepts

> No doubt every evaluator has had moments of glorious dreams in which a grateful world receives with adulation the findings of his or her evaluation and puts the results immediately and directly to use.
> (Rossi and Freeman, 1993:444)

5.1 Types of use

What then do we understand by utilisation? How are evaluations being used? In the literature one can find different classifications of use. The traditional attempt to classify use is based on three types of use (Leviton & Hughes, 1979, 1981, Rich, 1977)

Direct or instrumental use refers to the documented and specific use of evaluation findings by decision-makers and other stakeholders. Evaluations are used to control, manage and learn from an activity. Unhappiness about the usefulness of evaluations usually refers to the inability to see a direct link between an evaluation and instrumental use. This pessimistic view should, however, be somewhat modified. There are studies which have established a direct impact of evaluations on changes in projects and programmes (Leviton & Brach, 1983; Chelimsky, 1991; Brown, 1982).

Conceptual use refers to the impact of evaluations on policies, programs and procedures, through sensitising persons and groups to current and emerging social problems. This will eventually lead to changes in the way they think of social issues and in a later stage, changes in policy and procedures.

Persuasive use refers to enlisting evaluation results in efforts either to support or to refute a political position already taken - or to defend or attack the status quo (Rossi & Freeman, 1993:443)

A more extensive typology would include also other types of usage, some of them with their roots in organisational politics. Utilisation then becomes linked to stakeholder interests and the need to prove a thesis, consolidate power, acquire ammunition for political battles, give an impression of action etc. Vedung (1995) identified a typology that is based on this kind of legitimisation perspective: instrumental, conceptual, legitimising, interactive, tactical and ritual use. His typology gives us many interesting angles for studying use, although it suffers from certain weaknesses. First and foremost, it mixes different classification criteria, on the one hand the *process* of utilisation, on the other hand the *purpose* of utilisation. The category "interactive use" is the best example. Secondly, it is difficult to distinguish between some types of use. "Legitimising" and "tactical" are in practice difficult to distinguish from each other, and can therefore conveniently be grouped together.

Thus, our typology consists of the following types, where the purpose of use is the classification criteria.

* Instrumental use
The engineering model where the operations are tried and tested and where the results from such a testing are fed back into the planning and implementation of the operations.

* Conceptual use
Results and conclusions from evaluations trickle down into political processes in the form of ideas, new concepts and new ways of structuring operations.

* Legitimising use
The evaluation legitimises decisions and positions that have already been taken on other grounds. Often it is not the results but the evaluation process as such which is important and used for different purposes.

* Ritual use
Initiating and implementing an evaluation symbolises desired qualities of the management or the organisation.

* No use
The evaluation is simply not being used. Stakeholders do not find anything of interest in the evaluation to develop their action or their thinking on the subject that the evaluation analyses. A distinct reason for non-use can also be that the evaluation is not available to stakeholders. The evaluation does not matter to them, because they do not know it.

Can this typology be linked to quality? Are some types of use considered "better" than others?

The point of departure here is that the quality of use can only be determined from the perspective of the individual actor and his/her preferences and interests. It cannot be argued, for example, that the ultimate objective of improving use is the supremacy of the instrumental model of use. When evaluations are not being used in the way prescribed by textbooks and organisation manuals, frustration over the effectiveness of the evaluation system develops easily. It is rarely we see such a direct application of evaluation information as this model assumes. This does not, however, mean that evaluation results are not used. By and large they probably are, but use is a much more complicated psychological and sociological process than standard instrumental models would suggest. Simon (1986:210-211) pointed out that decisions are normally taken on the basis of bounded rationality. It is not possible for an individual to know all possible actions open to him, assess them and act according to his best choice. This is especially unlikely to happen under conditions of uncertainty, (which are quite typical of development aid). Decisions have to be taken on the basis of a rather limited set of alternatives which are, furthermore, highly unstable. Thus, there is a distinction to be made between the real world and the actors' perceptions of it and reasoning about it.

When we talk about an evaluation as being useful we refer to an evaluation process that has enabled stakeholders to find something in the evaluation which they can put to use according to their interest and needs.

5.2 Participation, partnership and ownership

In recent years, Sida has put a lot of emphasis on such concepts as participation, partnership and ownership. The Agency has formulated goals and strategies trying to incorporate these concepts into active use in its operations. Supposedly this means that they should also influence the Agency's evaluation policy.

The question of how the concepts of participation, partnership and ownership should be defined is a highly contested issue both when it comes to development aid in general and evaluation research in particular. The purpose of this section is to try to elaborate how these concepts are conceptualised in this study and how they can be helpful for an analysis of use.

During recent years, participation has arisen as a central concept in evaluation research, both within the general discipline of evaluation and among researchers involved more specifically with the issue of aid evaluation (Rebien, 1994, 1996; Forss & Rebien 1993, 1994; Lawrence, 1989; Swantz,

1992). There is no single definition of what constitutes a participatory evaluation. On this subject there are probably as many viewpoints as there are practitioners.

The controversies surrounding the meaning of the concept can partly be seen as a result of participation in evaluation being viewed both as a means and as an end in itself. The means aspect is reflected in the increasing arguments that participation contributes to greater effectiveness. Thus in the context of evaluation the argument is that participation facilitates use of evaluation results by increased learning and communication by stronger commitment to implementing recommendations, and that it enhances the quality and broadens the perspective of information by raising questions that would not be raised otherwise (Rebien, 1996; Guba & Lincoln, 1989; Lawrence, 1989; Patton, 1986).

Participation is also perceived as an end in itself and not as a strategy for improving aid effectiveness. Participation is by some presented as an empowerment strategy whereby existing power relations can be questioned. It is also seen as a "stakeholders right" in that stakeholders are put at risk in an evaluation (since their lives can in different ways be affected by it) and therefore that stakeholders are entitled to have their concerns and issues included in the evaluation. Guba and Lincoln add an ethical dimension to the justification of participatory evaluation:

> ... all stakeholders put at risk in an evaluation have the right to place their claims, concerns and issues on the table for consideration, irrespective of the value system to which they adhere (Guba & Lincoln, 1989, p.11)

According to them, failure to involve different stakeholders is seen as unfair and discriminatory. The idea of development co-operation as a partnership promoted by Sida, can in part be seen to belong to the dimension of participation as an end or right in itself.

On the other hand, Sida also views participation as a means of achieving other goals. Participation is presented as an important integral element in the process of achieving the main policy objectives, especially poverty reduction, gender equality and protection of the environment. (Rudquist & Woodford-Berger 1995)

The Poverty Task Force at Sida has defined participation as the process whereby people - especially poor women and men - can influence political life, policy formulation, determination of development directions and the choice of development investments and interventions which affect them (ibid.).

This definition of participation is quite similar to the conceptualisation we have employed for this study. What is central in this definition is that it builds on a notion of *influence* rather than merely "taking part" in a process. As Rahnema has pointed out, the concept of participation can mean many different things.[3] The fact of somebody participating in a process does not necessarily mean that they have influence over the process. Neither does it necessarily mean that the person is participating as a spontaneous free exercise arising from a wish to participate in the process. Instead they can be led to perform actions which are inspired or directed by others and which are of limited interest to them. These aspects have become evident to us during the course of this study. While some stakeholders have been invited to participate they have not had the possibility of

3 Rahnema (1995) distinguishes between transitive and intransitive, moral, amoral, forced and free, manipulative and spontaneous participation. By distinguishing between moral and amoral he wishes to point out that the term participation is often associated with desirable goals and is given a positive connotation, while in fact the act of partaking "may apply to evil or malicious purposes". The distinction between forced and free partaking can be seen as an example of this. He points out that while participation is often seen as a free choice exercise, people are sometimes "dragged" into taking part, in the very name of participation, in operations which are of no interest to them.

influencing the process. Also, some people have been made to participate in evaluations which they have not wanted or been consulted about. In these cases stakeholder participation has not affected their use or perception of usefulness in a positive way. Instead, the important aspect of participation – if it is to influence use in a positive way – is that those who participate feel themselves to be capable of influencing the process. A meaningful conceptualisation of participation in an evaluation context would thus be that it is *a process whereby a stakeholder is taking part in, and has influence over, the evaluation process*.

This conceptualisation of participation - a feeling that one has influence over the evaluation process - is very much related to *ownership*. In a recent document Sida has defined the policy of ownership as follows:

> Sida's approach where our co-operation partners are concerned ... is that they themselves must be responsible for their projects: they must "own" their own development. ... In order to be able to say that a partner in co-operation is the owner of a project, the partner must have full rights to use the resources provided within the framework laid down in the project agreement. But this is not enough. The co-operation partner must also be prepared to assume full responsibility, participate actively in the work, and be ready to implement the project on its own initiative. (Sida at Work - Sida's methods for development co-operation, p. 17)

The concept of ownership is important in this study. Furthermore, we have found that a feeling of ownership of evaluations has an important bearing on their use. But in contrast to Sida's concept of ownership, quoted above, which only involves one owner - the co-operating partner - we would prefer a conceptualisation of *shared ownership*. According to a notion of shared ownership there is no one single owner of an evaluation. By ownership we refer to a situation *whereby a stakeholder has the feeling that an evaluation is not only done by and for the interests of another stakeholder, but is initiated, carried out and influenced by and for their own interests and concerns.* Participation and ownership thus become very much related, in that participation - to participate and being able to influence a process - becomes a prerequisite of ownership.[4]

Lastly, we would like to say something about the concept of partnership which, at least in Swedish aid policy, is related, although it is by no means clear in what way, to the concepts of participation and use. But, this being a policy strongly promoted by the Swedish government and Sida, it is of interest to see how it applies to the cases here. It should, however, be kept in mind that the evaluations examined here were conducted before Sida officially announced its partnership policies. Since partnership is not introduced as an analytic concept we will take our point of departure in the definition given by the Swedish Government. The recent Government Communication "Africa on the Move - Revitalising Swedish Policy towards Africa for the 21st Century", gives the following definition of partnership (Government Communication SKR 1997/98:122, pp. 99-100):

> Partnership should be seen as an attitude, in a form of co-operation that is based on a shared basis of values and mutual trust. Joint and clearly formulated objectives, conditions, obligations, roles and responsibilities are part of partnership. The aspiration should be to bring about increased equality and mutual respect in the relationship, in awareness of the fundamental inequality represented by the donor's upper hand in terms of resources. Partnership is also aimed at strengthening the weaker party.

4 The difference between the two concepts would seem to be that ownership implies, not only a feeling that one has influence, but a feeling of the evaluation being initiated and carried out for one's own concerns and interests.

5.2.1 Partnership, ownership and control - contradictions in Sida policy

None of the concepts participation, ownership and partnership can be easily transferred into the evaluation context. If Sida's own notion of ownership is to be used in relation to evaluation, this would imply that the co-operation partners "own" the evaluation, that they "assume full responsibility", "participate actively" in the evaluation and are ready to implement the evaluation on "[their] own initiative". Transferring this notion of ownership to the context of evaluations, however, is somewhat problematic, given that Sida also has, as the expression goes, a "responsi-bility to exercise control". (see Sida at Work - Sida's methods for development co-operation, pp. 18-20). Even though there are different ways of exercising this control, evaluations are inevitably a part of the control function. In Sida's evaluation handbook from 1993, under the heading "Evaluations - for whom?", we read that financiers have "a legitimate interest to evaluate activities that they support" and that "it is natural that the financiers, in agreement with the recipient countries, reserve the right to evaluate the programme they support". One of the (three) objectives in Sida's Evaluation Policy (1995, p. 1) is also concerned with the issue of control:

> Approached with a view to *control*, or accountability, evaluation is an instrument for examining the use and results of Swedish development assistance. In this perspective, evaluations primarily serve the Swedish public, the Government of Sweden, and Sida's Board of Directors.

As this text shows, the notion of ownership, according to Sida policy, is not easily transferable to the evaluation context. Nor are the concepts of participation or partnership. This points to an inherent *contradiction* in Sida policy. One the one hand Sida has the right to exercise control, on the other hand, development co-operation should be organised in the spirit of partnership. It has become clear in the course of this study that evaluations are to a great extent viewed as instruments of control. From this it follows that participating in an evaluation should *not* be considered as the collaborating partner's *right*. At the same time Sida constantly stresses the spirit of partnership and the importance of ownership on the part of the collaborating partner.

Clearly, then, judging the patterns of participation found in this study on the basis of Sida policy is extremely difficult, and we do not propose making the attempt. What we will do is describe the patterns of participation found in the study and present the results of our analysis of how participation can be seen to affect use and usefulness. Our own position should, however, be stated since it nevertheless has implications for the study - participation in evaluations should be seen as a right of the collaborating partners. Without this being acknowledged, the notions of ownership and partnership become meaningless.

5.3 Participation - a necessary condition for use?

In mainstream thinking on utilisation and participation there seems to be general agreement on the positive link between participation and utilisation. The more involved a stakeholder is in an evaluation, the more prepared he/she is to use the results.

In general, there are few studies of this relationship and we have found none within the field of development aid. Greene´s (1988) case studies of implementation of youth employment and day-care information and referral programmes in a community in New York give some interesting insights. Greene found that stakeholder participation at three levels - cognitive, affective and political - led to enhanced utilisation of evaluation results. Cognitive participation refers to an ongoing active discussion and processing of information related to key program issues toward some decision or action. Affective participation refers to the individual's feeling of worth and value. Finally, political participation refers to the situation when an evaluation process gives voice to the

less powerful so that they receive interest and attention from the more powerful actors on the arena. An evaluation process which facilitates a participatory bottom-up perspective contributes to utilisation in four ways: 1) greater understanding of the results; 2) greater acceptance or ownership of the results; 4) heightened perceptions of the results as valid, credible and persuasive; 4) greater sense of responsibility and obligation to follow through on the results. Greene's findings support the hypothesis that there is a positive link between extensive stakeholder participation and the utilisation of the results.

It is also useful to differentiate between participation and the different stages of the evaluation process. Involvement is more important for utilisation at some stages of the process than others. In the literature the question formulation phase is singled out as the most important. If you are in a position to influence the questions to be answered by the evaluation, you obviously come to regard that evaluation as explicitly recognising your information needs (Guba & Lincoln, 1989; Patton, 1986; Lawrence, 1989; Rebien, 1996).

Thus, substantial stakeholder participation can be expected to enhance use since evaluation is a process of learning in itself. Utilisation then becomes something that does not necessarily follow, or involve, a reading of the evaluation report - but something that occurs during the evaluation process itself. Further, the inclusion of stakeholders in the evaluation process can enhance stake-holders perception of the credibility of evaluation results. This is of particular relevance if aid evaluations are primarily seen as instruments of control to be used by the donor (Lawrence, 1989).

6. The evaluation - useful for whom?

6.1 What is required to ensure utilisation

Utilisation means different things to different people in different settings, and is an issue subject to negotiation between evaluators and intended users. Thus, there is no use which can be regarded as the ideal type which every evaluation should strive to achieve. The reader should again be reminded that *when we talk about an evaluation as being useful we refer to an evaluation process that has enabled stakeholders to find something useful in the evaluation which they can put to use according to their interest and needs.*

There is no formula that can guarantee utilisation. There are, however, some basic requirements that have to be in place for utilisation to take place at all. Patton (1986) identified 9 basic premises for utilisation. An analysis of which premises were fulfilled, wholly, partly or not at all, gives us a starting point for understanding the pattern of use that we found in each of the case studies. We have, however, made some changes in Patton's original set up. Mainly because we felt that some premises were overlapping and some could be made more specific. Our final list of basic premises is presented below:

1. Evaluations should be user-oriented

> The evaluation should be aimed at the interests and information needs of specific, identifiable people, not vague audiences. Identification of primary intended information users is the first step in a utilisation-focused evaluation.

2. *The intended evaluation users should be involved in making decisions about the evaluation.*

> People who are personally interested and involved in an evaluation are more likely to use evaluation findings. The best way to ensure utilisation is to target the personal concerns of various stakeholders in every stage of the evaluation.

3. The multiple and varied interests around an evaluation should be identified and considered.

> Stakeholders representing various constituencies should come together at the beginning of the evaluation to decide whose issues and questions will be given priority in the evaluation.

4. High-quality, rather than high-quantity, participation.

> The quantity of group interaction time is often inversely related to the quality of the process. Careful selection of stakeholders for active participation, is more important than trying to reach even those more remotely linked to the evaluation. Evaluators must be skilled group facilitators.

5. Stakeholders should be trained in the evaluation process

> By training stakeholders in evaluation methods and processes the prospects for effective utilisation is increased.

6. Awareness of how situational factors affect use

> The evaluator need to be sensitive to and aware of how various factors and conditions affect the potential for utilisation. A situational analysis should be undertaken together with the stakeholders early in the process.

7. The timing of the evaluation must be seen as right by stakeholders.

> Utilisation is facilitated if an evaluation is done at a time when it fits into the decision-making process of the stakeholders.

8. The recommendations must be seen as realistic and feasible

> The recommendations of an evaluation must be such that stakeholders have the resources and the possibility to implement them.

9. The evaluation must be made available

> An evaluation can hardly be used unless stakeholders have an opportunity to read or in other ways get access to the evaluation and its analysis.

There is no priority order among these premises. Their importance is relative to that of the individual case. Depending on the circumstances, the various premises may assume varying degrees of importance. The important thing for us, in the first place, is not to establish any priority order based on generalisation from the case studies. It is the configuration of premises which is important, since it will help us to understand the pattern of use in each case.

6.2 Are the basic premises for use fulfilled?

The case studies yielded a considerable wealth of information on how evaluations are actually being used by various stakeholders. The picture that emerged was quite different from that expected by many handbooks and textbooks on evaluation. Instrumental use, for example, turned out to be less frequent than anticipated, and non-use appeared as quite a common experience for many stakeholders. The utilisation pattern we saw suggests that few of the basic premises for use were fulfilled to the extent that one would expect. Table 6.1 below summarises the outcome for each case study.

To facilitate an analysis of this kind of qualitative data we have used a ranking order from premises not being fulfilled at all, to premises being fulfilled to a considerable extent.

Table 6.1. Fulfillment of basic premises for use, as found in the case studies.

Premise	Road Safety BOT	Water BOT	NNRCCI NAM	HESAWA TAN	Soil KEN	Dodota ETH	Research ETH	Curricula ETH	Education sector ETH
1.Evaluation should be user-oriented	Moderate	Limited	Not at all	Moderate	Moderate	Not at all	Moderate	Moderate	Limited
2. Users should be actively involved in making decisions about the evaluation	Limited	Moderate	Not at all	Limited	Moderate	Moderate	Considerable	Not at all	Moderate
3.Identify and consider the multiple interests surrounding an evaluation	Not at all	Moderate	Not at all	Limited	Moderate	Not at all	Moderate	Limited	Limited
4. High-quality rather than high-quantity participation	Not at all	Limited	Not at all	Considerable	Not at all	Moderate	Not at all	Limited	Not at all
5. Stakeholders should be trained in the evaluation process	Not at all	Moderate	Not at all	Moderate	Not at all	Limited	Not at all	Moderate	Not at all
6. Awareness of how situational factors affect use	Considerable	Considerable	Moderate	Limited	Limited	Not at all	Not at all	Limited	Limited
7. The timing of the evaluation must be seen as right	Considerable	Considerable	Considerable	Considerable	Limited	Limited	Considerable	Not at all	Limited
8. Recommendations must be seen as realistic/feasible by stakeholders	Considerable	Moderate	Considerable	Considerable	Limited	Moderate	Moderate	Not at all	Limited
9. The results of the evaluation must be made available to stakeholders	Moderate	Limited	Moderate	Moderate	Limited	Not at all	Considerable	Limited	Moderate

Source: Developed from Patton, M. Utilisation-focused Evaluation. Sage, 1986, p. 333-336.

Premise (1) *Evaluations should be user-oriented*

The intended users must be identified at the beginning of the evaluation process. From the moment stakeholders and evaluators begin conceptualising the evaluation, decisions are being made that will affect use. This has clearly never happened in any of the evaluations. Whether the evaluation will be used, and for what purpose, bears little relation to the evaluation as such. Evaluations are rarely planned and designed with utilisation in mind. They are primarily designed to produce a report.

Premise (2) *The intended evaluation users should be involved in making decisions about the evaluation*

Premise (3) *Identify and consider the multiple and varied interests around any evaluation*

Both premises are central to the ideas of participation and ownership. They are not fulfilled in the evaluations included here. Intended users and stakeholders who could be expected to have an interest in the evaluation are seldom identified, consulted and involved in actual decision-making. The cases included in this study point to an evident discrepancy between the process of evaluation and the idea of ownership and participation in aid. While some stakeholders are often consulted at the beginning of the evaluation process, they are seldom participants in the sense of being able to influence the process. *The ToR are almost exclusively drafted at Sida* and then, in most cases, submitted for comments. In a majority of the cases, at least someone (though often not more than one or two persons) from the primary collaborating organisation has commented on the ToR for the evaluation. *Decisions on which evaluators to contract are generally taken by Sida.* In five cases the collaborating organisation had a say on this issue. But in four of these cases Sida selected the team-leaders and the main persons in the team, while the collaborating organisation chose a counterpart who was to be a part of, or rather, to assist the rest of the team. While the major collaborating partner is consulted in a majority of the cases, their perception of their influence over the decision-making often points in another direction. As mentioned in a previous section, formal participation in different phases in the evaluation process does not say anything about actual influence over the decision-making. Nor does it say anything about ownership of the evaluations.

> One of the most extreme examples of "forced participation", or participation and limited influence, is the evaluation of *Curriculum Development in Ethiopia*. From a formal point of view the process could be seen as quite participatory - the organisation to be evaluated chose one person as a counterpart to work together with the Swedish consultant. The organisation commented quite a lot on the evaluation draft and also a little on the ToR. However, the organisation to be evaluated, the central organisation ICDR (Institute for Curriculum Development and Research), was against the evaluation from the beginning. The primary reason was the timing of the evaluation. The whole education system, including the curriculum component, was, at the time of the evaluation, in the middle of a transition phase which involved a major shift in educational policy and also a radical decentralisation of the main responsibility for basic education and curriculum to the regions. According to one of the ICRD representatives, because of this transition phase, the time was not opportune for an evaluation. In his view it was meaningless to evaluate the previous work since it was all in the process of changing. Likewise it was meaningless to evaluate the present work, since it was still in a transition phase. According to him, it would have been better to have had an evaluation now that the transition phase is over. When we asked if they told Sida this, he replied that they did, but that Sida did not listen. According to him the response from Sida was that "we cannot give assistance if we do not do an evaluation". He was evidently very upset by this attitude and described the situation as that of being a "sort of hostage".

Although this is a particularly bad case of power demonstration by one major stakeholder, the unequal influence over actual decision-making it demonstrates is by no means unique. The same is

evident in most of the other cases as well. Even if Sida's co-operation partners are consulted, Sida has the last word.[5] The exception to this pattern of donor-recipient relationship was the evaluation of Sida/SAREC's support to research capacity building in Ethiopia. Sida's co-operating partner ESTC (Ethiopian Science and Technology Commission) expressed a strong feeling of ownership and commitment to the evaluation. It was not at all referred to as being a "Sida evaluation". Even so, this evaluation showed patterns of exclusion. The researchers in the various projects found themselves to be largely marginalised in the evaluation process. They submitted information, but their particular questions were never given much consideration. They received the report as a fait accompli.

Premise (4). *High-quality rather than high-quantity participation* and

Premise (5). *Stakeholders should be trained in the evaluation process.*

It appeared quite difficult to find an evaluation process based on "high-quality" participation. Except for the HESAWA evaluation and the evaluation of Curriculum Development in Ethiopia, the evaluations follow a pattern where people meet with the evaluators, often in big groups, for about one hour. This short time has been pointed out as a problem by several stakeholders. It was also pointed out that group sessions are problematic since it is difficult for those in lower positions to express their views and concerns. Moreover, the evaluators seldom return to continue the discussion, or present any preliminary findings. In other words, most stakeholders are used as providers of information and are only rarely involved in the whole evaluation process. The evaluators tend to prefer to keep the evaluation process more or less to themselves. It is generally a process which leaves most stakeholders unaware of what is going on.

Closely linked to the above is the general failure to train stakeholders in the evaluation process. The evaluators prefer on the whole to keep the design and organisation of the evaluation more or less to themselves. The general experience of the people working with implementation is that they meet with the evaluators, often at some group meetings, without being told much about the purposes of the evaluations. They are contacted by the evaluators, who request a meeting. They are then briefly informed what the evaluation is all about. These stakeholders are not involved in the preparation of the evaluation. Their only role is that of "information providers". As such they seldom have the possibility of putting their own concerns and issues to the evaluators.

> In the case of the evaluation of the *Soil and Water Conservation Programme in Kenya*, for example, the officers at the District and Divisional levels were only informed about the evaluation by a letter from headquarters. The basic tenor of these letters was that they were to arrange field-visits for the evaluators. Very little information was given concerning the purpose of the evaluation, and no ToR were attached. One of the divisional Soil Conservation Officers gave the following description of their meeting with the evaluators: "The evaluators had a brief introduction at the office, where they explained that they wanted to meet the catchment committees" (i.e. farmers). Some stakeholders pointed out that these meetings were not enough. There was no time to discuss various issues with the evaluators, nor did they get the opportunity to explain why the work was beset by certain problems, problems that might be due to other developments in the local community. One officer remarked that the meeting in his divisional office "was not a discussion where everybody had a chance to comment, it was only for facilitating the evaluation."

5 In the evaluation of the National Soil and Water Conservation Programme in Kenya, Sida and the central head office, the Soil and Water Conservation Branch, had different views on which evaluators to contract. According to the Branch it was Sida who had the final say on this issue. In the evaluations of The Dodota Water Supply System in Ethiopia, nobody in the region, district or project concerned was consulted about the issues to be covered in the evaluation.

Premise 6. *Awareness of how situational factors affect use.*

A project is situated in an external context that always impacts on its progress and long-term sustainability. For any evaluation it is of the utmost importance to consider any external influences. They need to be incorporated into the analysis as such, and they certainly need to shape the recommendations emerging from the evaluation. External influences present stakeholders with restrictions as well as possibilities in utilising the results from an evaluation. The situation of the country in which the project is being implemented is a critical variable for all development projects. The state of the economy is critical for providing financial and other resources. The current political situation is equally critical for providing long-term stability and clear "rules of the game". The conclusion we can draw from our cases is that few evaluations possess such an understanding. There is too often a tendency to disregard the situational context and focus on issues which are more directly connected with the project. It is unfair to say that evaluations are "ivory towers", but they certainly have difficulty in working within a broader analytical context.

Premise 7. *The timing of the evaluation must be seen as right by stakeholders*

On a more positive note, we found that, at least for the key actors in the project, the evaluation was often well-timed The evaluation took place during a period when it was useful for the decision-making process of the particular stakeholder. The two exceptions here are the evaluation of Curriculum Development in Ethiopia, referred to above, and the evaluation of the National Soil and Water Conservation in Kenya where for various reasons the evaluation was delayed.

Premise (8). *The recommendations must be seen as realistic and feasible*

The feasibility of recommendations varied a great deal from one case to another and also between stakeholders. In the evaluation of Research Support to Ethiopia, for example, the major collaborating institution, ESTC, found the recommendations feasible and realistic while several of the research groups did not. In other cases, stakeholders found the recommendations beyond their capacity. The recommendations required new policy decisions and/or access to financial resources. For many local stakeholders, finance was a constraint since most of them were running their operations on a shoestring budget. The evaluators could produce excellent recommendations from any technical point of view, and as such they were much appreciated by stakeholders. Too often, unfortunately, the same recommendations required new investments, or an expansion in operational costs, which were beyond the means of the organisation.

Premise (9). *The evaluation must be made available*

A most distressing finding was the very limited distribution of evaluations. We could not escape the feeling that evaluations are not really made for consumption by any but a very limited group of readers. Typically, it is only the Sida programme officers, their colleagues and the top executives of the most important local counterpart organisations who receive copies. In some cases we estimated the readership to be at best 25-30 people. It was not uncommon to find a readership of around 10 people.

In conclusion, it is no exaggeration to say that there is still some way to go before the typical evaluation process provides a good starting point for a broad utilisation of evaluation results. There is, across the board, a significant disregard of these critical factors determining the relevance or otherwise of an evaluation to all stakeholders. An evaluation remains the concern of a very limited proportion of all those who have an interest in a project and are affected by its outcome.

6.3 Patterns of use as found in the case studies

> Evaluations are long and costly and should not be used to postpone decisions nor to justify decisions which have, in practice, already been made. (Sida at Work - Sida's methods for development co-operation, p. 89)

One of the major problems with conceptualisations of use into different categories such as instrumental, conceptual, legitimating etc. is that there are no clear dividing lines between different kinds of utilisation. The same evaluation can be used in different ways by different stakeholders. That there could be a dimension of legitimating use does not, for example, preclude conceptual use. Even one and the same stakeholder can use the evaluation in different ways. Nevertheless, typologies of use are helpful since they provide tools for thinking about utilisation. In the following we will use these conceptualisations as the basis for structuring our findings on how different stakeholders have made use of the evaluations. When we refer to the different examples it is important to bear in mind that one type of use does not necessarily preclude another.

6.3.1 No Use

Non-utilisation has proved to be very dominant in these cases. If the number of stakeholders provides the starting point, it could even be said that non-use constitutes the most common pattern. This is because many stakeholders are not in a position to use the evaluation findings and recommendations. As mentioned briefly in the previous section, several people interviewed by us have never seen the final reports, nor have they received information on the evaluation findings from other quarters. One such group is the beneficiaries or the target groups. Evaluation findings seldom, it seems, reach the target group. The only exception here is the evaluation "Building Research Capacity in Ethiopia", where most, though not all, research groups participating in the evaluation also had access to the report. In the other cases the target groups had no access to the evaluation findings and recommendations.

Another group, which rarely has access to the evaluation results, are the people working with implementation. The general pattern has been that *the further away you are from the centre of the project, (in terms of decision-making) the less information you get on evaluation recommendations and findings*. This means that officers at Sida and, in most cases, officers occupying central positions within the major co-operation partner (i.e. the ministries) have received the report while people further removed from the project centre very seldom have access to evaluation findings:

In the NSWCP evaluation in Kenya, the evaluation was in some instances available in the regional offices, but not at district and divisional levels. A similar pattern was found in the HESAWA evaluation: several people at regional and district levels never received the final report. In the evaluations concerning the education sector support and curriculum development in Ethiopia, the officers occupying a central position within the Ministry of Education and the other institutions concerned had received the report while officers occupying a less central position and people at regional levels did not have access to it. These stakeholders are thus prevented from making use of the evaluation in the expected sense – that is, by reading, discussing and acting on the report. In the following section then, when we discuss different types of use, we are mainly talking about the people highest up in the government structures in the countries concerned.

One of the worst cases of non-use, where several very central stakeholders were unable to use the evaluation, is probably the evaluation of the *Dodota Water Supply System* in Ethiopia. Sida phased out the support to, and evaluated, the Dodota Water Supply System already in 1988. The Dodota project was a "participatory women's project" and it was praised already in the 1988 evaluation "as a prime example of a sustainable development project". In 1995/96 Sida decided to launch

another evaluation. The primary aim of this evaluation, according to a Sida representative, was to provide general information about the sustainability of the project. Another aim was to communicate lessons learned which could be used in future support, both generally but also more specifically in an integrated rural programme which was planned in the Amhara region. We will come back to Sida later, but what concerns us here is that nobody involved in the project, neither the regional (Oromia) or zonal office, nor the project itself, had access to the evaluation findings. The evaluators visited the regional and zonal offices once, but they never came back to communicate their findings. Nor did these offices receive the evaluation report.

The evaluators carried out quite extensive fieldwork at the project site. The discussions that we had with the project implementors, who were responsible for the day-to-day operation of the water system, suggested that the purpose of the 1996 evaluation (i.e. as directed towards "lessons learned") was not clear to them. It appeared that the evaluation mission had raised expectations that the evaluation could somehow benefit the Dodota project and the people involved in it. It was also quite evident that these expectations were not met. One of the first issues raised by the officers was that they had raised the problem of the present organisation and their concerns and problems regarding their salaries to the evaluators. One of the great problems for those working with the water system now, and at the time of the evaluation, was that the government had not yet taken over the formal responsibility of the water system. One of the implications of this was that the employees were not employed by the state and that they therefore did not get pensions and could not have their salaries raised. According to one officer, the women working in the project had raised this problem with the evaluators and asked them to broach it with the authorities concerned. According to her they did not do this; at least, she did not hear from them after they left.[6]

When we asked one officer if she had learned anything from the evaluation and whether the evaluators had provided any useful recommendations (outside the report, which was never received), her answer was that the only thing they were giving was moral support, nothing else. As she put it:

> Swedish people always come here to see it [the project] as a model ... it looks like a model but the employees do not get any increments in salary...When we put the problem of salaries and pensions to the evaluators, they said that we should not worry about salaries.. [they said that] if you work honestly it will work out ... that the project must help itself.

When considering the 1996 evaluation of the Dodota system, where the people involved themselves see many problems, and had pinned expectations on the evaluation, the outcome seems unfortunate. Those involved in the project did not benefit from the evaluation. They did not even receive a copy of the evaluation report. As noted above, neither the zonal nor regional officers received a copy. This was particularly sad, since the evaluation also dealt with, and pointed out, some of the current problems in the water system. The regional and zonal offices are the institutions with the mandate to make decisions regarding the water system. Had they had access to the evaluation, they could have been in a position to make decisions that could have resulted in improvements.

6 Another problem resulting from the fact of the water system not yet having an official owner is that it is "working in a vacuum". Only minor decisions can be made locally, while decisions regarding tariffs, major repairs, equipment and spare parts must come, but are most often not forthcoming, from the zonal authorities. As a result, the water system today, as at the time of the 1996 evaluation, is in need of equipment and spare parts.

6.3.2 Instrumental use

The conventional model of the instrumental use of evaluations has not turned out to be particularly relevant or useful for understanding utilisation in these cases. As outlined above, the idea of instrumental use is that decision-makers, on the basis of new information in an evaluation, make direct and quite immediate decisions about changing programmes. Only in a very few cases, and mostly on minor issues, have the evaluations provided new information that has been acted upon and has helped to bring about changes in the projects and programmes. One such example which could be mentioned is found in the evaluation *Building Research Capacity in Ethiopia* and the Flora research project, where the evaluation contributed to one important change – the appointment of a new project leader. It was pointed out in the evaluation that the project leader had too many other commitments that diverted him from the project and that a new leader was needed. According to one representative of the project, this problem had not been evident to the research group before (since they worked in a team) and the evaluation was therefore seen as valuable in pinpointing it.

These kinds of *direct decisions* arising from *new information or inputs* provided in an evaluation are, however, quite rare. In most other cases it has been difficult to identify direct operational changes in the programmes which could be attributed to new insights provided in the evaluations. A common statement is that the changes "would have happened anyway" - without the evaluation. The changes had already been discussed or even decided upon. Another frequently recurring comment from the stakeholders was that the evaluation did not provide much "new" information. For example, one of the stakeholders at the Department for Road Transport and Road Safety in Botswana found the report quite accurate "on the descriptive side" but said concerning its usefulness from a practical point of view that:

> ... the report contained nothing new; the consultant repeated what we had told him ... most of the recommendations are our own ideas.

Stakeholders communicated similar feelings in the evaluations *Building Research Capacity in Ethiopia*, and *National Soil and Water Conservation in Kenya*. In the evaluation *Building Research Capacity in Ethiopia* the views on how the evaluation was used differed greatly among the stakeholders. The co-operating partner, ESTC, claimed that the conclusions and recommendations of the evaluation to a great extent were considered and applied at the ESTC. They also claimed that the evaluation had resulted in major redirections in certain research projects. According to the research groups themselves, however, the redirections, which followed from the evaluation, would have happened anyway. Some even stated that the redirections already had been decided by ESTC before the evaluation was undertaken.

In the case of the evaluation of the *National Soil and Water Conservation in Kenya*, the ideas promoted in the evaluation had actually been included in work plans. The evaluation was delayed for various reasons, among other things, because of the conflicts regarding which evaluators that were to be contracted. At the time of the evaluation, the Branch had just completed its Plan of Operation for the next period in its draft form. The general attitude at the Branch was that several recommendations in the evaluation were already there in the Plan of Operation and their concept paper developed a year earlier. According to their judgements then, the evaluation could not be understood in terms of instrumental use, it rather repeated what was written in the Plan of Operation. This repetition also cased some frustrations among the officers at the Branch. The perceived problem was that the recommendations were listed as coming from the evaluators and no reference were made to the Plan of Operation. As one officer expressed it "they [the evaluators] do not refer to the Plan but say the same... as if we have not done anything, it is very annoying".

While in the case of the evaluation of the *Small Business Development Programme in Northern Namibia* non-utilisation certainly predominates, the evaluation was nevertheless used by some of the stakeholders. Firstly, Sida obviously needed the evaluation in order to better understand the current position of the project. That knowledge was needed in order to take decisions about the orientation of future support. The other stakeholder was the current Regional Secretary and co-ordinator of the project. He was involved in the evaluation and obviously in good position to understand the analysis and its practical implications. When Sida decided that an advisor should be attached to the project, the Regional Secretary became the natural counterpart to the advisor. Thus, the utilisation we saw in this project was limited to two stakeholders: Sida and, more importantly, the Regional Secretary.

6.3.3 Conceptual use

Evaluations, it would seem, rarely provide new information. This does not mean that they are not considered useful or that they have no impact. Instead their chief value lies in their repetition and/or re-interpretation of available information. This is a typical feature in the cases of conceptual use. It is useful at this stage to remind the reader of the definition of conceptual use. *Conceptual use* refers to the impact of evaluations on policies, programs and procedures, through sensitising persons and groups to current and emerging social problems. This will eventually lead to changes in the way they think of social issues and, at a later stage, to changes in policy and procedures.

In essence this means changing the knowledge structures of individuals, and in a following step exercising some impact on how an organisation thinks about its strategy, policies etc. And finally, new modes of implementing projects and programmes. Conceptual use therefore has two important properties: (1) it does not happen overnight; and (2) it is difficult to analyse and verify empirically. In other words, conceptual use can be a slow process where the same issues often need to be raised in a constant dialogue between stakeholders. New insights need to be communicated and processed before they can finally be accepted or refuted, i.e. impact on an individual's knowledge structure.

What would we call conceptual use? Obviously it can mean many things, from something quite simple to something involving drastic changes in the way we think about things. To illustrate somewhat the span involved, we would like to pick some examples from the HESAWA and NSWCP evaluations.

One of the major recommendations in the evaluation of the *National Soil and Water Conservation Programme in Kenya* (NSWCP) was that Sida should phase out its support and that the Programme must start to prioritise. The question of Sida withdrawal had already been raised both in the previous evaluation and in a sector review. It was thus not a new issue. As one officer at the Branch phrased it: "at the time of the evaluation we already knew that Sida was to withdraw...in fact we had already started to withdraw from certain areas such as nurseries at that time." At the same time, however, the officer stressed that the evaluation pointed at the need for prioritisation. The evaluation *made them 'more aware'* and '*made it more clear*' that Sida would phase out and that there was a need for prioritisation. Similar comments came from other officers at the Branch. According to another officer the evaluation made them "*more awake*"'. It forced them to "face the reality and think seriously on how to make the planning". In his view this was the major benefit of the evaluation. Most of the other recommendations were already there. Thus, in spite of the fact that it was not new information, it seems like the evaluation had an impact in that it made people in the Programme more aware of Sida's withdrawal and the need for prioritisation.

The mid-term 1996 evaluation of *Health Through Sanitation and Water in Tanzania*. (HESAWA) contains a number of findings and recommendations that concern strategic and policy-related issues. One concerned the role of the private sector in water development projects.

> It is recommended that if the Programme Management accepts this scenario, the Programme will assist the private sector to enable them to take over the construction component. This needs proper preparation. Time points for reflection on direction and progress of privatisation of construction have to be built in the planning. (p.75)

Another finding/recommendation concerned the issue of sustainability and focused on management strategies.

> The Mission therefore recommends that the district HESAWA Teams give more systematic support, particularly to phased-out villages in terms of advice on more efficient and effective institutional structures and organisational arrangements, and on the development of related Committee capacities. This would increase the sustainability of the systems. (p.73)

The information and the recommendations in the 1996 evaluation could by no means be considered as new. As most of the officers at the zonal office concluded, the 1996 evaluation, in general, merely repeated and dealt with the issues which were addressed in an earlier evaluation conducted in 1992. The Programme thus already knew most of the problems identified in the evaluation and several of them were, as the evaluators also noted, already being worked on. (One of the questions in the 1996 evaluation was to what extent the Programme had started to implement the recommendations from the 1992 evaluation.) Even if the repetitiveness and non-newness of the 1996 evaluation was presented as a problem, or rather a disappointment, the evaluation was presented as useful. The evaluation was described as a "reminder", and as a way to "see whether the programme is on the right track". As two officers put it:

> The evaluation helped me greatly...it gave me room to see difficult problems... the issue was not new but the evaluation reminded us.

> Even if there was not much new it is good to have the problems put in a structured way ... even if we think that we know certain issues, it is important to discuss them.

The evaluation also appears to have influenced the planning of the fourth and last phase. Most of the problems identified to be addressed in the next phase were similar to the problems identified in the evaluation.[7] The evaluation suggests that learning and changing directions in big organisations take time, and that the idea of instrumental use (i.e. direct decisions on the basis of new information) therefore tends to be unrealistic.

The evaluation of the *Dodota Water Supply System*, referred to under "no-use" is another quite good example of conceptual use. As mentioned above, the Dodota Water Supply System was a "participatory women's project" which was evaluated already in 1988 when Sida withdrew. Sida initiated the 1996 Dodota evaluation. One of the aims of the 1996 evaluation was to provide lessons learned which could be used in future support generally but also more specifically in an integrated rural programme which was planned in the Amhara region.

For most of the stakeholders, the ex-post evaluation in 1996 was a case of non-use. For Sida, however, the evaluation contributed knowledge of a conceptual nature. According to a Sida Programme Officer with long experience of water and sanitation projects:

7 E.g. inadequate management capacity of user groups, shortage of skilled resources at village level, unbalanced gender roles, more appropriate technology, slow uptake of household latrines, need for rehabilitation and retraining, weak support from the private sector

> The evaluation taught us about the importance of involving women in project implementation.

As a tool for lessons learned for Sida, the evaluation was praised. The evaluation had been used to a very high extent in discussions about gender and water, not only within Sida but also in discussions with other donors. According to a Sida officer, the evaluation should have been important for Ethiopia but the contention was that it probably was not utilised by any Ethiopian stakeholder. The argument was that "Dodota was participatory and a women's project" and that "the Ethiopians are not gender aware and work according to top-down directives". According to the Sida representative, "the original intention was that it would be used for lessons learned in Amhara but that it turned out to be a learning for Sida".

Sida officers often expressed disappointments over the lack of action in projects and programmes - "nothing happens". The implicit conclusion was that the programmes or projects are not learning from - or acting upon - evaluation recommendations. This could very well be a problem in some of the programmes. We do not have enough intimate knowledge of the programmes in order to judge this. What we have learned from this study is that organisational learning takes time. Learning and changes in orientation of a programme does not happen over night. It is a long process where evaluations seem to have some impact.

6.3.4 Evaluations as means to communicate and legitimise interests

> Evaluation work should be transparent and carried out in a spirit of co-operation. (Sida's Evaluation Policy, p.2.)

Evaluations are not situated in a vacuum, but in the context of different stakeholders interests. What these cases have shown is that stakeholders often try to use evaluations to communicate and legitimise their interests. Furthermore, the co-operation partners often see the evaluations as a way for Sida to convey its interests and concerns. We shall present some examples of this pattern.

Even if the examples below show that the co-operation partners perceive the evaluations as used by Sida to communicate or legitimate decisions, it does not mean that the co-operation organisation do not have interests which they bring to the evaluation. It rather means that, compared to Sida, they are often not so successful in their efforts. The evaluation of the *Road Safety Programme in Botswana* is a good example of this point. The Botswana Road Safety Improvement Project was carried out between 1985 -1995. It was based on a programme proposal drafted by the Swedish National Road Consulting AB, (SweRoad), for road safety improvements in Botswana. SweRoad was later given the assignment to co-ordinate the implementation of the project. The main objective of the project was to reduce the number of traffic accidents and their social and economic consequences. When the programme was finalised it had passed through three main phases and a fourth winding-up phase.

The present 1995 evaluation was carried out at the same time as the Swedish support was phased out. The evaluation was expected to give Sida feedback on the relevance and efficiency of this type of support to traffic safety and its applicability to similar projects in other countries. Another objective of the evaluation was to arrive at recommendations for the continued work at the Department of Road Transport and Road Safety in the short, medium and long-term perspective, as well as assess the Department's capacity (financial, manpower and material) to implement an extensive road safety programme. The evaluation should identify priority areas of need for external technical assistance.

In this case it is clear that the stakeholders at the Department of Road Transport and Safety in Botswana saw the evaluation, as being used by Sida to legitimise a decision already taken. As one stakeholder put it:

> Sida wanted to get information for the sake of phasing out the support ... Sida had already decided to withdraw from the project, we were made to understand that, and we had no choice. Sida left us halfway... it was Sida's evaluation, they wrote the TOR and we didn't object, they appointed the consultants and they paid for the evaluation.

The perception was thus that Sida wanted to use the evaluation in order to phase out the support. It was also strongly felt that this strategy had succeeded. According to one of the officers:

> The evaluation made Sida achieve what it wanted, namely to pull out of Botswana, to end the project.

What this case reveals is that also the co-operation partner had an interest that they wanted to get through in the evaluation, namely the continuation of the support. As one of the officers in the Road Safety Programme put it, the evaluation was not to any great use to the Department of Road Transport and Safety:

> "...because it did not convince Sida to continue the support."

The evaluation of the *Education Sector Support to Ethiopia* is perhaps the most obvious example of how evaluations are used to communicate and legitimate Sida interests. Sida took the initiative for this evaluation at an annual sector review meeting when one of the Sida officers brought up the idea as part of the planning for the new contract period. Although the idea was put as an open question - it was important to ask open and frank questions of what had been done and what should be done in the future - it seemed quite obvious to the co-operation partner that Sida used the evaluation to pave the way for a coming reorientation of the programme. It was felt that Sida had already decided to reorient its programme (to concentrate on a few areas and focus on qualitative issues and capacity building rather than school construction, the latter being a prioritised issue for the Ministry of Education) but that it would be easier to bring out this issue through the medium of an evaluation. Sida's interests were clearly expressed in the TOR and the evaluators did what was expected of them.

According to the Ethiopian side, the evaluation did not matter in terms of their own policy. They agreed to an evaluation since it would be difficult not to play along when Sida launched the idea with such force. Some stakeholders on the Ethiopian side said that the evaluation gave them the opportunity to understand Sida's intentions, which could be helpful to them in reorienting their donor strategy. As one of the stakeholders put it, one good thing about the evaluation was that they "*learned Sida's intentions*" and that the evaluation therefore increased their awareness and preparedness for the coming period. It should be pointed out that the Sida officials involved expressed surprise at the Ethiopians perceiving the evaluation in this light. They furthermore denied that any decisions had been taken beforehand.

The evaluation of the *Soil and Water Conservation Programme in Kenya*, (NSWCP), referred to in the previous section, is another example of evaluations being used as a means of communicating interests. The comments from some of the officers working on the programme indicate that the evaluation was not in all instances seen to reflect the view of some independent evaluators. The Kenyan evaluators had also, from discussions with Sida, understood that the issue of phasing out was important for Sida. As one of the evaluators put it: "the programme had been going on for too long and we were to 'provide reasons to trim down'... and 'find out how to terminate it painlessly".

The evaluators found their position between the Branch and Sida as problematic — "we felt that we were going to be the bad guys".

The remarks made by one Sida officer tend to confirm that the evaluation was seen as a means of communicating their interests. Certain things are communicated to the evaluators in discussions where you give the evaluators "hints" on what questions you consider important -"you don't put everything into the TOR". According to this officer, one of the issues raised in the discussions with the evaluators concerned the need to phase out. During this meeting they explained to the evaluators that "Sida has been here for a long time" and that they wanted "radical recommendations on new directions for future support" and "phasing-out strategies". Sida's comments on the draft report also seem to convey a picture of the evaluation as a means of getting a message across. It is stated, for example, that:[8]

> It is probably time for the branch to start using PRA in monitoring and evaluation and impact assessment. *We find it appropriate to indicate this in the report especially at this point in time as the branch is in the process of redirecting its focus for the future.* (Our italics)

This and other similar comments are not actually comments on the report itself. Rather they are expressions of Sida's interests in how the programme should be managed, interests which could easily be articulated in the semi-annual meetings. Moreover, these kinds of informal communication go against Sida's own policy of the importance of transparent evaluation processes (see quotation above).

The evaluation of *Curriculum Development in Ethiopia* is another very clear example of evaluations being used as a means of communicating and legitimating interests. In this case the evaluation was used by Sida to stimulate a debate and a rethink on certain issues. As mentioned before, the primary organisation to be evaluated, the Institute for Curriculum Development and Research (ICDR), was against the evaluation from the beginning, since the whole education system, including the curriculum component, was, at the time of the evaluation, in the middle of a transition phase which involved a major shift in educational policy and also a radical decentralisation of the main responsibility for basic education and curriculum to the regions. Given the attitude that the evaluation was meaningless (in the context of the transition phase), it is not surprising that the general view concerning the usefulness of the evaluation was quite negative. The evaluation was presented as important only for Sida and it was made clear that the Ministry of Education had their own evaluations.

The image of the evaluation as imposed by Sida on an unwilling "co-operation partner" was confirmed by one of the Sida officers. The evaluation was presented as very important, given the previous history of curriculum development in Ethiopia during the Mengistu regime. In the light of the new policies at Sida, stressing privatisation and decentralisation of curriculum development and textbook production, the evaluation was "100 per cent necessary" according to Sida. The intention was "to stimulate a debate on these issues in Ethiopia". It was emphasised that the evaluation should not be seen as a Sida report, but was intended to be useful for Ethiopia and the institution

8 It is also stated that "it is evident that the Branch does not have the capacity and experience to start serious interventions in the ASAL areas. As this is one of its future areas of focus... it is prudent for the branch to start exploring the possibility of collaborating with NGOs and other actors already active in the ASAL with the aim of working through them and eventually handing over during the phasing out period. It would be appreciated if this thinking is included in the final report."

in question and also to serve as a basis for discussion regarding future support by Sida. The contention was, however, that it was very difficult to make the co-operation partner understand this.[9]

This particular evaluation does not seem to have been useful to any of the major stakeholders. It has also not been used by Sida to learn more about curriculum development in Ethiopia. The evaluation was, by Sida officers, referred to as too detailed and as a "communication between experts" (that is between the evaluators and the curriculum developers in Ethiopia). A few people did, however, benefit from this evaluation, namely the persons who assisted the evaluator in his work. All of these persons have stated that they learned a lot both about methods and about curricular problems by assisting the evaluator in his data-gathering.

The Dodota evaluation also has a dimension of legitimating use. It seems quite clear that the evaluation was supposed to communicate some more specific Sida priorities to the authorities in the new programme in Amhara, namely those of gender and participation. The issues of gender (or rather women) and participation were dealt with also in this report and were, as such, nothing entirely new for Sida. But Sida wanted a stronger emphasis on these concepts in the future co-operation with Ethiopia. This evaluation can thus be seen as an example of both conceptual use and of how evaluations are used to communicate a position — in this case to the authorities in Amhara concerning the importance of gender and participation.

6.3.5 Ritual Use

The concept of ritual use refers to a situation whereby the initiation of an evaluation first and foremost come to symbolise desired qualities of management or of organisation. An evaluation takes the form of a ritual when it is seen as something, which has to be done, and is therefore done without much consideration of why it is done or how it should be used. Several of the evaluations included here, especially those which are included as a part of the agreement period, run the risk of becoming rituals. An evaluation has to be done but nobody is really sure what purpose it is meant to serve. The TOR often turn out to be wide and unclear and encompass "all and nothing" at the same time.

Ritual use is evident in the ex-post evaluation of the Rural *Village Water Supply Programme in Botswana* (RVWSP) which received Swedish support from 1971 to 1993. More than half a million people have been estimated to be served by water supplies installed by the programme. The main purpose of the 1995 evaluation was to produce lessons learned from the by then terminated support programme. Furthermore, the lessons learned were primarily directed at Sida, but the consultants also considered what lessons the Government of Botswana might learn from experience of the RVWSP:

> This study is an evaluation in its true sense. Its main purpose is to determine any lessons which might be applied to other programmes supported by Sida, or which could be used by Botswana in its future water development activities. (p.4)

The question, then, is to what extent this purpose was achieved. Did the stakeholders in the RVWSP utilise the findings for purposes of learning?

9 When it comes to the subsequent part of the evaluation process it could be added that the ICDR commented on the draft report in writing since the evaluator had left for Sweden. It is very evident that the ICDR and the evaluator(s) had divergent opinions on a number of central issues. The evaluators agreed that the comments by ICDR should be included in the appendix in the final version of the evaluation. In the final version the evaluators listed some of the ICDR's comments and addressed them at the beginning of the evaluation. On the basis of these comments they made new, quite radical suggestions (since they are concerned with recommendations to Sida for the future support). The ICDR did not receive this final report and were thus unaware of these recommendations or "comments on the comments".

The Ministry of Finance (MoF) set up a reference group consisting of representatives of the concerned Government authorities. They represent fairly accurately the major stakeholders in the project: the Department of Water Affairs (DWA), the Ministry of Mineral Resources and Water Affairs (MMRWA), the Planning Office of the same Ministry, the Ministry of Local Government, Lands and Housing (MLGLH), the Ministry of Health (MoH), the Development Planning Office, MoF. The Kgatleng District Council represented the District Councils. Sida was represented by two persons.

The views of the stakeholders on what they found useful in the evaluation and what they used it for are important. A few participants in the group described the whole process as a learning exercise. They knew very little or nothing about water affairs, and had little to do with them in their daily tasks. They were representing their Ministry. This presumably also meant that they had little opportunity to use their newly found knowledge in their daily work.

Of those in the group more familiar with water issues, few could confirm that they found anything new in the evaluation. The information it contained was more or less known to most of them beforehand. All of them had been providing the evaluators with information.

All of them said, however, that they found the report interesting; some specified the discussion of sustainability as particularly interesting. This suggests that there was an element of conceptual usefulness in the evaluation. Others said it was useful to have a programme of this size presented and discussed so comprehensively. However, almost three years after the evaluation was done no one could specify what they had learned from the exercise. Nor could they recollect any actions taken which had their roots in any of the recommendations or findings in the evaluation.

Regarding the usefulness of the recommendations, almost everybody commented that the recommendations were very good. But most of them were already in the process of implementation, a process which in most cases had started before the evaluation was made.

There were also stakeholders outside the reference group, the most important of them being the district councils. According to the evaluation they had benefited greatly from the RVWSP and its capacity-building efforts. The councils had one representative in the reference groups. The team had visited a few councils and interviews had been conducted, mostly for extracting information. Even though the District Council officers were important stakeholders in the project, and in the evaluation, not one of them had an opportunity to learn anything from the evaluation, nor could they be expected to learn anything from it, because (a) they had never seen a copy of the evaluation report, and (b) the recommendations were not really for them to act on. Those who had an opportunity of using the evaluation findings were, firstly, those who participated in the reference group and who had a position enabling them to act on the findings, and secondly, those outside the reference group who read the evaluation.

This case presents very little in the way of instrumental or conceptual use. For Sida and Botswana the chief value of this evaluation seems to have been that it crowned 20 years of development cooperation. It gave evidence that the programme had been efficient and effective. It is a good example of the evaluation as a ritual. For the stakeholders directly concerned with water provision in Botswana - the District Councils - the evaluation was a case of non-use. The report was never distributed to them.

6.4 Conclusion

> ...evaluations provide new knowledge and contribute to organisational learning rather than merely confirm what is already known. Evaluations are carried out from an independent perspective using innovative approaches.
>
> - Lewin 1993, Evaluation Group, Sida, on "The characteristics of a good evaluation system" p. 22

> We learned the intentions of Sida...
>
> – A stakeholder in the evaluation of the Education sector in Ethiopia commenting on some of the "good things" about the evaluation

The above quotations give an indication of how evaluations can be perceived in theory and practice. The value of evaluations in terms of usefulness to stakeholders is often quite different from what is believed in textbooks and expressed as desirable by evaluation policy-makers. The evidence from our case studies only serves to confirm this. There are two general observations that can be made from the case studies. Firstly, the pattern of use is much more complex than anticipated. Secondly, very few stakeholders have a direct possibility of acting on the findings of an evaluation. In Table 6.2 we have summarised the patterns of use as found in the case studies.

Table 6.2 Dominant Types of use as found in the case studies

	Instrumental	**Conceptual**	**Legitimising**	**Ritual**	**No Use**
1. SBDP Namibia	Financier PMC				TG
2.Village Water Supply Botswana		CP PMC		Financier	LI TG
3. Road Safety Botswana			Financier		PMC
4. Dodota Water Supply Ethiopia		Financier	Financier		LI CP PMC
5. Research Capacity Ethiopia	LI/TG (CP)	CP Financier			PMC LI/TG
6. Education Sector Ethiopia			Financier		CP LI PMC
7.Curriculum Ethiopia			Financier		CP PMC
8. HESAWA Tanzania		LI PMC			TG
9. Soil-Conservation Kenya		PMC	Financier		TG LI

Comments: Financier = Sida, Co-operating partner (CP) = The government agency which signs the contract with Sida, Project management central (PMC) = The agency that is responsible for the management of the project, Local implementor (LI) = A unit that works with the implementation of the project on the local level. Target group (TG) = The beneficiaries of the support

First of all, and as follows from the presentation in the former chapter, it should be pointed out that *many stakeholders are not in a position to use the evaluation findings and recommendations*. Several people interviewed by us have never seen the final reports, nor have they received information on the evaluation findings from other sources. The only exception here is the evaluation "Building Research Capacity in Ethiopia", where most stakeholders had access to the evaluation report. Evaluation findings seldom, if ever, reach the target group. They cannot make use of the evaluation in the expected sense – that is by reading, discussing and acting upon the report. But they can, provided that a participatory method is consistently pursued, still benefit through an evaluation process which in itself becomes an opportunity for learning.

The conventional model of the instrumental use of evaluations has not turned out to be particularly relevant or useful for understanding utilisation in these cases. Only in a very few cases, and mostly on minor issues, have the evaluations provided new information which have been acted upon and contributed to changes in the projects and programmes.

A common statement heard from stakeholders was that changes that had taken place, "would have happened anyway". Certain actions had been recommended in the evaluations, but they had already been discussed or even decided upon before the evaluation. Another frequntly recurring observation by the stakeholders is that the evaluation did not provide much "new" information.

Even if evaluations do not provide much in terms of new information and inputs, this does not mean that they are not considered useful or that they have no impact. Instead they seem to be important in *repeating and reminding* the people involved of certain issues. This reminding and repeating character of evaluations have been particularly evident in the HESAWA and NSWCP evaluations. These two evaluations could, together with the evaluation "Building Research Capacity in Ethiopia", be considered as the most elevating examples of use and usefulness.

Sida officers' in particular often expressed disappointment at the slow pace of change in projects: "Nothing happens". The implicit conclusion is that the projects are not acting on evaluation recommendations. This could very well be a problem in some of the projects. We do not have enough intimate knowledge of the projects in question in order to verify this. But it does happen that evaluations are discussed (at least in the NSWCP and HESAWA cases) and some of the recommendations are incorporated in work plans. A lesson learned from studying these evaluations is that organisational learning takes time. Learning and changes in orientation of a programme do not happen over night. You learn from a multitude of sources of which evaluations is one, but not necessarily the most important one.

Stakeholders among Sida's co-operating partners (those who have seen the reports) often knew quite a lot about the findings and had things to say about several, often more detailed issues, while Sida officers often do not. Their comments, and presumably interests, are often on a more general level of failures or successes in the project as a whole. Sida officers very often referred to the evaluation as a "starting-point for discussions" with the co-operation partners.

There could be several reasons for the picture outlined above. One could be the workload and information overload especially at Sida Stockholm. This made it difficult for officers to assimilate more detailed information. Ex-post evaluations of completed programmes naturally fall into the category of conceptual use, since the knowledge generated by these evaluations is supposed to provide, not only information on the programme or projects, but also lessons learned for the future.

Evaluations take place, not in a vacuum but in the context of different stakeholders' interests. What these cases have shown is that stakeholders often try to use evaluations to communicate and legiti-

mise their interests. The study has shown that the co-operation partners often see the evaluations as a way for Sida to convey its own interests and concerns. Even if the co-operation partners perceive the evaluations as used by Sida to communicate or legitimate decisions, this does not mean that the co-operation organisation does not have interests that it brings to the evaluation. Rather it means that, compared to Sida, they are often less successful in their efforts.

7. Understanding the current pattern of use

What can these cases teach us about utilisation? The starting point here is that there is a discrepancy between what transpires when we evaluate and the basic principles of Swedish aid. From the policy perspective of participation and ownership, the evaluation process leaves a lot to be desired. Most notably, stakeholder categories such as the beneficiaries and the ones involved with implementation are marginalised in the evaluation process. They are rarely in a position to use evaluations, for the basic reason that they are rarely given access to them. Even if there is a component of some formal participation by local organisations in the evaluation process, their sense of ownership is limited. Evaluations are too often regarded as the donor's concern and privilege.

Why is this? Possible explanations for the situation can be searched for both within and outside the evaluation process itself. Evaluation research often, naturally perhaps, searches for factors within the evaluation process itself. Many of these have proved to be important also in relation to the cases explored here, and we will touch upon them in this section. After that we will elaborate on important factors which affect evaluation utilisation but which lie outside the evaluation process and therefore are also more difficult to change.

7.1 Factors internal to the evaluation process

7.1.1 Participation, ownership and use

Participation may not necessarily lead to ownership, but it is nevertheless an important prerequisite for ownership to occur at all. Effective participation should lead to ownership in the sense of having an influence over the evaluation process. A stakeholder must be convinced that the evaluation is carried out in order to serve his/her own interests. The cases included in this study have shown that participation and ownership is more difficult to obtain than what perhaps many aid practitioners may have thought.

In this section we will particularly discuss the relationship between participation and ownership. What does participation do? And what kind of participation has actually taken place? Is effective ownership a matter of elaborate evaluation methods? Or can it be attained in other ways?

We shall start this section be presenting stakeholder participation in some of our cases. We have chosen four cases that illustrate the range of combinations between participation and ownership and how they relate to use. The first case is the evaluation *Building Research Capacity in Ethiopia*. This case is an example of strong influence and ownership from the part of the co-operating partner. This case will be contrasted with the evaluation of *Curriculum Development in Ethiopia* which is a case of limited or non-existent ownership. The third case is the *HESAWA* evaluation in Tanzania. In contrast to the other evaluations that were based on more conventional data-collection techniques, this evaluation was based on workshops where the participants could articulate their ideas and concerns in a more free way. The HESAWA case will here be contrasted with the evaluation of the *National Soil and Water Conservation Programme in Kenya*, which was based on more conventional data-collection techniques.

We end by discussing two issues that we have found to be particularly interesting with regard to the relationship between participation and ownership and its impact on utilisation. The first is the traditional focus of aid evaluations on written, rather than verbal, communication of findings and recommendations. The other is the role of the evaluator. From our cases this role appears much more complex than the picture of the dedicated, objective investigator might suggest.

The evaluation *Building Research Capacity in Ethiopia* forms an interesting case because it contains a strong element of local stakeholder ownership and marginalisation at the same time. Table 7.1 presents a picture of how the various stakeholders found themselves involved in the evaluation process. It should be noted here that the table presents a picture of stakeholder *involvement* rather than participation in the way in which it has been defined earlier in this report. As stated in section 5.2, a meaningful conceptualisation of participation in an evaluation context would be that it is *a process whereby a stakeholder is taking part in, and has influence, over the evaluation process*. As we will show by the example of *Curriculum Development in Ethiopia*, the fact of somebody participating in a process does not necessarily mean that they have influence over the process.

This is also evident in the case of *Building Research Capacity in Ethiopia* and the research groups (LI/TG). Three research groups read and commented, quite intensively, on the draft version of the evaluation report. But these, in several instances very critical comments, were not taken into account. While the research groups were "to a considerable extent involved in commenting on the draft", as is shown in the table below, they did not have much influence. A table presenting a picture of participation (considering influence) would thus turn out in a different way.

Table 7.1 Stakeholder involvement in "Building Research Capacity in Ethiopia"

Stakeholder/ Activity	Financier	CP (ESTC)	LI/TG 1.	LI /TG 2.	LI /TG 3.	TG/TG 4.
Preparing ToR						
Selecting evaluators						
Providing data						
Collecting data.						
Interpretation/ Dev.of Recommendations						
Commenting on draft/recom.						

Involved to a considerable extent | Involved to some extent | Involved to a limited extent | Not involved

Comments:
Financier = Sida,
Co-operating partner (CP) = The government agency which signs the contract with Sida,
Project management central (PMC) = The agency that is responsible for the management of the project,
Local implementor (LI) = A unit that works with the implementation of the project on the local level.
Target group (TG) = The beneficiaries of the support

The pattern can be summarised in a few points:

* Both Sida and the co-operating partner ESTC took, what seems to be, an equally active part in the evaluation process, particularly in the early, formative stages of the evaluation. According to ESTC, due to a conflict between ESTC and SAREC at the time, the two parties did not prepare the

TOR together but separately and the evaluators thus received two TOR to integrate and act upon. When it came to the evaluators the agreement was that the two institutions should choose one evaluator each and then together agree on a third person. This third person, who, after a lot of discussions eventually was agreed upon, was seen as a neutral person (at least from the view of ESTC).[10]

* Other local stakeholders, in particular the research groups, were mainly information providers. As noted above, they read and commented, quite intensively, on the draft version of the evaluation report, but these comments were not taken into account.[11]

The most central feature of this case is that the ESTC expressed a strong commitment to the evaluation. It was not at all referred to as a Sida initiative. The ESTC saw it as "their" evaluation, initiated and done to serve their interests and information needs. ESTC representatives stressed that they found the evaluation very useful for their work. The evaluation was presented as having pointed out important problems regarding the projects. According to the ESTC, the conclusions and recommendations of the evaluation were to a great extent considered and applied by the ESTC - especially those concerning questions of integration and capacity building. In this respect this case differs from the others.

Ability to exercise a great influence over the evaluation and the resultant feeling of ownership go a long way towards explaining the usefulness of the evaluation to the ESTC. Sida also considered the evaluation useful, even if they did not feel that it provided much new information. According to a Sida representative, the evaluation formed "a starting point for discussions with the ESTC".

Paradoxically enough, the evaluation process also contained a strong element of marginalisation, as shown above. Ownership coexisted with exclusion. The subjects of the evaluation - the research groups - did not share a feeling of ownership. On the contrary, they did not appreciate the evaluation at all. They were never involved or consulted in the initiation of the evaluation. And, as noted above, their comments on the draft report were not considered. They were subjects of control rather than anything else.

This evaluation can be contrasted with the evaluation of *Curriculum Development in Ethiopia.*

10 One interesting point in relation to the preparatory phase is that the ESTC put 'their man' into work during the process of preparation, before the formal evaluation exercise was started.

11 Both the Swedish and Ethiopian representatives of the Flora and BRTP projects received the draft report and all of them submitted highly critical comments either verbally during discussions or by written documents. These very critical comments were partly concerned with basic 'incorrect information' in the evaluation such as the number of students who had achieved their exams, when the projects started, the number of persons in editing boards, outlines of budget and costs, and utilisation of laboratories, and the purposes of the research projects.

Table 7.2 Stakeholder involvement in "Curriculum development in Ethiopia"

Stakeholder/ Activity	Financier	CP	PMC/LI	TG
Initiation				
Preparation of TOR				
Selecting evaluators				
Providing data				
Collecting data				
Interpreting data				
Commenting on draft report				

Involved to a considerable extent | Involved to a moderate extent | Involved to a limited extent | Not involved at all

Comments:
Financier = Sida,
Co-operating partner (CP) = The government agency which signs the contract with Sida,
Project management central (PMC) = The agency that is responsible for the management of the project,
Local implementor (LI) = A unit that works with the implementation of the project on the local level.
Target group (TG) = The beneficiaries of the support

In contrast to *Building Research Capacity in Ethiopia,* the evaluation of *Curriculum Development* is an example of very low level of influence, ownership and usefulness from the perspective of the co-operating partner. The organisation to be evaluated, the central organisation ICDR (Institute for Curriculum Development and Research), was against the evaluation from the beginning. This evaluation is also a good example of the fact that involvement does not necessarily imply influence. From a formal point of view the process could be seen as quite participatory; 1) the ICDR chose one person as a counterpart to work together with the Swedish consultant, and 2) the organisation commented quite a lot on the evaluation draft and also a little on the TOR. Their actual influence over the evaluation process itself was, however, very limited.

Even if ownership is one of the most important variables for explaining usefulness, it is not the only explanation. The HESAWA case suggests that an evaluation can be perceived as useful even if there is a limited sense of ownership among some stakeholders. Table 7.3 presents the involvement found in the HESAWA evaluation.

Table 7.3 Stakeholder involvement in "HESAWA in Tanzania"

Stakeholder/ Activity	Financier	CP	PMC (Zonal HESAWA)	LI (Regional HESAWA)	LI (District HESAWA)	TG
Preparing ToR						
Selecting evaluators						
Providing data						
Collecting data.						
Interpretation/ Dev.of Reco-mmendations						
Commenting on draft/recom.						

Involved to a considerable extent | Involved to some extent | Involved to a limited extent | Not involved

Comments:
Financier = Sida,
Co-operating partner (CP) = The government agency which signs the contract with Sida,
Project management central (PMC) = The agency that is responsible for the management of the project,
Local implementor (LI) = A unit that works with the implementation of the project on the local level.
Target group (TG) = The beneficiaries of the support

The pattern of participation in this case is different from the other cases. There is participation from a much broader range of local stakeholders. Furthermore, they are participating not only as providers of information, but also in discussions of the findings of the evaluation. One prerequisite for usefulness is of course that the *stakeholders feel that their information needs are addressed* in the evaluation. Evidence from the HESAWA evaluation shows that the participatory character of data-collection, and subsequent analysis, largely explains why this evaluation was regarded as useful by stakeholders who never took part in the initial preparatory steps of the evaluation. An important reason for this is the methodology (SWOT) used by the evaluators. This methodology draws on the opinions of the participants, how they see the programme's past, present and future in terms of successes, weaknesses, opportunities, and threats. Some of the interviews at district level revealed that a properly conducted participatory exercise could become an important learning exercise in itself. The exercises were judged as very useful, as a kind of "self-evaluation" whereby one had the possibility to identify different problems in the implementation and discuss different ways to handle these problems. The evaluation is not seen primarily as a control mechanism, as something to fear or as a phenomenon where problems and difficulties should be concealed for the controller. As one officer put it:

> Before the evaluation people came as auditors .... asking "What is this?!" and "Why do you do this?!" but the 1996 evaluation was encouraging. I found myself evaluating myself... I learned what I did wrong and where I should improve.

The issues treated were to a large extent based on the concerns and the information needs of the people working in the Programme and, as a result of this, the evaluation was perceived as relevant and important.

The HESAWA evaluation, and the pattern of involvement in the process of collecting and interpreting information, could be contrasted with the evaluation of the *National Soil and Conservation Programme in Kenya*. (a programme with a similar structure). In contrast to HESAWA, the ones working directly with implementation in the NSWCP were marginalised. The officers at the District and Divisional levels were only informed about the evaluation by a letter from headquarters. One of the divisional Soil Conservation Officers gave the following description of their meeting with the evaluators: "The evaluators had a brief introduction at the office, where they explained that they wanted to meet the catchment committees" (i.e. farmers). Some stakeholders pointed out that these meetings were not satisfactory. There was no time to discuss various issues, which they considered important, with the evaluators. They also felt that they did not get the opportunity to explain why the work was beset by certain problems. One officer remarked that the meeting in his divisional office "was not a discussion where everybody had a chance to comment, it was only for facilitating the evaluation."

Table 7.4. Patterns of involvement in "National Soil and Water Conservation Programme in Kenya"

Stakeholder/ Activity	Financier	CP Branch	LI (Provinsial/ District level)	LI(Divisional level)	TG (farmers)
Preparing ToR					
Selecting evaluators					
Providing data					
Collecting data					
Interpretation/ Dev. of Recom-endations					
Commenting on draft/recom.					

Involved to a considerable extent	Involved to some extent	Involved to a limited extent	Not involved

Comments:
Financier = Sida,
Co-operating partner (CP) = The government agency which signs the contract with Sida,
Project management central (PMC) = The agency that is responsible for the management of the project,
Local implementor (LI) = A unit that works with the implementation of the project on the local level.
Target group (TG) = The beneficiaries of the support

The cases treated here give some important indications of how participation is related to ownership and use.

The cases Building Research Capacity and Curriculum Development show that:
* "taking part in" a process and influencing a process are two different things.
* even though ownership and influence can not be seen as the only prerequisite for use and usefulness, a feeling of real influence promotes ownership and use.

The HESAWA and NSWCP cases point at another issue which is directed more towards participation in the collection and interpretation of data and how use can be enhanced from the perspective of those working with implementation. These cases show that:
* one prerequisite for usefulness is that the stakeholders feel that their information needs are addressed in the evaluation
* participation in the collection and interpretation of data (through for example workshops) can make the evaluation process a learning process in itself, a theme which will be elaborated on more in the next section.

7.1.2 Focus on written rather than verbal communication

One would have thought that an evaluation becomes a rather meaningless affair unless those concerned and interested can access the analysis and findings. A fundamental part of evaluation ethics suggests that at least those who contribute to the evaluation should automatically have a copy of the report. Across the board, we have seen that this principle is sadly forgotten in aid evaluation. This is not only a matter of weaknesses in the distribution part of the evaluation system, but also reflects a lack of attention/understanding of what makes an evaluation effective. It also reflects a lack of respect of the needs, interests and aspirations of the various people who have a stake in a development project.

However, distributing the final report cannot be seen as the only, nor a sufficient, way of disseminating evaluation findings and recommendations. Even if the evaluation report had been accessible to the different stakeholders, it would have been quite pointless without workshops where the recommendations are discussed. Many of our cases have shown that the evaluation report is over-emphasised, at the expense of other ways of disseminating findings. Several stakeholders pointed out to us the lack of, and the need for, workshops where the results and recommendations could be discussed. One of the officers at a regional office in the HESAWA Programme (who did not participate in the workshop arranged) presented the importance of workshops and follow-ups in the following terms:

> This has been very common for evaluations in this Programme...sometimes we get a draft... we read it, pass it to another and put in on a shelf...after this we get a beautiful final copy but nobody says what to do with it.

This officer proposed that in big programmes like HESAWA, a number of workshops should be arranged at different levels of the programme so that everybody could have the chance to participate and discuss, not just a few as the case was at the workshop arranged in the 1996 evaluation.

A stakeholder in the evaluation of NNRCCI in Namibia voiced a similar sentiment:

> In Africa we like to talk to each other and discuss problems. We are not living in a culture where you write to communicate with each other.

It is therefore unfortunate that both evaluators of the NNRCCI left Namibia without having discussed their findings with all the stakeholders, not even all the major ones. The report, furthermore, once it was ready, was distributed by Sida only to a select group of readers. The First National Bank, to take one example, had been interviewed for the 1997 evaluation, but it never

received a copy of the report. This was particularly remarkable, as the Bank was severely criticised for its management of the credit fund.

7.1.3 The role of the evaluator

The quality of the evaluation, as well as that of the evaluator, cannot be judged only by the technical properties of, for example, the analysis and the recommendations. An assessment of quality must also include the evaluation process itself. The role of the evaluator in shaping utilisation is critical, since it is largely he/she (the evaluation profession tends, however, to be a male reserve) who decides how aspects of participation and ownership should be considered in the process. The case studies have given us ample material for reflecting on the role of the evaluator. Being evaluators ourselves, it has been a very useful exercise in self-reflection. Below we have put together some of the more important observations.

* The evaluator is a professional with considerable experience, but he/she rarely possesses any expertise in the theory and practice of evaluation. The evaluator knows a lot about a particular sector (water, road safety etc.), but much less about how to design and carry out methodologically more qualified evaluations. There is a standard evaluation procedure that is often followed. The framework is set by budget considerations, as well as the donors' own ideas of what is best practice.

* Few evaluators attach much importance to use and the factors that determine and condition use. Their responsibility ends with the completion of a report. Utilisation is the responsibility of the client who asked for the evaluation. This, however, does not prevent many evaluators from complaining bitterly about the non-utilisation of their work.

* It is not uncommon to launch evaluations that require a team of 2-5 persons, sometimes even more. Leading such an evaluation is an imposing management task. The team leader is by far the most important member of the team. He/she has a considerable say in shaping the design, work schedule etc. of the evaluation. Often he/she is also quite influential in preparing the analysis, conclusions and the recommendations.

* Many aid evaluations make a point of having local evaluators in the team. The primary reason is that they bring on board a knowledge of local conditions which few foreign evaluators can be expected to have, in addition to their presence often having been seen as an necessary ingredient for imparting some legitimacy to the evaluation. It is often the case in aid evaluations that the relationship between the local and foreign evaluators tends to be overlooked. Our cases prove no exception. There was evidence from many case studies that the marginalisation of the local evaluator is a reality. In the evaluation of the "Education sector support in Ethiopia", a local team member said that he did not participate in the final analysis, but was instead asked to prepare a paper on education policy issues in Ethiopia. The case is interesting because the team member had been appointed by the Ministry of Education to participate in the evaluation and assist with knowledge about the local context. Obviously, he was regarded by the other team members as a representative of the Ministry and hence not a "real" member of the team. Another example comes from the evaluation of the "Road Safety" project in Botswana. According to the report there was only one author. After some time in Botswana we understood that there had actually been two evaluators - the road safety expert and a local evaluator. Her name never appeared in the report. Yet she had been participating throughout the evaluation.

* There is a complex interaction between the evaluator and the client which cannot be described in simple client-server terms. Although the TOR of the evaluation, often formulated in neutral terms, stipulate what the client wants done, there are sometimes messages, intentions, views etc., which need to be communicated to the evaluator. The existence of such communication between the sponsor and initiator of the evaluation and the evaluator is particularly difficult in aid evaluation. It can easily disturb an already unbalanced relationship between the donor and various local stakeholders. The fact that it exists can possibly help to explain why so few evaluation contracts in Swedish aid are awarded in a process of open bidding (Carlsson, J., et.al. 1997)

7.2 Factors external to the evaluation process

In this section we shall discuss how factors external to the evaluation process have an impact on the utilisation pattern we have seen in the case studies. These factors are not at all related to how we conduct the evaluation in terms of method. They exercise their impact from the outside and often in a subtle and not very transparent way. We shall discuss two such factors. First, the evaluation relationship as a reflection of the aid relationship. Second, organisational constraints within the donor agency as well as recipient country organisations.

7.2.1 The evaluation relationship

Rossi and Freeman (1993) point out that any understanding of use must be situated within an analysis of the social ecology of the arena in which the evaluation is undertaken. The key determinants of utilisation are then to be found in the social and political contexts in which evaluations are undertaken. Let us explore this proposition a bit further.

Development aid involves a range of various stakeholders. But essentially it is a matter of two parties - the donor and the recipient or co-operating partner. The image promoted in various official documents is that of the aid relationship between donor and recipient as "friendly co-operation where two parties work together for a common goal - namely the development of one of the parties". (Saasa & Carlson, 1996) This representation is based on a consensus model, where conflicts are not a major disturbance. Essentially, everybody is working together for the common good. Related to this model of the aid relationship is the conceptualisation of the aid intervention process as a rather linear process with a kind of step-by-step progression from policy formulation, via implementation to outcomes.

The metaphor of an arena can also be used to describe the aid relationship. Each activity, which we will call a project, also constitutes an arena where there are many actors operating. Each has a stake in the outcome of the project. They are all concerned about the effectiveness of the project and they also bring different values, beliefs and attitudes into the project. The stakeholders also exercise power, some more than others. What is particularly distinctive about the aid relationship is that it is a relationship between two parties with quite different access to resources: "there is one who gives and one who receives". This imbalance is a fundamental property of aid and permeates almost all cases of development projects where aid is involved. This suggests that aid interventions tend to be less harmonious and conflict-free than the official picture would suggest. It is a process that is political by nature and shaped by different interests and power relations (Long & Villareal, 1993).

Aid evaluation has for a long time been based on a concept of knowledge which implies "discovering the facts", as if they were out there waiting to be uncovered. This is an approach that has its roots in the belief that the world is composed of objectively verifiable facts and that the goal of knowledge is to provide a literal account of what the world is like. (Knorr- Cetina, 1981) Our approach to evaluation is different. Just like the aid relationship, we see evaluation as a socially

constructed and negotiated process. The evaluation process is an arena for communication of such constructions. It is an arena where the parties in the aid relationship exchange their models for explaining and interpreting the environment, i.e. exchanging views on what is effective aid.

It is the adequacy of this arena with respect to such processes which is a key determinant of how the evaluation and its findings will be received, interpreted and finally used. The functioning of evaluation in terms of communication and learning cannot be understood purely in terms of formal systems for appraisal, monitoring and evaluation. Production and transformation of knowledge are situated in processes whereby social actors interact, negotiate and accommodate to each other's perceptions of the world. These processes and outcomes are shaped by sources of power, authority and legitimisation available to the different actors involved (Arce & Long, 1992:214). Given the power configurations inherent in the aid relationship, it is safe to assume that they are also replicated in the evaluation relationship This has contributed to a connotation of evaluation as an instrument for control and as an audit that assigns blame for mistakes committed (Lawrence, 1989). Such a view of the evaluation process is of course an obstacle for utilisation since it tends to make recipients of aid less interested in expending their time and energies on evaluation.

What does this imply for the question of evaluation use? If it is assumed that there can be different interests, and different strength/power to promote these interests, the issue of whose questions and information needs are included in the evaluation becomes important. The question of information needs becomes especially important, given the fact that in development aid the heart of the evaluation system is normally located within the donor agency.

How did the aid relationship manifest itself in the evaluation system as it was observed in our case studies?

In the production of an evaluation, the aid relationship appears primarily in the crucial, early phases of the evaluation. The initiative to undertake an evaluation almost always comes from within Sida. In none of the 30 cases studied in the first phase of this project had the recipient country taken the initiative (Carlsson, et al, 1997). The needs of the decision-making process within the donor agency, rather than any of the recipient organisations therefore determine the timing of the evaluation. The most common explanation for this stable pattern is that the recipient lacks the control and/or resources to take an initiative to evaluate. A local proposal to launch an evaluation needs to have the support of the donor, since the recipient often lacks money to go ahead on its own. Another explanation was given to us in the course of the interviews. Since the financier has the right, whenever deemed necessary, to evaluate/control how the money has been utilised, it is considered natural for the donor to take the initiative. This is in itself a good illustration of how difficult it is to design an aid relationship characterised by partnership and ownership.

The feeling of control is probably reinforced by the fact that the purpose of the evaluation is seldom made very clear. Evaluations are conducted because it is in the plan to do so or "we need an evaluation to be able to take a decision". They are often made routinely "because they have to be made." (Carlsson, et.al, 1997)

Consultations with the recipient do take place, but during our interviews we gained the distinct impression that the recipient organisations usually had very little to add. The terms of reference for the evaluation are almost always written in Stockholm. The only exception in our sample is the SAREC – ESTC evaluation in Ethiopia.

The recipient organisations lack not only money, but also knowledge of internationally reputable evaluators. It is difficult to challenge a proposition from Sida when you cannot propose an

alternative. In our sample we found only one case where the recipient organisation actively involved itself in the selection of the evaluation team. It was the Ethiopian Science and Technology Commission who had a long argument with SAREC about the composition of the team.

Sida very seldom recruits evaluators by open bidding. Programme officers tend to prefer persons they have previously worked with and whom they think have a good understanding of the subject matter of the evaluation. It goes without saying that this practice further undermines the idea of an evaluation as a joint exercise. Instead it strengthens the notion of the evaluation as mainly something of interest to Sida.

To conclude, the donor exercises a significant impact on the evaluation process. To some extent this may be a result of a limited understanding of what evaluation is all about - "they don't know any better". But more importantly, it shows how the unequal nature of the aid relationship is transmitted to the evaluation relationship. The evaluation itself becomes the subject of the same rivalries, promotion of interests' etc. as the aid relationship. Stakeholders try as actively as possible to influence the evaluation to suit their own purposes. Given the power configuration in the aid relationship, this often leaves the donor with the upper hand.

7.2.2 Donor agency and recipient organisational constraints

Within both donor and recipient organisations there are structures which actively determine the use of evaluations. One concerns the mechanisms in place for receiving an evaluation and disseminating its findings among the members of the organisation. Another factor is the basic conditions and terms for work in the organisations and the incentive structures they provide for the staff. We shall exemplify these points by using examples from two of our case studies - the evaluation of the water programme in Botswana and the evaluation of the support to the NGO in Namibia. There are some comparative advantages in choosing these examples. In Botswana we are dealing with key organisations in the public administration, the most important being the Department of Water Affairs (DWA) and the water departments of the district councils. In Namibia we are working with something completely different. The Northern Namibia Regional Chamber of Commerce and Industry (NNRCCI) is a small, poorly funded, membership-based, voluntary organisation.

The immediate impression of the water evaluation is that it was organised in such a way that it would allow for involvement and feedback to the concerned organisations. The reference group was composed by representatives from the most important stakeholders such as the DWA, the Ministry of Finance, Sida, the district councils etc. The group was supposed to function as a support to the evaluators. It was also expected that they should take care of feeding back information from the evaluation to their respective organisations. For various reasons this expectation was not fully realised.

Firstly, the reference group was not appointed primarily for its knowledge and experience from water issues. Instead they represented their respective government departments. They were there because they were involved, in one way or another, in the general planning of foreign aid and development projects. For many of them the evaluation was an opportunity to learn about the basics of providing water. Thus, their professional interest was limited to issues that specifically concerned their department's role in the overall government administration. The lessons they could be expected to bring back from the evaluation would therefore only concern a rather narrow part of the evaluation and its findings.

Second, there was a rapid turnover of staff within the Botswana public administration. "We were always being shifted around," as one member of the reference group put it. A few years after the evaluation had been completed, many members of the reference group had either left the government or been transferred to other government agencies at best remotely connected with the subject matter of the evaluation. The knowledge they gained from participating in the reference group was of little relevance to them in their new positions. A high turnover of staff also characterised Sida. None of the Sida representatives in the reference group work with water issues any longer. It was instructive to learn that among the present staff of the Swedish Embassy in Gaborone, nobody was very familiar with the evaluation, its findings and recommendations. It was explained to us that the programme had been terminated and the Embassy was now working with other issues.

The exception to this pattern was the DWA. The staff had been remarkably consistent over the years and they were clearly motivated in terms of their professional capacities and responsibilities. At the same time, a senior member of DWA explained to us that he regarded the evaluation as marginally interesting, except for the sustainability discussion which brought some important insights. The recommendations by the team were basically workable, although not new to him.

Thirdly, since using individuals to transmit the evaluation findings seemed to be a less viable option, one would have expected some institutional mechanisms to be in place for making the evaluation available to members of the organisation. Neither in Sida, nor in any of the Botswana organisations, were there any clear ideas of how the findings should be disseminated in the organisation and add to the institutional memory, rather than being confined to a few individuals. The typical dissemination pattern seems to be that the report is discussed at a joint meeting soon after the completion of the evaluation. After that very little occurs in terms of follow-up and further dissemination. An evaluation quickly becomes "yesterday's news". It is consumed here and now and is remembered by very few people at best. The reality of the organisations - new projects and new evaluations - soon makes an evaluation redundant and forgotten.

The evaluation of the NNRCCI exhibits another kind of organisational setting than that of the more traditional public sector organisation. NNRCCI is basically an NGO, a type of organisation with other strengths and weaknesses than the public, government financed organisation. It is the members of NNRCCI who determines whether it will continue to exist or not. There is a lot of flexibility, but also a lot of uncertainty in its operations.

The weak organisation of NNRCCI is a major factor in explaining how an evaluation can be used. Rather than being an important document for the new board, it became a document that was used by basically two persons - the advisor and the regional secretary. A pattern of use which may easily jeopardise the long-term sustainability of the support, and possibly also of NNRCCI.

The examples cited here show that utilisation of evaluations is closely linked with the settings of the organisations that are supposed to consume the evaluation. Generally there is very little thought given to the important issue of how an evaluation shall be disseminated in an organisation and. how organisational learning shall be brought about. There seems to be a rather naive belief in the value of the report as such. The existence of a written document seems to guarantee that it will automatically be read and absorbed by the members of the organisation. Evaluations can also easily become victims of internal power struggles between various stakeholders within organisations. The potential usefulness of an evaluation then runs a risk of being undermined. The lack of awareness of this situation may develop further and instead of assisting in solving conflicts, the use of the evaluation may lead to an escalation of an unproductive conflict.

8. The link between the evaluation process and use – conclusions

The evaluations analysed in this study have been used in various ways and for various purposes. What the experience from this study tells us is that it is rare for evaluations to be primarily put to instrumental use. Other types of use are much more common. Conceptual use is one example. The results of evaluations are supposed to trickle down into various decision-making processes and result in changes that are more concerned with strategies of doing things, rather than addressing more immediate operational concerns. Using evaluations for legitimising decisions already taken is common, as is the use of evaluations as ritual flag-waving.

The fact that other types of use than the conventional instrumental use are so dominant suggests that evaluations are not "pieces of objective evidence". Rather the evaluations seem to form part of already ongoing processes of organisational encounter. This is evidenced by the fact that we found few examples where the evaluations actually contributed something new in terms of knowledge. Neither the issues, nor the questions were new. They had been discussed before through other means and perhaps in other ways. The evaluation is just another means for the stakeholders to conduct their dialogue.

Evaluations occur in the context of various stakeholders' interests. They are by nature political. There are always different stakeholders with different interests who try to influence the evaluation in the direction of their interest. From this it follows that evaluations cannot be seen as neutral or objective or as representing some outside neutral perspective. In addition to this, the evaluators themselves have values and assumptions of their own which they bring to the evaluation. By choosing a particular evaluator you also choose a particular perspective, a preference for how things should be done and even the direction of the conclusions.

The fact that various stakeholders influence evaluations could be seen as problematic. Our conclusion is that this can never be avoided. It is in the nature of all evaluations in every kind of context. It becomes a problem when the evaluation arena is characterised by dominance arising out of dependency. Sida as donor packs a hefty clout in relation to most recipient governments and does not hesitate to use the evaluation function as a means of furthering its interests. The problem is not so much the lack of objectivity - evaluations can never be entirely neutral - as Sida's domination of the evaluation processes. It is not surprising that almost all of Sida's co-operation partners saw the evaluations as a way for Sida to convey its interests and concerns.

The evaluations analysed here are not good examples of partnership. There is not much openness and clarity about the way in which interests enter into an evaluation. Mutual trust is not what immediately comes to mind. In terms of adherence to the present policy framework of Swedish aid, the evaluation process as we know it is in dire need of reform.

Are the case studies examples of "good" use? Remember our definition of usefulness: *an evaluation is useful when stakeholders find something in the evaluation which they can put to use according to their own interest and needs.* The conclusion from this study is that evaluations are useful to a very limited group of stakeholders. For a majority of stakeholders the evaluation process could just as well have been left undone.

There are reasons for this attitude. The dominance of Sida constitutes a major obstacle to use. It limits the possibility of establishing concrete ownership of the evaluation on the part of the co-

operating partners. The evaluations are seen as a part of the agreement with Sida, and the process itself - in a word, as Sida's business. A serious application of a policy, which emphasises ownership, should enable the co-operation partner to "own" the utilisation and implementation of evaluation findings to a much greater extent. But in order to do this they must also have greater ownership of the whole process.

To assume that the co-operation partner should implement and "own" the implementation process (following an evaluation) without granting them an ownership over the rest of the evaluation process is quite unrealistic. A creation of ownership requires a change in the decision-making and influence over the evaluation process. It is not enough to "consult" the co-operation partner. Concrete and real use requires (i) that the co-operating partner have the right to choose evaluators, (ii) that they have a real influence on what issues are to be treated, (iii) that the co-operation partner take evaluation initiatives, (and of course that they have a say on whether an evaluation should be undertaken or not) and (iv) that the process becomes more transparent and open.

The experience from this evaluation has indicated the risk of evaluations becoming rituals. In order for evaluations to be useful there must be clear ideas of why an evaluation is undertaken, which issues that should be treated and how the evaluation should be used. This is one of the central prerequisites for use. The fact of the evaluation being included in an agreement does not guarantee that it will be used. This is an important issue from the viewpoint of utilisation. Evaluation purposes can be in conflict with each other. For example, evaluation for the purpose of establishing accountability is quite different from evaluation for the purpose of learning and management of projects. Different purposes may require different evaluation designs etc. Purposes of learning and management require an evaluation process which allows for involvement and effective participation of local stakeholders.

The stakeholders most frequently excluded from the possibility of influencing the use of evaluations are often those working with day-to-day implementation and the target groups. This is partly an ethical problem, if the stance is taken that everybody with an interest in and/or put at risk by an evaluation should also have a right to put their concerns and issues on the table. It is partly a problem of quality. The quality of evaluations will most probably, as witness the HESAWA evaluation, be enhanced if these groups are allowed to participate in the evaluation process. There are several reasons for this. Firstly, if these groups participate in identifying the issues to be addressed and in the gathering of data, the evaluation results tend to be more realistic and more closely attuned to the project/programme context. Secondly, if they are involved in this way, the findings tend to be more accurate, since participation tends to remove the tendency to hide and exclude problematic issues. Thirdly, if these groups are included in the evaluation process, both in the way outlined above and in final workshops, the process of "trickling down" directives can be shortened. Perhaps participatory evaluations like these are more time consuming. On the other hand, the very process of trickling down directives and evaluation recommendations will most probably be shortened. And the evaluation will tend to become a learning process in itself.

9. Recommendations for improving use

There is no doubt that it is necessary to increase the usefulness of evaluations made within Swedish development co-operation. At the moment they are the concern of a small group of stakeholders, leaving a majority of them, particularly in the developing countries, largely uninvolved in the full evaluation process. Increasing the effectiveness of the evaluation system as it currently exists should aim at giving many more the possibility to participate in the evaluation process, and hence, a chance to improve their capacity and the effectiveness of their work. Making the evaluation system more democratic is an important step towards increasing the effectiveness of aid.

The task of improving the evaluation function will require a combination of small, as well as larger measures. We shall here present a set of recommendations for, as we believe, necessary changes. They are primarily directed at Sida - the Department for Evaluation and Internal Audit as well as all other departments conducting evaluations to act upon. Some of them can be carried out by Sida alone, others will require collaboration with Sida's partners in development co-operation.

1. Current Swedish aid policy revolves around two basic concepts - ownership and partnership. At the same time, Sida's evaluation system is not characterised by any significant adherence to these guiding policy concepts. It is a rather undemocratic exercise with little concern for any other needs than Sida's own. The first step in changing the evaluation system is a revision of Sida's evaluation policy. It must be revised in order to more clearly reflect the fact that development aid, and aid evaluation, is a matter also for the recipient countries.

2. Sida is in need of a new evaluation manual. The point of departure for a manual should be the problem of utilisation. This will affect the way in which the different parts of the evaluation process are organised. It will also mean less emphasis on the report itself, and more on the process and on what will happen after the report has been produced. In essence, a new manual should try to give advice on how the "basic premises of use" are to be fulfilled.

3. Building evaluation capacity in the co-operating countries. Organisations involved in development aid should be given access to resources that will enable them to initiate and conduct evaluations independently of the donor. Such a fund should be located outside individual projects.

4. Evaluation results must be made more accessible. Today too many potential readers never gain access to an evaluation report. Improving distribution involves different small measures, which are not very difficult to implement as long as one is aware of their importance.

   * Plans for distributing an evaluation should be established as part of the ordinary planning process of an evaluation. Who shall have a copy? Who shall be responsible for the distribution? These are examples of questions that traditionally are never given much consideration, but can be easily included in evaluation planning.

   * The embassies in the recipient countries should be involved in distributing information about evaluations. Today it is very difficult for anybody in a recipient country interested in, for example, water issues or child health, to gain access to experiences from Swedish development co-operation. The embassies should, as part of their information services, become involved in dissemination of information about completed and ongoing evaluations, where they can be found etc.

* The western-based bureaucracy's need for written documentation should not overshadow the fact that there are other means of communication which are just as effective as a report. In several instances it is probably much more efficient to distribute the findings from an evaluation through verbal communication than through the traditional medium of the report. The usefulness of seminars, workshops etc. has been demonstrated in some of our cases, and they should be used much more consistently for ensuring the feedback of results to various stakeholder groups.

5. Our study has clearly demonstrated the value of participatory evaluation methods. Such methods have many advantages. They can counterbalance the tendencies towards inequality and dominance that are part of the aid relationship. They allow the evaluation process, and not necessarily the evaluation report, to become a source of learning. Sida and UTV should launch a consistent strategy towards introducing participatory evaluation methods. This is not only a matter of techniques, but to a large extent also a matter of changing attitudes within Sida to what evaluation is all about.

6. The evaluator has an important role for determining the nature of the evaluation process, and hence the usefulness of the evaluation. The experience from this study has demonstrated that, 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. Although a difficult thing for Sida to achieve, there are a some things that the agency can do:

   * This study has found that the recruitment of evaluators often follows the principle of the "old boy network". We believe that an increasing use of competitive bidding is a cost-effective way of increasing the quality of evaluators.

   * The purpose of the evaluation must be made much clearer than is the case today. A clear purpose - what, why and for whom - acts as an efficient guide to direct the work of the evaluator.

   * Sida's evaluation strategy should offer such clear guidelines that it becomes clear to an evaluator that the design of the evaluation process should focus on facilitating utilisation.

7. Representatives of the recipient country should become much more involved in the process of evaluating tenders for an evaluation. Sida is too dominant, which effectively contributes to the image of the evaluation as a "Sida evaluation" and not a joint effort.

Bibliography

Carlsson, J., K. Forss, K. Metell., L. Segnestam. & T. Strömberg. (1997) Using the Evaluation Tools. A survey of conventional wisdom and common practice at Sida. Stockholm: Sida UTV.

Forss, K. & Rebien, C.C. (1994). Listen to the People — a case study to describe participatory evaluation of a water and sanitation project. New York:UNDP/CEO.

Forss, K. (1994). Några aspekter på kvalitén hos utvärderingar av svenskt biståndssamarbete. Stockholm:SASDA.

Greene, J.G. (1988). "Stakeholder Participation and Utilization in Program Evaluation", Evaluation Review 12 (2): 91-116.

Guba, E.G. & Lincoln, Y. (1989). Fourth Generation Evaluation. California:Sage.

Lawrence, J.E.S. (1989). "Engaging recipients in development evaluation: the stakeholder approach", *Evaluation Review*, vol. 13, no. 3.

Lewin, E. (1993). Utvärderingshandbok för Sida. Planning secretariat. Stockholm: Sida.

Leviton, L.C. & E.F.X. Hughes. (1979) Utilization of Evaluations. Evanston, IL: Northwestern University Center for Health Services and Policy Research.

Leviton, L.C. & E.F.X. Hughes. (1981). "Research on the Utilization of Evaluations: A Review and Synthesis", *Evaluation Review* 5(4): 525-548.

Marsden, D., P.Oakley & B.Pratt. (1994). Measuring the Process: Guidelines for Evaluating Social Development. Oxford: Intrac Publications.

Rebien, C.C. (1994). Evaluating Development Assistance: the potentials of participatory evaluation, in *Forum for Development Studies*. No. 1-2, pp. 295-307.

Rebien, C.C. (1996). How to Evaluate Development Aid — A Theory and Methodology in the Making — with a special focus on participatory evaluation. Handelshöjskolen i Köbenhavn: Samfundslitteratur, Ph.D. serie 8.96.

Rich, R.F. (1977). "Uses of Social Science Information by Federal Bureaucrats", in C.H. Weiss (ed.) Using Social research for Public Policy making, pp. 199-211. Lexington, MA: D.C. Heath.

Rossi, P.H. & H.E. Freeman (1993). Evaluation — a systematic approach. London: Sage.

Patton, M.Q. (1986). Utilisation-focused evaluation. Beverly Hills: Sage

Rahnema, M. (1995). "Participation", in Sachs, W. (ed.) Development Dictionary. London: Zed Books.

Rudquist, A. & P. Woodford-Berger. (1996). Evaluation and Participation. Stockholm: Sida UTV.

Sida at Work - Sida's methods for development co-operation.(1997). Stockolm: Sida

Srinivasan, V. (1981). "The Methodology of Participatory Evaluation", in Fernandes, W. & R. Tandon (eds.) Participatory Research and Evaluation. Indian Social Institute: New Dehli.

Swantz, M-L. (1992). "Participatory Research and the Evaluation of the Effects of Aid on Women", in Berlage, L. & O. Stokke (1992). Evaluating Development Assistance: Approaches and Method. Frank Cass: London.

Vedung, E. (1995). "Utvärdering och de sex användningarna" in Rombach, B. & Sahlin-Andersson, K. (Eds.) Från Sanningssökande till styrmedel. Moderna utvärderingar i offentlig sektor. Stockholm: Nerenius & Santérus.

Weiss, C. (1972). Evaluation Research: Methods for Assessing Program Effectiveness. Englewood Cliffs, NJ: Pren tice-Hall.

Weiss, C.H. and M.J. Bucuvalas. (1980). "Truth Tests and Utility Tests: Decision-Makers Frames of reference for Social Science research", American Sociological Review 45 (April): 302-313.

Åberg, J-O. (1996) Det rationella och det legitima. En studie av utvärderingars teori och praktik. Monograph from the Department of Sociology. University of Gothenburg

Annex 1. Terms of Reference for the Assessment of the Impact of Evaluations Financed by Sida

1. Background

A survey of the use of the evaluation tool within Sida was performed in 1996.[1] The purpose of the survey was to illustrate how the evaluation process at Sida is conducted and how practice relates to Sida's evaluation policy. As a follow-up on this survey an assessment of the impact of a sample of the evaluations analysed in the survey will be carried out.

2. Purpose and Scope of the Evaluation

The purpose of this evaluation is to analyse the impact of evaluations on stakeholders through their involvment in the evaluation process and their use of evaluations and their results. The Consultants are also to assess to what extent there is a linkage between role and involvment in the evaluation process and use of evaluations and their results. The Consultants are to give recommendations on an evaluation approach and a design of evaluations that maximizes impact of evaluations and their results for stakeholders.

3. The Assignment

The Consultants are to analyse the issues listed below, starting with stakeholders within Sida, such as regional, sector and policy departments and continuing, as well as putting the emphasis on, partners in recipient countries holding a decision-making and/or implementing position for the project. An analysis for consultants responsible for the implementation of the project is also to be carried out. At least ten projects recently evaluated are chosen as case studies (see methodology section below).

Role and involvment in the evaluation process

The Consultants are to make an analysis for each project of stakeholders' roles and involvment in the evaluation process. As a starting point, the initiative and purpose of the evaluation is to be analysed from the different stakeholders' point of view. The Consultants are also to analyse whether there are differencies in how the different stakeholders view each others roles, involvment in the evaluation process and use of the evaluation and its results.

Use of evaluations and their results

The Consultants are to analyse the following questions for the projects and stakeholders:

- Did the evaluation design and questions address the needs and priorities of all stakeholders?
- To whom and how were the evaluation results distributed (report, other channels of dissemination of results)?
- How did the stakeholders find out about evaluation results (e g report, seminar)?
- What parts of the evaluation report has been read by the different stakeholders?
- To whom are the recommendations and lessons in the report directed?
- Are the recommendations and lessons in the report designed in a way that makes them possible to act upon?
- Are the recommendations and lessons in the report equally useful for all stakeholders and are they acted on?

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.

- How was the actual use manifested in concrete action?
- What was the role of evaluation results in changes, if any, in policy; procedures/methods; and projects/programmes?
- For each case study, the Consultants are to discuss other project and evaluation characteristics that may have influence the use of evaluation results (type of project and evaluation, working methods, country characteristics). The possibilities to make this kind of analysis should be further discussed in the inception report.

Linkages between role and involvment in the evaluation process and use of evaluations and their results

The Consultants are to assess whether it is possible to discern a linkage between role and involvment in the evaluation process and use of evaluations and their results.

Recommendations

The Consultants are to give recommendations on an evaluation approach and a design of evaluations that maximizes impact of evaluations and their results for stakeholders.

4. Methodology, Evaluation Team and Time Schedule

The Consultants will start with a preparatory phase before embarking on the main phase. The preparatory phase will lead to an inception report, with a methodological approach and an outline for the collection and analysis of data, as well as a budget for the main phase. The inception report is to include a presentation and discussion on which reports the Consultants have selected for the case studies.

The starting point for the evaluation is the 30 evaluation reports that formed the basis for the study previously mentioned: "Using the evaluation tool". Out of these 30 reports the Consultants are to make a representative sample of at least 10 reports that may be used as case studies. The case studies are to include both reports where the partners in recipient countries have been involved in the evaluation process and reports where they have not been involved in the process. If possible, the case studies are to represent different sectors of support and working methods.

The assignment is to be carried out by Andante Tools for Quest, Jerker Carlsson and Maria Eriksson-Baaz.

The evaluation is to begin in September 1997. An inception report will be presented no later than 30 September 1997. A draft report is to be presented no later than 7 August 1998.

5. Reporting

The preparatory phase will lead to an *inception report*. The *evaluation report* shall be written in English and should not exceed 75 pages, excluding annexes. The outline of the report shall follow Sida Evaluation Report - a Standardised Format (see Annex 3, p 71 of Evaluation Manual for SIDA). Three copies of the draft report shall be submitted to Sida no later than 31 January 1998. Within two weeks after receiving Sida's comments on the draft report, a final version in one copy and on diskette shall be submitted to Sida. Subject to decision by Sida, the report will be published and distributed as a publication within the Sida Evaluations series. The evaluation report shall be written in Word 6.0 for Windows or a compatible format and should be presented in a way that enables publication without further editing.

The evaluation assignment includes production of a summary according to the guidelines for Sida Evaluations Newsletter (Annex 1) and the completion of Sida Evaluations Data Work Sheet (Annex 2). The separate summary and a completed Data Work Sheet shall be submitted to Sida along with the final report.

Annex 2. Interview guide

The purpose of this short guide is, first of all, to clarify the approach to interviews as a research method followed in this project. Secondly, I shall also outline the practical approach which we shall use when conducting interviews in the field.

1. Interviews as a method

In one way or another evaluation researchers usually end up to talking to people in order to get their data. Evaluators will therefore, sooner or later, have to confront questions such as "how to do the asking?" and "who to ask". Inevitably the answer to these question will arise methodological disputes, so common in the social sciences, that you might get the impression that the different methodological camps are locked in a position of mutual hostility. The issue of interviewing techniques is no different from the general tendency in the social sciences. There are distinct camps, each defending with vigour its position. As we shall see the barriers between are not necessarily unbridgeable.

With respect to interviewing strategy, there are basically two camps - those favouring "structured" interviews and those favouring more "unstructured" approaches. The first follows a more formal style. The researcher is bound to a fixed set of questions, with fixed wording and subjects are bound to closed, predetermined answer alternatives. The rational of the formal, structured interview is to provide a simple, neutral "stimulus", in the form of questions, in order to tap the true responses or true values of individual interviewees. The usage of identical questions and set responses with all subjects is said to allow for proper comparison to be made across the entire field of potential viewpoints. The researcher begins with a theory of the information required from the subject; these are then operationalised into set questions and set response categories; the respondents answer by saying which of the categories applies to them; and finally responses are analysed to gain an overall picture of the studied population.[1]

The second approach follows a more conversational style. The researcher works with a broad list of themes to be explored and subjects frame answers according to their understanding of these issues. This implies a critique of the structured approach where the researchers conceptual system is imposed on the flow of information. Set questions and predetermined response categories offer little opportunity for the interviewee to question, or even understand, the evaluators ideas and intentions, with the result that the subjects own ideas may be misinterpreted.

In the unstructured interview information flows differently. Data collection has the task of creating a conversational framework in which the respondents frame of reference is properly accounted for and respected. The evaluator offers minimal steerage of the topics within broad areas of discussion. The idea is that a mutual understanding emerges via the in-depth exchange of ideas. The evaluator then selects out for his report those extracts from the total dialogue which most thoroughly crystallise the perspective of the subject. Analysis then consists of a descriptive narrative of the key world views. Critics of this approach stress that the information collected in such a way is diverse, and discursive and thus hard to compare from respondent to respondent. Researchers are accused of selecting the pieces of information that best fits into their own preferred explanatory framework.

The selection of an interviewing strategy is then often seen as a choice between standardisation or sensitivity, enumeration versus emancipation, anonymity versus ardour etc. In evaluation practice

1 1 An interview involves two persons; one who asks the questions and one who answers them. In this paper the first is called either the researcher, the evaluator and/or the interviewer. The second is referred to as the subject, or the stake holder or the interviewee.

the inclination towards structured interviewing often runs parallel with a preference for measuring outcomes, which often presupposes an experimental orientation, which in turn is more geared towards meeting the concerns of program managers and policy makers. Researchers who prefer to understand process tend to use the unstructured interview. This often presupposes some kind of constructivist research strategy. Such a strategy is more open to the concerns of practitioners and subjects.

This polarity is to a large extent an artificial construction. In practice, evaluators tend to be pragmatic and adopt a combination of interview approaches (Stern, 1995). *".....is that pragmatic, get-your-hands-dirty researchers should have no truck with the supposed polarities, since in real research it is often sensible, indeed advantageous, to work with a combination of methods"* (Pawson & Tilley, 1996, p.154) This would mean using both structured and unstructured interviewing, or a combination of the two - the semi-structured interview - depending on the actual situation. Common to them all is that they are data-driven strategies. The task of data construction is to ascertain information which is as true as possible to the thoughts and experiences of the person being interviewed.

This means that I believe that using a particular data collection strategy does not commit the evaluator wholeheartedly to a particular explanatory package. It is quite possible, and feasible to follow a more pluralist and pragmatic route where different approaches and purposes are combined. But pluralism for the sake of pluralism is not very valuable. It is not sufficient to offer a pragmatic suggestion that methods can be "combined". Combination can mean anything. The combination of qualitative and quantitative data, structured and unstructured qualitative data etc. should offer something more than just weighty evidence. It should also offer some explanatory value. This leads further to the need for theory. Theory which is needed to guide the combination of methods to obtain an explanatory value. And methods which help us developing theory.

Interview strategies could also be theory-driven and not only data driven . The researcher's theory is then the subject matter of the interview, and the subject (the stake holder) is there to confirm, to falsify and, above all to refine that theory. This comes quite close to the purpose of our study of utilisation of aid evaluations. It is important for us to understand how a stake holder has participated in an evaluation, what he feels about it etc. Given the state of knowledge about how evaluations are being used, empirical data of this kind is important. Equally important, however, is to understand what are the mechanisms that shapes certain utilisation patterns.[2] In other words, our study has set itself certain theoretical ambitions in that it does not only attempt to describe a situation, but also explain why it look the way it does. Such an ambition requires theory and, as said above, one purpose of the interview is to help in building such a theory.

2. Who knows what - stake holders in aid evaluation

Programs and project are complex social organisations. They have therefore a division of labour, which suggests a division of expertise between the actors on the project arena. Each of them can be expected to know something about the project, which none of the others are familiar with or have thought about from that particular perspective. The task of the interview strategy is to tap this combined knowledge. In a project evaluation it is possible to distinguish between four general types of actors (Pawson & Tilly, 1996). There are the *practitioners, the subjects, the evaluators and the policy*

[2] 2 Aid in general builds on the assumption that it is a partnership where the donor and the recipient enter into an agreement to do something together. Evidence from a range of studies shows, however, that this ideal is far from being realised. The aid relationship is too often beset by dominance by the donor. This is problematic since it might well lead to reduced aid effectiveness. The same problem is encountered by aid evaluation. Too often it is an activity led by the donor. A pattern which one could assume is equally detrimental to high aid effectiveness. It is therefore an important task to understand what factors shape the aid relationship, in aid programs as well as in aid evaluation, in order to reform it and make it functional. This can not be done without building theory.

makers. Each category can of course be further subdivided into more specific types. Furthermore, in aid each category can be placed in two distinct groups - the local and the foreign. At this juncture it is not necessary to elaborate on all possible types and their characteristics. It would probably only complicate the matter further. As a first approximation I shall try to outline their major distinctions in terms of what aspects of the project they are the most interested in.

a. Subjects (the more familiar term in development aid is the target group) are likely to be more sensitised to the mechanisms in operation within a project, than they are in relation to its contextual constraints and outcome patterns. Project mechanism provide the reason and resources which encourage and provide some benefits to the subject which will enable them to change the situation they are in. Subjects normally know fully well whether the project has been successful in providing such resources. The subjects are likely to have a rather personal view of choices made and capacities changed, and are therefore not really able to talk about other subjects experiences. Subjects are not very knowledgeable or sensitive to the influence of the context. The reason is that they encounter the project in a setting which is, for them, entirely routine. Finally, a representative of the target group can not be expected to have a full understanding of the total outcome pattern of a project. The experience only a limited part of it.

b. Practitioners are those who translate the ideas and intentions of a project into practice. In many evaluation situations they are the ones being evaluated. This make them very sensitive to the information provided by an evaluation and its utility. Their experience is often operational, they know what project mechanisms that work. But we can not expect them to have a systematic idea about what works and for whom under what circumstances. Their understanding of the project is based on the kind of working relationships they establish during the course of the project. Experienced practitioners work best by tailoring their efforts closely to the current personalities involved without necessarily anticipating the needs of the potential target population as a whole. There are a range of other factors which prevent practitioners to abstract and to typify and generalise their understanding of programs.

c. Evaluators enter the project equipped with another pool of knowledge. They carry experiences from other projects with them, and other theories about what works and what does not work. Their approach is often based on analysing context, mechanism and outcome.

They establish their hypotheses on the basis of previous experience, as well as the propositions of social science theory. The evaluators wisdom is probably stronger on form than content and hence lack the local detail that practitioners and subjects are so good at providing.

d. Policy makers are important. It is up to them to agree to or initiate a project. It is normally policy makers who commission or agree that an evaluation shall be undertaken. Policy makers are also often the prime recipients of an evaluation report. They are often familiar with the project, but do not know all the details - they see much but they are not all-seeing. As the other types they are locked in their own expertise.

On teaching and learning:

Evaluator: Do you understand the general thrust of the evaluation and what we are trying to investigate?

Respondent: Yes, I understand the general theoretical ground you are exploring, this makes your concepts clear to me, and applying them to me gives the following answer.....

The division of expertise among major actors sets a clear task for data construction. It is to enable a cross-fertilisation between these different interpretative currents. Each of these stake holders have something to teach the others, and something to learn from them as well. This suggests that the kind of interview to be conducted in this study of USE follows a teacher-learner pattern. In the next section I shall discuss in some detail how such an interview situation can be constructed.

3. The interview

The evaluators task is to get an idea of the actions and beliefs of each stake holder within a wider framework of their causes and consequences. There is always a fringe of incompleteness in self-understanding, which means that a stake holder can not be expected to have a full picture of the context and outcomes of a project. These gaps in self-understanding is the task of the evaluator to fill during an interview session. There are two features of the interview situation which are important in shaping an interview situation which is characterised by dialogue and interaction, rather than the orthodox " researcher ask question, subject answer them" type of interview.

First, we have the "teacher - learner function. The issue here is to consider the nature of flow of understanding between interviewer and interviewee, or, to put it differently, how can we know that the subject is attending to the researchers understanding of the problem? In the structured interview this is solved by relying on precision in question wording and clarity in operationalisation. Questions should be clear, precise, unambiguous and not leading. My point here is that precision and clarity alone is not sufficient to orient the subject to the underlying research task. The evaluator knows full well the nature of the hypotheses which are embedded in a research task. The interviewee, on the other hand, can remain blithely unaware. What I am suggesting here is that the evaluator should be much more active in teaching the overall conceptual structure, or the general purpose, of the evaluation to the subject. Respondents always want to know more than what is conveyed in the formal structure of questions. They often think to themselves: "who is he", "what is he after", "why am I being asked" etc. The objective of the "teacher- learner" approach is to put the subject in a position where these hidden questions are answered.

The second feature of the interview situation is the conceptual refinement function. In a conventional sense any interview is trying to encourage the respondent to explicate their own understanding of a situation. This is the pint in time when the respondents will have their say about what is being evaluated. The point about conceptual refinement is that respondents will deliver their thoughts in the context of the evaluation. The overall structure of the questions will, generally speaking, mark out the areas in which the subject will make decisions. The subjects task is to agree, disagree and categorise themselves in relation to the preferences and attitudes as construed in the questions, but also to refine their conceptual base. The respondent is asked for attitudes. He is assisted in explaining them by, first, being offered a formal description of their own thinking, followed by an opportunity to explain and clarify that thinking.

This approach can perhaps be made more clear by confronting it with some weaknesses of standard interviewing tools. The standard way of measuring attitudes is usually by using indicators along a line from "strongly agree" to "strongly disagreeing" in relation to statements such as "I think that the recommendations made by the evaluation has helped me in my job as project manager". Such questions often have respondents wishing for an opportunity to place their answer in a context. There is, for example, a need to explain the project setting to be able to understand why the recommendations have been useful to the respondent. Other expressions of the need for context are questions such as: "am I along the right lines here?", "you wouldn't want to know about that, would you?" or "is this the kind of stuff you want?" The basic idea of conceptual refinement is to avoid such difficulties by carefully contextualising the domain in which the subject reflect on their own thinking. Only then can a mutual understanding emerge.

Now the teacher-learner- and conceptual refinement function of interviewing leads us to designing an interview strategy which combines features of the structured and unstructured interview technique. The interview strategy which we will follow is not only data driven, but is also theory driven, i.e. we have an ambition to explain why a certain pattern of use actually exist.

The point of departure for our interview design is that we have a basic idea of how evaluations are being used by different stake holders. The task of the interview is to refine that idea by obtaining as much detail as possible from the participants on how they relate to the evaluation and its findings and recommendations. The first tool in our tool box follows the format for a structured interview. It has all the outward resemblance of a questionnaire and can be described as a sort of attitude rating questionnaire to which the respondent can agree or disagree (in reality is more of a discussion document). The items on the questionnaire represent the current stage of the evaluators ideas and perceptions of he general research problem - the utilisation of evaluations. The immediate task of the interview is to "teach" this conceptual structure. This is done by making it clear to the respondent that they are being asked to reflect upon and respond to a series of statements which *other*s have made about the utilisation of evaluations.

Refining the context:

Respondent: This is how you have depicted the potential structure of my thinking, but in my experience of those circumstance, it happened like this.....

The questionnaire developed for our study is found on the next page. The stake holders are presented with a list of statements about the evaluation process and the use of the evaluation reports findings and recommendations.

Using evaluation results

Please consider each of the statements and assess to what extent they apply to your experience from the way the evaluation was conducted, as well as from using the findings and recommendations of the evaluation. Each of the statements below applies to me..

to a considerable extent	1
to a moderate extent	2
to a slight extent	3
not at all	

Response rate

In the evaluation process, I was……:

1 among the ones who initiated the evaluation.
2. against making an evaluation.
3. reading and commenting carefully on the terms of reference for the evaluation.
4. participating in selecting the evaluating team.
5. meeting several times with the evaluators.
6. providing them with extensive information

Regarding the results of the evaluation, I ……..

7. have never seen the final evaluation.
8. found that the evaluators did not use my information very much.
9. think that the results can not be used for practical purposes.
10. think that the findings were more important for the donor, than for me and my organisation.
11. used the recommendations to improve my work and that of my organisation
12. feel that the recommendations enabled me to improve my participation in the project
13. used the evaluation to improve my use of the services produced by the project.
14. think that the evaluation helped us to solve conflicts in the project.
15. think that in general the evaluation was used to improve the management of the project in order to become more efficient
16. think that in general the evaluation was used to change the orientation of the project.
17. think that in general the evaluation was used for learning more about the impact from the project.

What we have at this point is a piece of attitudinal scaling which may irritate everybody. The analyst with a quantitative instinct would go for more formal analysis. The qualitatively minded analyst on the other hand, would disregard the lot as useless number-crunching and instead emphasise more personal involvement (participation) to get the real meat out of the respondents. For the respondents this kind of exercises can be very irritating simplifications of their rich experience from a project.

These questions are neither specific stimuli to obtain a response, neither are they an invitation to the respondent to enter into a discourse about everything he feels is important. These formal questions set a clear agenda which represents a body of theory, offering up the evaluators potential explanations to the use of evaluation. The key role of these questions should therefore be to involve the respondent in a clearer articulation and clarification of the hypothesis and theories put forward by the evaluator.

In reality these questions are more important than their somewhat superficial character may suggest. They lead further to an in-depth conversation with the stake holder on the evaluation he has experienced. In very practical terms the typical interview would take the following course:

The key feature of this interview strategy is the creation of a situation in which the theoretical postulates and conceptual structures under investigation are open for inspection in a way that allows the respondent to make an informed and critical contribution to them. The interview follows a teacher-learner pattern. The interviewer teaches the subject - the stake holder - what the interview is designed to find out. The subject then teaches the interviewer about how evaluations were actually used as seen from his perspective.

The interview procedure

We can only assume that there will be only one meeting between the evaluator and the subject.

When the evaluator meets the interviewee for the first time in order to conduct the interview, the first thing he/she does is to introduce the questionnaire.

The interviewee then fills it in, in the presence of the evaluator. This should take around 15 minutes.

The evaluator then goes through the answers with the interviewee. This will enable the both of them to fully penetrate the answers and what they really mean.

Annex 3. Summary of the 9 evaluations chosen as case studies

Country Criteria	Ethiopia	Ethiopia	Ethiopia	Ethiopia	Kenya	Tanzania	Botswana	Botswana	Namibia
Sector	Education	Education	Research	Health	Soil Conservation	Health	Infrastructure	Health	Industry
Purpose of support	Production of textbooks for primary schools	Development of curricula for primary schools	Research support to the Ethiopian research council, and research projects	Water supply through gravity-systems in villages Training of villagers	Soild conservation	Improved water supply, sanitation and health through better techniques and knowledge	Improved road safety through driver training, vehicle control and accident analysis	Water supply to 354 villages. Training of villagers	Technical and management advice to the Northern Namibia Regional Chamber of Commerce and Industry
Project period	1992-1996	1992-1996	1984 - 1995	1982-1986	1994-1996	1985 - 1998	1985-1992	1971 - 1993	1993 - 1996
Place	Addis Ababa	Addis Ababa	Addis Ababa	Dodota, Arsi and Orominia	National	Kagera, Mara and Mwanza	National	National	Northern Namibia
Responsible for implementation	Educational Materials Production & Distribution Agency (EMPDA)	Institute for Curriculum Development and Research (ICDR)	Ethiopian Science and Technology Commission Addis Ababa University	ARDU Orominia Water, Mineral and Energy Resource Development Bureau	NSWCP PDA och PSCO Farmers	Ministry of Community Development, Women Afffairs and Children. Distrikt councils and villagers	Department of National Transport and Communication (DNTC)	Dept of Water Affairs och District councils through Rural village water supply programme	Northern Namibia Regional Chamber of Commerce and Industry
Target group	EMPDA, teachers and students	ICDR, teachers and students	Research departments and researchers in the projects and researchers in the projects	Inhabitants of Dodota district	Farmers	Villagers	DNTC	District councils Villagers	Small businesses
Evaluation	Valdelin, J. et.al. Sida Support to the Education Sector in Ethiopia. Sida Evaluation 96/27	Palme, M. et.al. Curriculum Development in Ethiopia. Sida Evaluation 96/40	Thulstrup, E. et.al. Building research capacity in Ethiopia. Sida Evaluation 96/9	Olsson, B. et.al. Water supply system in Dodota, Ethiopia. Sida Evaluation 96/23	Tiffen, M. National Soil and Water Conservation Programme. Sida Evaluation 96/25	Smet, J. et.al. HESAWA, Health through sanitation and water. Sida Evaluation 97/12	Davey, R. Botswana Road Safety Improvement project. Sida Evaluation 96/43	Valdelin, J. Rural village water supply programme - Botswana Sida Evaluation 96/10	Sahlin, Å. Evaluation of the Small Business Development Programme in Northern Namibia. January 1996

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

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