

Secretariat for Analysis of
Swedish Development Assistance

Evaluations in Multilateral Development Cooperation

Kim Forss
Stockholm, Juni 1994

SASDA

The Secretariat for Analysis of Swedish Development Assistance

The Swedish government has appointed a committee with the task of analysing the results and effectiveness of Swedish development aid. A special Secretariat, SASDA, was set up on 1 March 1993 to carry out the work.

The Secretariat will work until the end of 1994 and will have as its main task to propose to Government suitable mechanisms for evaluations and policy analyses of Swedish aid. In its work SASDA will give priority to carrying out a set of selected studies world-wide, at country, sector and subject level and to studies of individual organisations to provide a basis for decisions on development co-operation in the future and to gain experience on how policy evaluations should be carried out. A major study concerns Sweden's co-operation with Central and Eastern Europe.

SASDA's point of departure is the aim of a better understanding of the mechanisms of development in order to enhance the results and increase the effectiveness of aid in achieving the five goals set by the Swedish parliament: increased resources, economic and social equality, economic and political independence, the democratic development of society, and the long-term management of natural resources and care of the environment.

The studies and analyses will be managed partly by the Secretariat's own staff and will include studies commissioned from different specialists in the committee's areas of priority.

The staff are :

Ambassador Claes Sandgren	Head of the Secretariat
Dr Olle Edqvist	Senior Policy Analyst
Mr Enrique Ganuza	Senior Policy Analyst
Mr Per Johan Svenningsson	Senior Policy Analyst
Ms Kerstin Sandling	Assistant Secretary

Postal address: SASDA
P.O. Box 16418
S-103 27 STOCKHOLM

Telephone: +46-8 791 21 80
Telefax: +46-8-791 21 88
Visiting address: Klarabergsgatan 23, Stockholm

Evaluations in multilateral development cooperation.

A study commissioned by the Swedish Secretariat for Analysis of Development Assistance

by
Andante Consultants AB
Kim Forss

1. Introduction

1.1 Purpose

The Secretariat for Analysis of Swedish Development Assistance (SASDA) has started a review of evaluation activities in the Swedish aid organizations. The review covers SIDA, BITS, SAREC, SwedeCorp and Swedfund, and the multilateral department of the Ministry for Foreign Affairs. This paper forms part of that review, and it treats the evaluations of the Ministry for Foreign Affairs; that is, the evaluations pertaining to the Swedish contributions to multilateral organizations. The terms of reference are found in Annex 1.

The purpose of this paper is threefold. My first aim is to describe the evaluation information coming to the department; that is, how much is it, where does it come from, and is it of any use? The sheer quantity of information is interesting in itself; if there is a scarcity of information the most important problem is to search for more and that has budgetary consequences. However, if there is no shortage of information, the problem may be one scanning, selecting and aggregating. This too has budgetary consequences, but may reveal opportunities to save money. The quantity and quality of evaluations are described in chapter 2.

The second aim is to discuss means of asserting the effectiveness of organizations. The multilateral organizations are complex systems and they cannot only be understood by looking at the effects of the development cooperation they are engaged in. In addition, if a study of effects indicate that one of the organizations has little impact, the political choice will not necessarily be to stop supporting that organization. Another choice might be to seek to reform the organization to get better results in the long run. However, if the latter option is to be realistic, the department needs tools for thought on organizational issues. Chapter 3 is a brief review of some ingredients in organizational analysis.

Even if this review only presents a brief indication of the contents of organizational analysis, the next step of my analysis will be to compare the "shopping list" of tools for thought to the quantity and quality of information at hand. The comparison is presented in the concluding analysis, in chapter 4. I should add that the analysis of gaps in the information flow is not only a theoretical review but also builds on observations by people in the organization.

1.2 Definitions

The title of this paper immediately suggests the objects that must be defined; multilateral development cooperation and evaluation. To start with the former, I write about the work being done by the department for multilateral development cooperation. Consequently it is information about the object of their decisions and actions that are of

interest. To make it easy, these are decisions and activities regarding the funds that the government allocates for development cooperation under the heading C1 in the budget. These funds are used to support the economic and social activities of the UN system, the international financing organizations, the UN activities for peace and reconstruction, and some other institutions. During 1993/94, the government allocated 3.148 million Swedish crowns under this heading, and has proposed 3.295 million during the coming budget year. I look for evaluations that tells something about the use of this money.

There are two things in addition to be said about the multilateral development cooperation that concern how evaluation instruments are used. The first is that the Swedish contributions to the UN agencies and the other multilateral institutions also come from other organizations. SIDA and the other aid agencies contribute directly to the multilateral organizations and sometimes cooperate closely with them in the Third World and Eastern Europe. The department at the Ministry for Foreign Affairs works closely together with the other agencies and with the Ministry of Finance. The total Swedish support to the multilateral organizations amounts to a third of the budget for development cooperation, or around 4.600 million crowns.

The second fact is that the Swedish contributions are concentrated to a handful of organizations. The government's budget mentions twentyfive. Some more may receive support for program or project activities directly from the aid agencies. The Swedish government channels funds to UNDP, UNICEF, UNFPA, UNRWA, WFP, UNHCR and the IDA to mention the destination of the largest sums. Other organizations are not mentioned in the budget, for example the FAO, UNIDO, UNESCO and ILO. Both past and future decisions need(ed) evaluation information on the organizations that do not receive funds today.

Definitions are one of those things that are hard to live with and hard to live without. It is obviously necessary to define terms in order to communicate. It is also obvious that many practical tools as well as ideas do not fit neatly into narrow definitional pigeonholes. In my view it is neither respectful nor practical to adopt a single definition of evaluation and then work at making the situations one encounters fit that definition. In the following I try to recognize the variations and the complexity within the field of evaluation: the word evaluation information is often used as a synonym to evaluation in this text. It signifies information on the value of something, in this case the value of development cooperation through the multilateral organizations. This may be past value, present value or future value, and evaluation information in this sense includes policy analysis.

Such a broad definition of evaluation differs from common practice in development cooperation which has been established by the OECD (1986). However, I think it is more important be responsive to new situations than to stick to formal definitions. Above all, evaluation is an activity which has a mandate to be practical, and practicality also means having sufficient flexibility to understand which definition(s) are

appropriate and meaningful in a particular context, with a specific group of people.

1.3 Methods.

The first step in this review was to peruse evaluations used by the department. The desk officers were informed about my assignment and were asked to provide me with evaluations they have come across. This gave 15 evaluations, totalling some 2.000 pages (listed in annex 2). These evaluations were supplemented by a review of incoming correspondence, and I was helped to classify and structure this and single out documents with an evaluative content.

The evaluation reports were then analysed with the format established by SASDA in mind, though it was necessary to add other indicators of quality. Towards the end of this phase, I held a meeting with staff members of the department, presenting my analysis of the supply of evaluations and a tentative assessment of how this matches their information needs.

In addition, I have interviewed eight people from the department. Each interview was centered around open questions of how many evaluations the interviewee normally reads, what the interviewee thinks of their quality, and what type of information is needed to supplement the evaluations (if any). The open ended questions often lead to a discussion of the nature of work at the department, policy issues regarding multilateral assistance, and the substance of the work done by the multilateral organizations.

These rather unsophisticated methods provide the empirical data which is used in the review of evaluation information in chapter 2, and they make an important input to the analysis of information shortcomings in chapter 4. The section on organizational analysis builds on literature and experience; it has the character of reflection on practice rather than empirically based research.

2. The quantity and quality of evaluation information

2.1 Sources of evaluation.

In this section of the report I analyze the information at hand in the department for multilateral development cooperation at the Ministry for Foreign Affairs. The term evaluation is used in a broad sense to cover analytical written reports on the policies, performance and effects of the multilateral organizations. Information comes from three different sources.

Many studies are undertaken by the multilateral organizations themselves, as part of the ordinary evaluation activities or as separately

commissioned tasks. The World Bank has an evaluation unit which is well reputed. Much of the knowledge concerning impact and effectiveness of development assistance comes from the studies undertaken by the bank's evaluation unit, for example the Cassen report of 1986 and the Wapenhans report of 1993. The regional development banks also have evaluation units, although their contribution to the research frontier is less prominent. Among the UN agencies, it is primarily the UNDP, UNICEF and UNFPA who have evaluation units of significance and general interest. However, most of the others also have organized evaluation activities, although their results are less widespread.

Other studies are made by independent research organizations, on their own initiative, or on commission from donor countries. All countries who support the development activities of the UN system or the multilateral development banks are interested in their effectiveness. All donors face the same dilemma of knowing where to channel money for the best effects, and according to their own national priorities. Consequently all at some time or the other undertake evaluations of the multilateral organizations.

The Nordic countries often cooperate in evaluations, but some studies are undertaken separately. For example, Denmark commissioned a major study of the effectiveness of 11 multilateral agencies in 1991. Norway, together with Canada and the Netherlands commissioned a study of WFP which was completed in 1993. Other major donors such as Japan, the U.S.A, Germany and France also seek to make as rational decisions as possible and they try to base decisions on written analytical information. The studies that are commissioned often circulate among the donors, but it appears as if the "like-minded" nations share more information with each other than with others. I have not found any references to German, French or Japanese donor evaluations, but I would be very surprised if there were none.

The UN organizations receive attention from numerous independent researchers. Some publish doctoral dissertations that are studies of the effects, effectiveness, policies and procedures of the system. Williams (1987) study of the specialized agencies and the United Nations is an example of a study that still has considerable impact. My own study of planning and evaluation in UNDP (Forss, 1985) might have had some relevance in the mid-1980ies. The research team at the ODI and the University of Manchester have done several important studies of the World Bank and IMF and the problems of structural adjustment lending (Mosley et al; 1991). Books and reports such as these at times also reach the department and are read by people there.

Finally, the Ministry for Foreign Affairs commissions studies directly. The Nordic UN project lasted for three years and it led to several studies of effectiveness and efficiency. These and other studies have as their main objective to guide and formulate policy decisions. Information on effects and effectiveness is one ingredient in such decisions - but not the only one. Consequently, the studies contain evaluation information. The public commission on multilateral development

assistance (SOU 1990: 48) is another example of a study that could play a role in the evaluation of the department.

In conclusion, information comes from many sources. Aid evaluations often have a short life, but timing is less of an issue concerning many of the reports here. A study such as Williams's (1987) is as relevant today as it was 7 years ago, and Gordenkers study (1976) on international aid and national decisions is echoed in Berg (1993). Two factors contribute to give the evaluations from the sources above a long life; one is that the multilateral system itself is changing slowly, and the other is that the studies are often based on solid research undertaken over long periods of time.

2.2 The amount of evaluation information.

Let us now look at how much information is available. Even though sheer quantity is not an end in itself, it opens up opportunities for aggregated analyses and may indeed pose problems of a qualitative nature in information processing that are different from a situation of information scarcity. I discuss quantity in relation to each of the sources identified above.

Evaluation from the multilateral organizations

There is no shortage of information on multilateral development assistance. In 1993 5.860 documents came to the multilateral department from the joint Nordic office in Washington, which handles support to the World Bank group. Every week a short list of incoming titles is printed and circulated in the office and to other instances (such as the Ministry of Finance, the bilateral department, the aid agencies). Not all of these, but still quite a few, might be called evaluations.

Out of these 5.860 documents there were 160 staff appraisal reports. These reports were completed by bank staff who had been out to visit field activities. They assess project or programme performance against objectives and they usually contain a financial analysis. Another thirty reports were entitled evaluations; that is, more formal, external assessments. The list also encompassed 1 571 project documents. A project document is an ex ante evaluation of a development project. If it concerns an extension of a project, it will normally consist of a summary evaluation of preceeding activities.

In each of these reports there is information on objectives of cooperation, targets, achievements of targets, analyses of short-comings, recommendations concerning organization and implementation, economic analyses of various kinds, presentation of key indicators and many other questions that may be asked in respect of the banks' operations. Not least do many of the reports contain environmental impact statements.

commissioned tasks. The World Bank has an evaluation unit which is well reputed. Much of the knowledge concerning impact and effectiveness of development assistance comes from the studies undertaken by the bank's evaluation unit, for example the Cassen report of 1986 and the Wapenhans report of 1993. The regional development banks also have evaluation units, although their contribution to the research frontier is less prominent. Among the UN agencies, it is primarily the UNDP, UNICEF and UNFPA who have evaluation units of significance and general interest. However, most of the others also have organized evaluation activities, although their results are less widespread.

Other studies are made by independent research organizations, on their own initiative, or on commission from donor countries. All countries who support the development activities of the UN system or the multilateral development banks are interested in their effectiveness. All donors face the same dilemma of knowing where to channel money for the best effects, and according to their own national priorities. Consequently all at some time or the other undertake evaluations of the multilateral organizations.

The Nordic countries often cooperate in evaluations, but some studies are undertaken separately. For example, Denmark commissioned a major study of the effectiveness of 11 multilateral agencies in 1991. Norway, together with Canada and the Netherlands commissioned a study of WFP which was completed in 1993. Other major donors such as Japan, the U.S.A, Germany and France also seek to make as rational decisions as possible and they try to base decisions on written analytical information. The studies that are commissioned often circulate among the donors, but it appears as if the "like-minded" nations share more information with each other than with others. I have not found any references to German, French or Japanese donor evaluations, but I would be very surprised if there were none.

The UN organizations receive attention from numerous independent researchers. Some publish doctoral dissertations that are studies of the effects, effectiveness, policies and procedures of the system. Williams (1987) study of the specialized agencies and the United Nations is an example of a study that still has considerable impact. My own study of planning and evaluation in UNDP (Forss, 1985) might have had some relevance in the mid-1980ies. The research team at the ODI and the University of Manchester have done several important studies of the World Bank and IMF and the problems of structural adjustment lending (Mosley et al; 1991). Books and reports such as these at times also reach the department and are read by people there.

Finally, the Ministry for Foreign Affairs commissions studies directly. The Nordic UN project lasted for three years and it led to several studies of effectiveness and efficiency. These and other studies have as their main objective to guide and formulate policy decisions. Information on effects and effectiveness is one ingredient in such decisions - but not the only one. Consequently, the studies contain evaluation information. The public commission on multilateral development

assistance (SOU 1990: 48) is another example of a study that could play a role in the evaluation of the department.

In conclusion, information comes from many sources. Aid evaluations often have a short life, but timing is less of an issue concerning many of the reports here. A study such as Williams's (1987) is as relevant today as it was 7 years ago, and Gordenkers study (1976) on international aid and national decisions is echoed in Berg (1993). Two factors contribute to give the evaluations from the sources above a long life; one is that the multilateral system itself is changing slowly, and the other is that the studies are often based on solid research undertaken over long periods of time.

2.2 The amount of evaluation information.

Let us now look at how much information is available. Even though sheer quantity is not an end in itself, it opens up opportunities for aggregated analyses and may indeed pose problems of a qualitative nature in information processing that are different from a situation of information scarcity. I discuss quantity in relation to each of the sources identified above.

Evaluation from the multilateral organizations

There is no shortage of information on multilateral development assistance. In 1993 5.860 documents came to the multilateral department from the joint Nordic office in Washington, which handles support to the World Bank group. Every week a short list of incoming titles is printed and circulated in the office and to other instances (such as the Ministry of Finance, the bilateral department, the aid agencies). Not all of these, but still quite a few, might be called evaluations.

Out of these 5.860 documents there were 160 staff appraisal reports. These reports were completed by bank staff who had been out to visit field activities. They assess project or programme performance against objectives and they usually contain a financial analysis. Another thirty reports were entitled evaluations; that is, more formal, external assessments. The list also encompassed 1 571 project documents. A project document is an ex ante evaluation of a development project. If it concerns an extension of a project, it will normally consist of a summary evaluation of preceeding activities.

In each of these reports there is information on objectives of cooperation, targets, achievements of targets, analyses of short-comings, recommendations concerning organization and implementation, economic analyses of various kinds, presentation of key indicators and many other questions that may be asked in respect of the banks' operations. Not least do many of the reports contain environmental impact statements.

Obviously these reports differ widely, some are no more than twenty or thirty pages, representing a few weeks investigative work, and others are the result of intense research efforts and consist of several hundred pages. If the average length of these reports were 50 pages, the department would receive around 90 000 pages of evaluation information during a year from the World Bank group, the IMF and the regional development banks.

It is not that easy to estimate how much information that comes to the department from the other UN organizations. There is no separate registry. The documents enter into the archives of the Foreign Office and cannot be analyzed with the help of search words as can the incoming documents from the banks. Still, much less evaluation information comes from them. Not because there is a shortage of such information. Many of UN organizations have evaluation departments that are as productive as those of the banks. However, projects are rarely discussed by the donors and the organizations do not send out project documents, appraisal reports or even country or sector evaluations. The organizations do send out policy documents, country programme studies and various reports. A rough estimate would be that the UN organizations send less than a third of the amount coming from the banks; that is, around 30.000 pages per year.

Evaluation reports from other donors or independent research.

As part of the assignment, I have looked at evaluation reports commissioned by other donors. During the interviews I got reports that people use and that play a part in the assessment and decisions on the multilateral support at present. These reports covered the activities of UNICEF, UNFPA, WFP and UNDP. One of the reports covered the subject of technical assistance, and the analysis pertained to both bilateral and multilateral aid organizations. (The reports are listed in annex 2)

The evaluation of UNFPA was a joint study by BMZ, the German Federal Ministry for Economic Cooperation and Development, CIDA, the Canadian International Development Agency, and FinnIDA, the Finnish Ministry of Foreign Affairs, department for development cooperation. The evaluation of UNICEF was commissioned by the Australian, Canadian, Danish and Swiss donor community.

Evaluations such as these are major undertakings, involving large teams of consultants and researchers. On the average, several man years of work go into each evaluation. Consequently there are not that many evaluations completed per year, but on the other hand the information content in each could sustain policy initiatives and reform programs for several years.

"Rethinking Technical Cooperation - reforms for capacity building in Africa" is an illustrative example of this type of evaluation information. The study was coordinated by Elliot Berg, and it was financially supported by UNDP, but it is an independent evaluation/policy study.

It covers technical cooperation as such, whether implemented by multilateral or bilateral organizations. However, the efforts of the UN system play a major role in the analysis. The report refers to 231 evaluations, studies, articles and books, most of them published in the late 1980ies or in the 1990ies. Those who take the time to read the 300 pages of "Rethinking Technical Cooperation" thus get a solid discussion of a central aspect of the operations of all multilateral development cooperation, and they can search for more information with the use of the list of references in this book.

How much information is available and how much is used by the department? Based on the evidence of the material given to me, it is mainly the evaluations commissioned by other donor organizations that are used by the Swedish authorities. It is evaluations commissioned by the other English speaking donors that are adopted. I have not come across references to French, Japanese, Italian, Austrian, Belgian or Spanish evaluations in the department. For all I know, there might not be any, but given the size of the contributions from at least some of these nations, and knowing how concerned they are about evaluation, I would be greatly surprised if there were not some studies that would be worth taking part off.

In addition, I suspect there are plenty of independent researchers who publish articles, books, essays and dissertations within the field, but few of these receive attention by the department. The network of EADI (European Association of Development Research Institutions) have many members who study various aspects of multilateral development assistance, for example at the Copenhagen School of Business Administration, the universities in Leiden, Limburg, Frankfurt, Dublin, the ODI and many other. Obviously, only a small fraction of this would be relevant and perhaps there are others who digest the information and bring forward conclusions that are worth taking note of.

A study such as the previously mentioned by Elliot Berg (1993) serves such a purpose. However, this raises the important point of how reports are written. There is a tendency by many, readers and as well as authors, to shy away from the use of references. But references are an essential instrument in disseminating "information about information", and an important means of documenting how knowledge accumulates.

Studies commissioned by the Swedish authorities.

Evaluations of the multilateral organizations are usually expensive undertakings. It is useful to pool resources with other donors, both to share the costs and to use the results in cooperation. Even if there is nothing to stop the multilateral department from undertaking evaluations, it is not common practice to do so. First, as I demonstrated above, there is no shortage of information in general. Second, the absorptive capacity to "digest" yet another study put a limit on the use of evaluations. It takes time to formulate terms of reference, to recruit a team of consultants, to monitor their work and receive the

report in the end. For evaluation exercises like those mentioned above, it would probably be necessary to devote at least 6 to 8 weeks of work in the department, and so much time is rarely available.

Nevertheless, in spite of the shortage of personnel, two major studies have been coordinated in recent years; the Nordic UN project and the public commission on multilateral development assistance. Both were major studies involving several subcontracted studies as well as final reports summing up the work. The Nordic UN project started in 1988 and continued through 1992, but the follow-up activities are still going on. The main report as well as several of the subcontracted studies are still much to the point and guide the policy initiatives as well as the more operational issues that the Nordic countries raise within the UN system.

The public commission was a parliamentary inquiry into the effects of multilateral development assistance. However, even though it was a public commission, led by a chairman appointed by the government, the terms of reference were drafted in cooperation with the department for multilateral development cooperation, and one of its staff was appointed secretary to the commission. The investigative work of the commission therefore had direct links to the information needs of the department.

In spite of the constraints mentioned above, the department commissions a few studies and evaluations per year. During the last year, two studies were commissioned; one an evaluation cum policy analysis of the ITC/UNCTAD/GATT, the other an inquiry into the future of UNDP. If we now look at the total number of evaluations and policy papers commissioned by the department over a three year period, this amounts to two major policy studies and five to six minor studies (of approximately one month's working time each). If we include the annexes to the major reports, there is no less than around 2.000 pages of analytical material.

2.3 The quality of evaluation information.

The estimates above indicate that the department receives around 370.000 pages of evaluation information during a three-year period. It receives an estimated 120.000 pages per year from the organizations themselves, and over a three year period roughly around 4.000 pages from other donors and independent research and some 2.000 pages of studies that are directly commissioned. Much of this is never used, in particular the project documents, staff reports and reviews from the development banks. So there is no lack of information as such. But what about the quality of these reports, are they reliable and do they give the department the information it needs? This is in fact two separate issues, and I will start by analyzing the quality of the information.

SASDA has developed a format for assessing evaluation reports in Swedish development assistance, and as a starting point I will look at

the profile of the reports at the department according to these criteria. Before doing so, it is necessary to bear the following limitations in mind;

- the SASDA format is meant to analyze evaluations of projects and programmes, not to analyze organizations,
- the qualitative criteria are established in respect of operational decisions on projects, not on organizations,
- the department uses reports that are commissioned by others. When that happens, it has no influence on how the studies are done,
- the SASDA criteria focus on qualities of evaluation reports. This is not necessarily the same as qualities of policy analysis.

Bearing this in mind, I will now turn to the reports. The typical evaluation report takes around a year to produce and involves a team of researchers. The methods usually consist of a combination of case studies at country level, plus an analysis of headquarters operations, policies and procedures, organizational structure and hierarchical relations. The studies concerning the banks build on extensive use of economic analysis. Sometimes calculations of efficiency and cost-benefit relations are included in the reviews of the UN agencies.

The evaluation teams usually have between 5 and 10 members. The teams are international, and they mostly have participants from developing countries. When there are country studies, it is common to work with local consultants to supplement the international team. The research teams are dominated by men with an academic background in economics. Many have working experience from the agency they are evaluating. Neither donor agencies nor the organizations being evaluated form part of the team.

The format of the evaluation reports usually follow the best practices possible. They include terms of reference, work plans and lists of persons interviewed. They contain summaries and extensive reference lists. Some of the reports have comprehensive methodological statements, as major research projects would have. However, the methodological skills are more pronounced on the subjects of economic analysis of the effects and efficiency of aid projects. They are strong in respect of methods analysing effects, but weak in causal analysis.

It is often practically useful to structure a research task around a set of hypothesis, but none of the studies I reviewed made use of this method. The methods of data collection were many and varied; typically they include all methods shortlisted in the SASDA format for analysis (agency documents, other documents, agency personnel interviews, beneficiary interviews, other donors interviewed, direct observation). Most reports also discuss the reliability and validity of the findings, but mostly regarding effects. In the few cases where there is an analysis of organizational issues, I have not yet come across a discussion of validity or reliability of conclusions.

Let us look at some of the other criteria for assessing quality according to SASDA's format. Projects often have explicit objectives and thus it makes sense to analyze whether the objectives were reached. The multilateral organizations also have objectives, but the evaluations I have reviewed do not put the question in direct terms. They do not answer whether the organizations reach their objectives or not, and the truth is of course that they never reach objectives fully, the objectives are not of that nature. Similarly, there are not any formal analyses of whether the organization is needed, or whether there is an alternative. The issues of sustainability and related aspects of project performance are also largely irrelevant at the organizational level. The evaluations treat objectives, issues of sustainability and the so-called cross-cutting issues in the analysis of effects; that is, in country studies and the like.

The characteristics of the evaluation reports are an approximation of quality, but even a report that possesses many of these desirable characteristics may still be a very poor report - and vice versa. The most comprehensive effort at developing standards of excellence in evaluation was hammered out over 13 years ago by a 17-member committee appointed by 12 professional organizations, with input from hundreds of practicing evaluation professionals (Patton, 1982). The standards call for evaluations that have four features:

- utility
- feasibility
- propriety
- accuracy

The order between the features is not coincidental. The rationale is that an evaluation should not be done at all if there is no prospect for its being useful to some audience. Second, it should not be done if it is not feasible to conduct it in political terms, practicality terms or cost effectiveness terms. Third, evaluations should not be done if it cannot be demonstrated that they will be conducted ethically and fairly. Finally, if we can demonstrate that an evaluation will have utility, will be feasible and will be proper in its conduct then, and only then, could we turn to the difficult matters of technical adequacy.

How do the evaluations at the department fare when assessed in respect of these criteria? Again, I will concentrate the analysis to the major evaluations and policy studies commissioned by the department, by other donor organizations and independent research, and with occasional reference to the studies coming from the organizations themselves. Evidence from my interviews at the department as well as the content of the evaluations indicate that their utility is high. The personnel make extensive use of the evaluations. The text in the budget proposal contains several references to the evaluations, quotes them extensively and builds the arguments for future support on the evaluation information. The Nordic UN project has formed a base for many countries' policies towards the multilateral system. The correspondence between the department and the UN delegations in New York and Geneva, as well as the embassies in Rome and other places

where there are UN agencies, indicate that the evaluations are spread to those concerned with the agency in question.

However, as there is so much information available, it is usually only the desk officer at the department who takes part of evaluation reports. For example, the evaluation of UNICEF was first and foremost read by the desk officer in charge of UNICEF and by the department's management. The desk officer would discuss the evaluation, form policy on its bases, etc in cooperation with the colleagues at the delegation in New York, with colleagues in the other Nordic countries and possibly some others. The link to other Swedish aid agencies is crucial, and in particular SIDA is an important partner in using evaluations. However, the nature of cooperation varies from excellent to nonexistent.

The evaluations are shared with others in the department, but often they have no time to peruse evaluations concerning other agencies than those of their own desk. But the conclusions in some evaluations might have relevance for other agencies as well, and it might at times be an easy matter to verify if the conditions described in respect of one agency are also present in another. There is a danger that evaluations are not used to their full potential.

What about the *feasibility* of evaluations? The feasibility appears to be weakest link in the chain of excellence. First, there are political constraints. Sweden is only one of several financiers, and the possibilities of acting on the information may be limited. Other donors or stake holders may have other interests or ambitions and may thwart the changes suggested by evaluations. This is of course not a reason to refrain from attempts to reform the system, but it suggests a limit on the quality of evaluations.

Secondly, feasibility is low for practical reasons. Some of the evaluations have very high ambitions but use little resources and are conducted rapidly. The study of 11 UN agencies at country level is a case in point, and so are many of the other country studies that often form part of the analysis of effects. It is simply not feasible to answer complex questions of impact, efficiency and effectiveness of several years of development cooperation from several agencies within a few months of evaluation work. It may be possible to answer some questions, if a rigorous analytical framework is applied. The problem is that the donor agencies have little money to spend on evaluation, they have many questions that need answers, and the approach seems to be that it is better to get poor answers to several questions, than to make sure that you get highly reliable answers to few questions.

The Swedish parliamentary commission on multilateral development assistance illustrates the same deficiency. The commission was to analyse the effects and effectiveness of Swedish assistance to the multilateral system of development cooperation. Obviously, such a task is not feasible to undertake over a few years time, and the report from the commission also failed to say much of substance concerning impact and effectiveness of multilateral cooperation, that was not already known.

Propriety is a standard which is seldom raised concerning evaluations of development cooperation. Maybe the psychological distance between evaluators and those commissioning evaluations on the one hand, and those assumed to receive the benefits on the other hand, is so large that we in the profession fail to perceive the problems. Be that as it may, I do not find any reason to question propriety of the evaluations of the multilateral organizations here and now. None of the research methods raise the questions of propriety that would usually be relevant concerning evaluations, mostly because the methods are not very sophisticated. Thus they often do not even enter the territory where one can raise a discussion of whether they are ethically conducted or not (which may constitute an ethical problem in its own right).

The standard of *accuracy* concerns what I discussed above as reliability and validity, but then I only commented on whether the evaluations contained an assessment of those qualities. They did, but the question is of course whether that also means that the evaluations are reliable and valid; that is, whether they are accurate. The fact that they discuss the issue does not mean that they draw the right conclusions. Accuracy is interwoven with the problem of feasibility and the major threat to accuracy is the ambition of the evaluations; that is, their level of ambition might be too high - and at the same time too low. The evaluations seek to assess qualities and impacts that cannot realistically be answered in depth and with reasonable certainty by the means afforded, thus their ambition is too high. But the ambitions are too low when it comes to set stringent and precise objectives and allocate the means to reach those objectives.

Accuracy is an attribute of empirical data, and the evaluations mentioned above are based on three kinds of empirical sources; country studies, economic analyses of impact and organizational analyses. Often the core of the empirical data consists of judgements and opinions by project or programme personnel (in the country studies). Similarly, the organizational analyses seldom apply research methods, but are based on selected interviews or experiences from those working in the system, and more or less casual observations. They rarely contain references to scientific works in organizational studies, and in general they do not apply or discuss recent advances in organizational sciences. My hypothesis would therefore be that the accuracy of the organizational analyses is rather low.

This is a general observation, but there are exceptions to the rule. At times the evaluations discuss very concrete and specific organizational issues, such as the financing mechanisms or the nature and scope of boardwork. Here accuracy is higher and it is possible to deduct normative prescriptions that are firmly based in organizational theory. The problem is larger in areas such as the configuration of roles, best managerial practices, the functions of monitoring and evaluation systems, etc.

The highest accuracy is found in the economical analysis. In reports such as the Wapenhans study, or the reviews of the quality of bank

lending, the methods of assessment are clearly presented and of the highest standard. I have not checked the calculations, but from what is presented in the studies, there is no reason to doubt their accuracy. It is of course always possible to raise questions regarding pay-off times, the size of intangible benefits, or opportunities foregone, but such issues are always estimates. The most important thing is that the reports enable an inquisitive reader to assess the outcome of different assumptions.

2.4 Summary

The review of evaluations at the Ministry for Foreign Affairs' department for multilateral development cooperation shows that there is no shortage of evaluation reports as such. A rough estimate yields around 370.000 pages of information over a three year period, or around 125.000 pages per year. The evaluations come from three sources; (a) directly from the multilateral organizations evaluation units, (b) from other donors and independent research, or (c) studies commissioned by the department itself, alone or together with other donors.

In this study I have concentrated the analysis to a handful of evaluations and policy studies from all three sources; evaluations and studies that assess the operations of one or several of the multilateral agencies. The majority of the reports are of a high quality, most of them research work at the frontier of knowledge about development cooperation. The evaluations are undertaken by international teams, often with participants from the Third World. The economics profession dominates strongly. The reports are comprehensive in their assessment of impact and effectiveness of cooperation. Several of the reports also build on country studies. This level of aggregation seems unfortunate and the conclusions at country level appears to be a weak link in the chain of evidence.

What do the reports cover? An overwhelming share of the information concerns the impact of assistance, and this is also the most accurate and most reliable part of the information. There is less information on the organizational qualities that deliver a high impact. A rough estimate might be that around 80% of the information concerns impact and effectiveness of cooperation, and around 20% treats the delivery systems; that is, the causal link between organizational design variables and the effects of assistance. In brief, both the quality and quantity of information on the effects of multilateral aid are high, and both the quality and quantity of information on the system itself, its organizational characteristics, are low.

Table 2.1 summarizes the argument. The vertical axis lists the sources of evaluations and the horizontal axis consists of the qualities of evaluations, plus at the far end a column that gives the quantity of information in each category. The patterns in the boxes reflect a judgemental short-note of high, medium or low quality from the different sources. Again, it is clearly seen that most of the information

Table 2.1 Summary of the sources, quantity and quality of evaluation.

	Utility	Feasibility	Propriety	Accuracy	Quantity * (no of pages)
(a) from the international organizations			n.a		360.000
(b) from other donors or independent research			n.a		4.000
(c) commissioned by the department			n.a		2.000
Code: ■ High ■ Medium □ Low * over a period of three years n.a - not applicable Note: Within each group, the diversity of studies is wide. Some of the studies in category (a) have high utility, and some (c) low utility.					

concerns effects and impact, and is of relatively low utility - even though it is of high quality in other respects. The information generated by donors have higher utility, but ranks lower in respect of the other qualities.

3. A note on the art of organizational analysis.

3.1 Introduction

In this section of I discuss elements of organizational analysis. The department for multilateral development cooperation consists of 20 people. Many of them have a background in economics and business administration, and that implies some familiarity with organization theory. However, that familiarity mostly stops at the level of undergraduate studies, at least in theoretical terms. On the contrary, many have several years experience working in different organizational contexts. The department's mandate requires much thought on organization. The practical experience is considerable. The challenge is to make the practical experience confront theoretical concepts and issues, with the purpose of structuring and developing the level of inquiry.

I write of organizational analysis as an art precisely because it is an exercise which builds not only on formal inquiry, but also on intuition and a sense for holistic systems and aesthetic qualities. There is no standard format for organizational analysis. Organizations are open

systems, and must be confronted as such. There are many overlapping routes to understanding them, but most understanding is partial. Some areas of analysis are more technical than others, for example budget examinations and systems of financing. The design of monitoring and evaluation systems may be a relatively simple technical issue, but when we move to the level of organizational learning (which contains systems of monitoring and evaluation) we stand in front of a more complex research issue.

Chapter 2 demonstrated that the mass of information concerns the effects at field level of the multilateral organizations. This has been the starting point in evaluation for the last 20 years. The logic is that if the effects are sound, then the system producing the effects functions well. Consequently there would not be any need for more extensive organizational analysis. Unfortunately, the effects have never been good enough to let us leave the organization producing them alone. To put it mildly, all reviews of effects have indicated that there is room for improvement in organizational design.

The first task is to develop an understanding of the organization which is the object of inquiry, be it a small unit as the ITC or a larger one such as the UNDP or the World Bank group. Second, the department needs to develop strategies and policies to improve the performance of the organization. Third, the department must assess as accurately and objectively as possible, the performance of the organizations. These tasks overlap each other and develop jointly, but in the following I treat them in turn.

3.2 The nature of organizations.

The multilateral organizations can be said to have both a technical and an institutional environment, in the terminology adopted by Meyer and Scott (1983). A technical environment evaluates and supports an organization in terms of its products and results; the institutional environment judges it on its structures, processes and ideologies. The institutional environment is not equally important to all organizations; it is presumably more important to banks and hospitals than it is to a restaurant or a power supply company. The norms concerning structures, processes and ideologies are often explained and justified on grounds that they should contribute to efficiency. Unfortunately, there is no guarantee that the demands imposed by the institutional environment are compatible with effective action. Brunsson (1989) shows how modern organizations often are exposed to inconsistent demands from their environments, and describes two ways organizations cope.

The first principle is associated with the assumption that the survival of an organization depends on an exchange of resources with its environment; that is, the organization can supply goods and services for which it receives money to finance its operations. The other principle is associated with the assumption that the survival of an organization depends upon its ability to reflect and create a symbolic

accord with its environment. An organization designed in accordance with the first principle can be dubbed an action organization, and the second a political organization.

Obviously, the multilateral organizations do not occur in these pure states; they have to be good at both producing action and at creating a symbolic accord with donors. But by looking at the ideal types we can see the differences between the two kinds more clearly, and this will provide a better understanding of the organizations. Let us briefly look at the characteristics of the two ideal types.

An action organization formulates rules to govern the behaviour of its members. Powerful organizational ideologies tend to evolve under their own steam; people who meet and work together easily come to think along similar lines. But because organizational ideologies are such crucial instruments of coordinated action, organizations also make concrete efforts to create and maintain them. Managers define goals and work out policies and business ideas; they try to create a corporate spirit, a corporate culture, and so on. In this way the variety of ideas, perspectives, notions and opinions which members presumably have when they join the organization is reduced and absorbed into the common ideology. In a complex world which can be understood in many ways, the organization circumscribes its members' understanding. To put it bluntly, the action organization makes people narrower and stupider than they were before they joined it.

A powerful organizational ideology cuts down the need for making decisions. It is often obvious what action should be taken. The decision is meant to reinforce people's expectations that a given organizational action really will be undertaken, to motivate the actors to carry it out and commit them to action. Decisions with this kind of action-reinforcing function should not obey the norms of rational decision-making; action rationality is very different from decision rationality. Rational decision-making all too easily produces uncertainty and doubt; the decision process should instead be systematically irrational.

The action organization specializes; it may limit itself to one type of technique, one type of product or addressing one market. Business ideas are often ideas about niches, or about special needs which the organization intends to satisfy. There are many reasons for specialization; the most frequently cited being to achieve scale economies. But there are also ideological arguments. Organized action calls for powerful organizational ideologies, and if we subscribe to one then many otherwise possible actions are excluded. Organizational ideologies should be precise, complex and consistent - which also makes them narrow and concerned exclusively with a limited section of reality.

The action organization is geared to solutions rather than problems. Only when we have a solution can we act, and shared loyalty to an organizational ideology will make it easier to find solutions. (The car was the idea Ford had in mind, not the problem of transportation; ball-bearing was the idea that launched SKF, not the problem of friction).

Action is easier when a powerful organizational ideology obtains and conflict is suppressed. Thus the action organization cultivates great confidence. The members of the organization are convinced that their view of the world is the right one. This faith may sharply conflict with what appears reasonable to an outsider, but an outsider can afford to be sceptical and critical since he has no need to generate action. Firm confidence affects the capacity for action, and it is difficult to see how it can be abandoned without reducing the ability to act. The action organization cultivates a spirit of enthusiasm rather than criticism.

The antipole of the action organization is the political organization. Its main bases for legitimation is that it reflects inconsistent norms. Instead of seeking niches like the action organization and satisfying one need or interest at a time, the political organization reflects a variety of ideas and demands and satisfies the expectations of diverse groups in its environment. Important structures and processes in the political organization are aimed directly at the environment. The idea is to demonstrate them to the outside world and to show that the organization is attending to the inconsistent demands made upon it.

A good way for the political organization to maintain support is to develop several different organizational ideologies. The political organization is multi-ideological; it includes lots of ideas about the nature of the organization and its environment, and about what the organization should do. The organization reflects a complex environment full of inconsistent ideas in a series of ideologies that are also inconsistent.

The political organization is intellectual in three respects. To begin with, it embraces several ideologies. Its members' ideas are not regimented. Countless ideas and opinions are elaborated and exploited. Analysis often benefits from diverse approaches and an abundance of conflicts provokes criticism. Lively criticism in turn reduces the risk of coming to the wrong conclusions. Thus the political organization should be better equipped than the action organization to understand a complex and change-prone world.

Secondly, the political organization is intellectual in that it tends to follow the norms of rational decision-making. It devotes a good deal of time to decisions, describes activities in terms of decisions and indicates clearly when decisions have been made. This is logical. In the absence of a dominating ideology, there is no obvious way to act. Because of inherent conflict in the organization, several different alternatives to act are certain to be suggested. The rational decision process provides an excellent way of demonstrating conflict. Establishing different goals, handling many alternatives, describing all possible positive and negative consequences of alternatives and discussing which one is best - all these activities give the organization opportunities to reflect inconsistencies in its environment.

The political organization is also intellectual in that it is largely concerned with problems. Problems, not solutions, are its strength. It also entertains a special affection for tough or insoluble problems.

Insoluble problems are a splendid vehicle for reflection of many ideas or values. They can be endlessly discussed from all sorts of angles and without ever reaching a conclusion. Instead of specializing, the political organization tends to generalize itself. The reflection of inconsistent values is itself an expression of a generalizing strategy. But the political organization that wants to grow will actively seek to incorporate new ideas in the environment into its own organization. No group needs to be left outside its domain. The organization grows by reflecting an increasing number of inconsistencies. As it does so, it becomes less and less likely to act.

Remember that the caricature above is meant to show archetypes of organization. In real life organizations combine elements of both. The point is that some organizations are more like action organizations and some more like political organizations, and the recipes for success depends on how they are identified. I would like to end the typology that was briefly reviewed above with the following observations:

First, to understand the multilateral organizations, it is necessary to bear in mind to what extent they are political organizations. When an organization is built on the logic of a political organization, there is a very clear limit on its ability to act effectively - meaning according to notions of cost-effectiveness etc developed for action organizations. In fact, doing so might threaten its survival to the extent that its legitimacy depends on reflecting inconsistent environments. The same applies in reverse, an organization which is more like an action organization may have a limited capacity for decision rationality and a more limited capacity to understand and analyse the world.

Second, in practical terms, the multilateral organizations need to develop instruments both to act and handle inconsistent demands from the environment. These are rather sophisticated tasks, and require a subtle balance between parts of the organization, their roles and strategies. Nowhere in the review described in chapter 2 have I found an awareness or even a reference to the existence of these distinctions and the significance they may have for the future of the organizations. It is not likely that organizational reforms will yield results as long as the basics of understanding organizations are not in place.

Third, organizations change and the dynamics of changing from an action organization to a political organization are poorly understood. Several of the multilateral organizations seem to change in that direction; for example the World Bank and UNICEF. Others may want to change in the opposite direction, such as the UNDP. Many new organizations are set up as action organizations, but they grow as political organizations.

Fourth, the system of international organizations in itself reflects environmental inconsistencies and some of its members have been established with more of the characteristics of actions organizations and others as political organizations. But they still constitute a system, and they are meant to cooperate with each other. It is commonplace to comment on the difficulties of coordination between

the organizations, but it appears as if the problems of coordination are largest when an action organizations operates alongside a political organization. This is often the case in emergency operations, and the problem may well be aggravated when a political organizations is set to coordinate several action organizations.

3.3 Design variables and patterns of effective management.

Now that we have reviewed the basic type of organizations, which both raises theoretical issues and presents a highly practical challenge, I will turn to some aspects of management that may have particular relevance when assessing the multilateral institutions. These are issues to look out for in evaluation reports, to request information on from boards, and possibly to observe and assess during visits to the organizations.

The essence of management comprises three factors (Eccles and Nohria, 1992): *rhetoric*, the use of language to create meaning, frame action and prompt creativity; *action*, which highlights getting things done; and *identity*, which refers to an appreciation of how people's identities get built and maintained in organizations. The purpose of management is to foster action and then make that action meaningful both collectively and individually.

Let us first review some elements of design that are essential to the identity and rhetoric of an organization. Weick (1993) identifies three systems properties that are fundamental to an organization's ability to cope with change and with chaotic environments: (1) improvisation and bricolage, (2) virtual role systems, and (3) the attitude of wisdom. As I understand it, these are also fundamental aspects of organizational culture.

Improvisation and bricolage.

Organizations need to let members develop skills in improvisation and bricolage (Janowitz, 1959). Bricoleurs remain creative under pressure, precisely because they routinely act in chaotic conditions and pull order out of them. Thus when situations unravel, this is simply normal natural trouble, people proceed with whatever material are at hand. Knowing the substance of their tasks intimately, they are able to form new insights when they face challenging situations. These properties would be particularly essential in organizations handling emergency situations, catastrophes and other complex missions. Weick and Roberts (1993) describe the fundamental organizational features of such systems.

Virtual role systems.

The title may seem mystifying, but the point made is not that complicated. In each organization there is an element of social construction of reality. Part of that social construction constitutes the

role system of the organization. However, the organization cannot depend on continuous social construction to formalize its role system, instead the role system must be reconstituted and "run inside each persons head". In the manner of a holograph, each person should be able to reconstitute the organization and assume whatever role is vacated, pick up activities and run a credible version of the role. Furthermore, people may to some extent be able to envisage the organization in their minds and use the vision for continued guidance of their own individual action (Weick and Roberts, 1993). This is the meaning of a virtual role system.

The attitude of wisdom

To understand the role of wisdom (Bigelow, 1992), consider what Maclean (1992) calls the first principle of reality: "little things suddenly and literally become big as hell, the ordinary can suddenly become monstrous, and the upgush breezes can suddenly turn to murder". The same point is illustrated with more general references by Waldrop (1993). Still, most organizations fail to understand the consequences and thus fail to learn, because "Reality backs up while it is approached by the subject who tries to understand it. Ignorance and knowledge grow together" (Scott, 1987). To put it in a different way: each new domain of knowledge appears simple from the distance of ignorance. The more we learn about a particular domain, the greater the number of uncertainties, doubts, questions and complexities.

The role system best able to accept the reality that ignorance and knowledge grow together may be one in which the organizational culture values wisdom. Meacham (1983) argued that wisdom is an attitude rather than a skill or a body of information:

To be wise is not to know particular facts but to know without excessive confidence or excessive cautiousness. Wisdom is thus not a belief, a value, a set of facts, a corpus of knowledge or information in some specialized area, or a set of special abilities or skills. Wisdom is an attitude taken by persons toward the beliefs, values, knowledge, information, abilities, and skills that are held, a tendency to doubt that these are necessarily true or valid and to doubt that they are an exhaustive set of those things that could be known.

In a fluid world, wise people know that they don't fully understand what is happening right now, because they have never seen precisely this event before. Extreme confidence and extreme caution both can destroy what organizations most need in changing times, namely, curiosity, openness and complex sensing. The overconfident shun curiosity because they feel they already know most of what there is to know. The overcautious shun curiosity for fear it will only deepen their uncertainties. Both the cautious and the confident are close-minded, which means neither makes good judgements.

Organizational culture is a phenomenon that can be described in many ways, and several researchers have suggested dimensions to analyze culture and to describe it (for example Frost et al 1991, Weinshall, 1993). There is basically only one thing to add here. Even though there

are several approaches to the phenomena and some may lead to dead-ends, no practical or meaningful analysis of an organization should ignore the cultural aspects of management.

Organizational learning

I have never heard anybody argue that an organization could do without learning. The approach to learning is closely connected to the organizational culture, it presumes an attitude of wisdom and it is a precondition for improvisation and bricolage, as described above. Still, this raises the question whether organizations really can learn, or whether learning is a capacity vested in individual human beings. Certainly, learning is first and foremost identified at the level of individuals, but it may also be possible to speak of organizational learning.

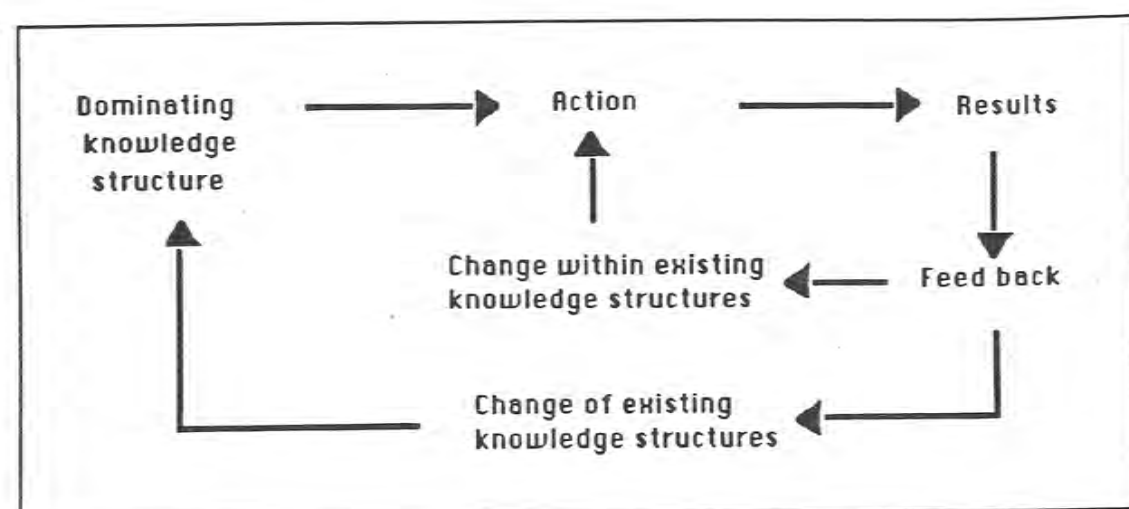
However, research on organizational learning shows that people do not think and learn as much as they should, and consequently organizations are not as effective as they might have been (Argyris & Schön, 1978; Mitroff, 1983; Schein, 1985). Researchers have pointed at the lacking capacity to change thinking (Starbuck, 1983) and to stimulate several ways of thinking (Gray, 1987) among the organization's members as the primary obstacles to more effective learning - and consequently more effective performance.

If we start by defining what it is that needs to be changed, research in cognitive psychology and epistemology suggests that people as individuals and in groups construct reality. Berger and Luckman (1966) portray society as a social construction, produced by peoples action and thinking. Social cognition develops because of the individual human beings need for order (Douglas, 1987). These social processes make up the foundation of the knowledge structures that each one of us possesses (Kelly, 1955; Perry 1986). The key to learning is to affect individual knowledge structures.

Evaluation is by definition synonymous to feedback, and feedback is the link between performance and knowledge structures. Figure 3.1 below illustrates the connection. What exactly is it that happens when people learn? The keyword is change; learning implies a change in what we know (our knowledge structures) and, by consequence, in how we act. Organizational learning similarly means changes in what the organization knows and how it acts.

These knowledge structures may be more or less amenable to change. But how do knowledge structures change? For one thing, that depends on how well developed they are (meaning rich, diversified, advanced). Piaget (1976) pointed out that well developed knowledge structures facilitates change, and then change occurs by assimilating new knowledge and further refining the knowledge structures (also exemplified by Sevón, 1984). But less well developed knowledge structures may find it have difficult to absorb new knowledge without

Figure 3.1 The learning process.



drastic reorientation. It is then common to speak of accommodation. In plain terms, the more you know, the easier is it to learn to more - and to do so without painfully drastic changes of knowledge structures.

An organization, like an individual, may receive information inputs that demand a total overhaul of knowledge structures. Assume that an atheist suddenly finds evidence that God exists, or a fundamentalist who is confronted with evidence of the Big Bang. Both will have to revise their view of the world. Their old knowledge structures break under the impact and new ones take their place.

But the input of information may be quite substantial but perhaps it still fits with existing knowledge structures. A student of physics may be engaged in a research programme to explore the content of black holes in space. In the course of investigation large amounts of information may be encountered, but perhaps nothing that alters his or her basic view of the world. Learning takes place, but the new information is incorporated in the old knowledge structures and leads to their further development and refinement.

In the theory of learning it is common to distinguish between single-loop and double-loop learning (Argyris & Schön, 1978). The difference between the two is that single-loop learning is supposed to take place within existing knowledge structures, developing and refining them. Double-loop learning implies that the old knowledge structures have become redundant and that they no longer help us perceive and act in the world (to put it drastically). The authors also identify a third level of learning which they call deuterio-learning. This means the capacity of learning how to learn; that is, how to systematically generate and make use of single- and double-loop feedback mechanisms.

My point here is that organizational learning can be operationalized into research tasks, but it may also be possible to assess learning more quickly. For a start some of the basic concepts and categories must be identified, and once they are in place, it might be possible to assess learning practically. One of the key instruments in a learning organization is the systems for evaluation. Evaluation must of course not be confused with learning; it has many other purposes. But evaluation is one of the most important channels of documenting and assessing past performance, thus it provides much of the fuel to fire learning. Let us look in some detail at evaluation systems.

Evaluation.

Many of the evaluations that were reviewed in chapter 2 do point at the necessity of building systems of evaluation and monitoring in the organizations. Some point at the lack of monitoring and follow-up, others discuss the role of the evaluation unit. As far as I know, there are not two aid organizations that have designed their evaluation systems similarly. Even in Sweden, the variety between our five aid agencies is considerable. Which are the most important structural features of an evaluation function? It is possible to discern at least five aspects of the design of an evaluation unit that distinguish between different practices in the donor community (Samset, Forss and Hauglin, 1993). These are: (1) independence, (2) differentiation, (3) centrality, (4) openness and, (5) the formalization of the systems. Although these refer to bilateral agencies in the text that follows, it might be useful to discuss the evaluation systems in the multilateral organizations in similar terms

(1) Independence. In evaluation theory it is often maintained that evaluators should be independent of the subject they evaluate. Similarly, an evaluation unit should be independent of the system it evaluates. Among the major donors it is only Sweden and the World Bank who have an independent evaluation unit, placed outside the agency(ies) and with its own governance. Both report directly to the main stakeholders; that is, parliament (Sweden) and the Board (World Bank). The Netherlands Joint Inspection Unit also has considerable independence from the system, but it reports to the Minister concerned with development cooperation, not to the "public". In the other multilateral organizations the evaluation unit is a part of the administration.

(2) Differentiation. In principle evaluations may cover a broad range of issues concerning the more limited aid delivery, impact assessment at project, programme and sector levels, or policy issues and other fundamental issues concerning sustainability, relevance of different objectives, etc. Some of the administrative systems limit the scope of the evaluation unit to one or a few aspects. Other systems assign responsibility for different types of evaluations at separate levels, yet others give different levels freedom to pursue the questions they find relevant over the whole field of possible evaluation tasks.

(3) Centrality. An evaluation unit which is centrally placed may be more likely to reach top-level decision-makers. But in some systems the top-level decision-makers are notable through their absence as power is decentralized. In that case a central evaluation unit faces the risk of being marginalized. But that is a risk even if decisions are taken centrally, because a central unit may lack a grasp of real issues, that are relevant at the moment. In some administrations the evaluation unit is centrally placed, as in the UNDP and the U.K. Yet other systems have evaluation units that are not close to the centers of power and others have evaluation units at central as well as other levels.

(4) Openness. The practice of making documentation from the civil service open to the general public varies from one country to another, and the practice of making evaluations of aid public follow these general patterns. In several systems such studies are confidential whereas they are actively made public in others. However, there is a trend towards greater openness.

(5) Formality. The evaluation system could be more or less formalized in the process of working; that is, the content of the evaluations set and the objects determined according to rules. The evaluation system could also be more or less standardized as to what follows after a report. In some systems there are procedures for writing, circulating and using summaries and in other systems there are virtually no rules for anything, the reports are made and spread ad hoc. It is more common to have standardized the content and process of initiating evaluations than to have standardized the follow-up, but there is a trend towards formalization and standardization in most donor countries, not least initiated by the DAC.

The most important distinction seems to be between external versus internal units and between widespread evaluation functions or a centralized unit. If anything, the trend has been towards a more fragmented use of evaluations. As aid amounts increase and more actors become involved the need for feed-back grows. Each actor needs some control of its operations, and some need an overview. Systems go toward more fragmented structures, and consequently with evaluation functions spread throughout the administration.

In a way this resembles an experimental situation. Many actors try to do the same thing in the same environment, and they have evolved different ways of accomplishing their objectives. True, the size of their aid contribution varies and some have a longer history of cooperation than others. They give different emphasis to the same set of objectives. But they all have evaluation systems intended for management and learning. So, which system works best, which makes the most important contribution to an aid program of high quality?

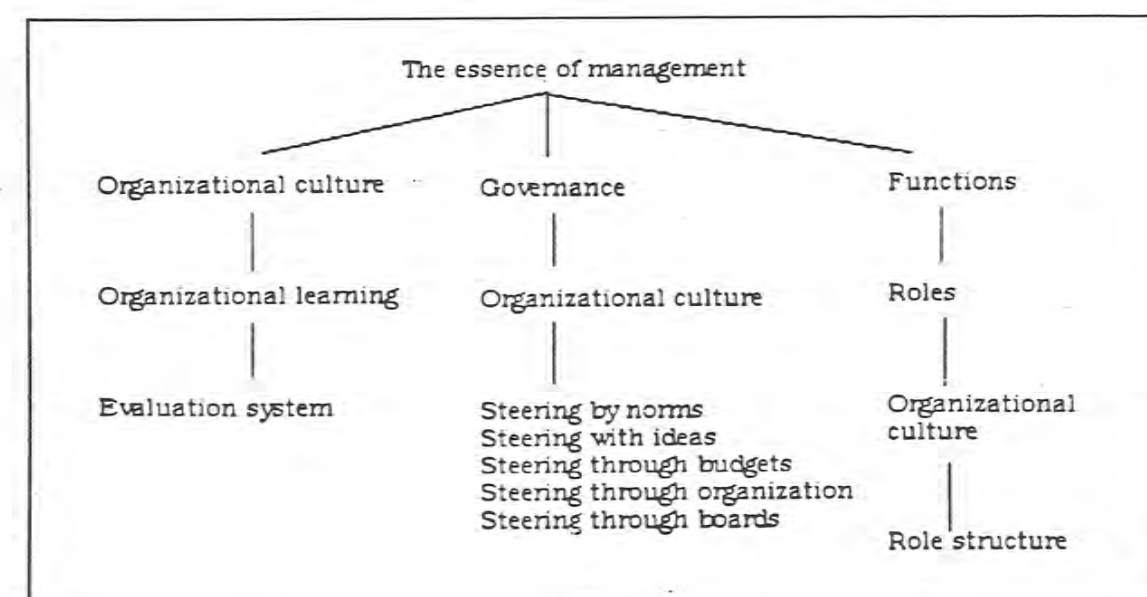
The question cannot be answered. There is no research and no studies that attempt to make such a comparison. Only two system have been assessed in terms of learning, and that is Sweden and Norway. Neither was comparative. The few comparative studies that exist (DAC, 1989; Cracknell, 1991, Stokke, 1992, Bovaird, Gregory and Stevens (draft)) do

not go so far as to assess performance or contribution. They are descriptive and analytical, and possibly there is not yet any sufficient methodological framework for a comparative, comprehensive review of the effectiveness of aid evaluation systems (perhaps not an audience either). There is a danger in that descriptive studies tend to put a virtue on the formal aspects of the system, and underrate the effectiveness of informal channels of feedback. They may also miss the subtle connections between learning and decision-making that only a close and intense scrutiny of a system can start to yield.

3.4 Conclusion.

In the review above, we have moved from the more general levels of organizational analysis to the more specific issues of evaluation. The path has led from the general and comprehensive to the specific and concrete. It illustrates one of the many possible paths of organizational analysis. The paths to the right of the one we took highlight other issues that may be put in focus of an organizational study. Now, given the task of the multilateral department, are there any paths that provide better hunting grounds than others - that leads to more fruitful organizational analysis?

Figure 3.2 Paths of organizational analysis.



Reversing the order compared to the discussion above, I would suggest starting with the concrete and possibly technical issues and moving to the more general and comprehensive:

(1) Financing mechanisms. The Nordic UN project has pointed at the necessity of providing money to the organizations with reasonable predictability and in phase with activities.

(2) Systems of planning and budgeting. How are plans formulated and developed from the strategic levels, to tactics and operations? How are budgets designed and how does one analyze the components of a budget proposal?

(3) Systems of monitoring and evaluation. Key elements were discussed above.

(4) Personnel policies, in particular levels of competence, recruitment and mobility within the organization.

(5) The organizational structure. How is labour divided between units in the organization, how do they relate to each other and how are hierarchical relations formed and maintained?

(6) Coordination. What types of decisions and activities need to be coordinated, how is coordination achieved, what are the costs involved and which instruments are used for coordination?

(7) Organizational functions can be described in terms of roles; which roles do the organization play, how do they interact and overlap, and how can the role structure be maintained?

(8) Systems of governance. The boards are an important ingredient in the system of governance, but organizations are governed through other instruments as well. It is essential to develop a clear picture of the system of governance, not least to see where the functions of management start and those of the board end.

I would like to see aspects of organizational culture included in the list of organizational elements. In the long run organizations change. The evaluations listed above imply that some of the multilateral organizations need to change rapidly; and some seek change actively. Whether they will succeed or not, and through which processes, largely depends on the organizational culture. In addition, change agents would normally start by affecting the organizational culture. The list is not comprehensive and the discussion above only touches on some issues of organizational analysis. However, I hope it has demonstrated some issues that may be of concern, and perhaps it has sharpened the appetite for additional knowledge.

4. Concluding remarks.

The multilateral department is a small unit of around 20 people handling the more than 3.000 million Swedish crowns allocated to multilateral development cooperation. The department has a mandate to follow-up the disbursement of funds to ascertain that these really contribute to the objectives of Swedish development assistance. As part of the information on whether they (the funds) do or don't, the department receives some 125.000 pages of evaluation information per year.

The question is if the department receives the information it needs to assess the use of funds accurately. The tentative answer from my review is that there is plenty of information on the impact and effectiveness of the assistance provided by the multilateral organizations, particularly that provided by the large organizations such as the World Bank, UNDP, UNICEF, UNFPA and WFP. The feedback on performance at project and programme level is quite satisfactory, but it shows that the performance of the organizations leaves much to be desired.

The government presents effects of multilateral cooperation to the Parliament in the budget proposal. The text in the proposal contain many references to evaluations and provides a shortlist of the problems and opportunities faced by the multilateral organizations. It is frank when it describes impact and effectiveness, and I suspect the quality of information to Parliament on multilateral development cooperation is considerably higher than on bilateral development cooperation, in spite of the greater difficulties involved in getting information from these organizations.

The weak link in the chain of information consists of organizational knowledge. There are several evaluations of organizations, but for reasons described above, these are not always the most accurate of evaluations. In addition, the department does not only need evaluations, it needs knowledge on how to act and what issues to raise in board work and through other mechanisms of governance in relation to the multilateral organizations. There is no need for additional studies of impact and effects, but there is a need for organizational analysis and policy studies that could help the department act as a change agent in respect of the organizations Sweden contributes to.

In chapter 3 I touched upon a number of central issues in management theory. Few of the evaluations provide any information in respect of these, and those that do have often been initiated directly by the department - in particular through the Nordic UN project and studies that have followed suite. Not only do the major part of evaluations provide limited and selected treatment of organizational design, but this fact is also recognized by people at the department.

However, this is not only a question of evaluation. First, it is simply not possible to commission large numbers of organizational evaluations cum policy studies. The capacity to do so is not present.

and even if it were, the capacity to receive and absorb the information would still pose a bottle-neck.

Second, the need for organizational knowledge is not only a question of information but of competence and organizational learning. In the long run, to have a sustainable impact on the operations of the multilateral system, the organizational knowledge needs to be internalized at the department. I do not assume the department could be transformed into a research institution on multilateral development cooperation - but the level of skills could still be upgraded without overdoing it. In the short run, the following steps could be initiated:

- Focus professional development on issues of organizational analysis. The department could invite lecturers to conduct seminars on topics such as coordination, governance, organizational culture, organizational learning. By devoting half a day every month to the subject, it would be possible to enable most of its personnel to encounter the research frontier in each of these areas.

- Focus environmental scanning techniques on issues of organizational analysis. The department should subscribe to at least one of the important journals publishing advances in research, and in academic data bases search for articles or books on multilateral organizations regularly.

- Make sure that the analyses of organizational variables in the incoming evaluations are disseminated among the staff, reviewed, criticised and compared other organizations.

- The long run effects of the Nordic UN project are considerable. I do not suggest that the experience as such be repeated, but the value of having an ongoing policy study were considerable. Perhaps there are other forms of financing and conducting such studies that could be attempted, for example by closer cooperation with academic establishments, "think tanks" or even individuals.

- The internal organization may be changed into a full matrix structure, perhaps incorporating three dimensions; organizations, topics and organization. I have not studied the option, but it could be elaborated.

- Competence may also be developed through recruitment. I cannot weigh the relative merit of different additional competence profiles at the department, but if it is essential, it might be possible to give priority to one of two persons with a considerable background in organizational analysis.

Annex 1. Terms of reference.

Utrikesdepartementets system för utvärderingar och bedömning av multilateralt bistånd

Uppdraget syftar till att beskriva och göra en kvalitativ bedömning av utrikesdepartementets underlag inför beslut om anslag till, och utvärderingar av tidigare bidrag till multilaterala organisationer. Utredningsuppgiften är en del av SAUs analys av utrikesdepartementets och biståndsorganisationernas förmåga att lära av egna erfarenheter samt att systematiskt följa och självkritiskt granska sitt arbete.

Uppdraget utgår från Regeringens uppfattning om hur stödet till multilaterala organisationer bör uppfattas och de krav som kan ställas på deras effektivitet i arbetet. Utrikesdepartementets stöd till dessa organ styrs av strävan att utifrån svenska biståndspolitiska mål och prioriteringar skapa förutsättningar för dessa multilaterala organ att verka på ett effektivt sätt. Kraven på effektivitet från medlemsländerna måste således i första hand riktas mot organisationerna själva (se budgetproposition 93/94:100, sid 85). Handläggningen av stödet bör alltså karakteriseras av denna målsättning. Därtill bör den skötas med god inre effektivitet vad gäller uppföljning av hur dessa mål uppnås.

Det innebär att i uppdraget ingår att analysera följande frågeställningar:

- 1) vilken typ av information behövs för att avgöra om biståndet genom en viss organisation kan förväntas vara ändamålsenligt och om verksamheten bedrivs effektivt av organisationen utifrån ett svenskt biståndspolitiskt perspektiv?
- 2) i vad mån finns sådan information tillgänglig på utrikesdepartementet inför beredningen av beslutsärenden och i vilken utsträckning används den?
- 3) genom vilka kanaler har utrikesdepartementet skaffat sig tillgång till den information man har enligt punkt 2) ovan?
- 4) i den mån det brister mellan 1) och 2) ovan, hur skulle utrikesdepartementet kunna förbättra sitt arbete med att skaffa sig den information man saknar och/eller bättre använda den information man har?

Syftet är alltså både deskriptivt, analytiskt och normativt; arbetet ska bidra till att ge en analys av utrikesdepartementets arbete med beslutsunderlag och utvärderingar och skall bidra till ett mer ändamålsenligt beredningsarbete.

I det deskriptiva och analytiska arbetet är det två typer av material och två typer av frågeställningar som primärt bör granskas:

a) svenska och i andra hand andra givares utvärderingar (eller motsvarande material, retrospektiva rapporter, studier, forskningsrapporter etc) av multilaterala organ deras verksamheter; och i samband med detta användningen av dessa utvärderingar i det egna arbetet;

b) organisationsbedömningar och annat beslutsunderlag; och i samband med detta användningen och refererandet av materialet i de slutliga besluten samt i vilken mån detta beslutsmaterial lägger grunden för en systematisk resultatvärdering.

Materialgenomgången vad gäller utvärderingar dokumenteras med stöd av bifogade mall för värdering av utvärderingar.

Bakgrund

Uppdraget utgår från på en tes om utvärderingars roll i en organisations arbete. Denna tes kan kort formuleras som att en organisations förmåga att lära av egna utvärderingar av genomförda uppgifter fungerar väl endast under vissa förutsättningar. En viktig sådan förutsättning är att organisationen redan vid initiering en av en ny fas i stödet gör det tydligt vilka mål och kriterier som kännetecknar framgång i verksamheten och därmed hur graden av framgång eller misslyckande skall kunna avgöras. Detta underlag bör göras tydligt i planeringsprocessen, inklusive projektbedömningar och beslutsunderlag. Med denna utgångspunkt kan sedan en reell kontroll göras av arbetet både under gång och efter avslutande genom utvärderingar, där de ursprungliga föreställningarna om mål, kriterier och värderingar ställs mot verksamhetens resultat (inklusive ändrade mål om så behövs). Detta sätt att arbeta med utvärderingar kan ge organisationen en systematisk återföring av erfarenheter i en process som kan karakteriseras av substantiellt lärande. Mot detta står utvärderingar använda som mera formella, ytliga instrument för verksamhetsuppföljning och -kontroll i processer som karakteriseras av mer eller mindre symboliskt lärande. För referenser och utförligare kommentar hänvisas till Bra Beslut (Ds 1990:63), sid 89-97.

Genomförande och omfattning

Arbetet genomförs i nära samarbete med Lennart Båge vid utrikesdepartementet. Hjälp kan påräknas från UD med framtagning av material.

Uppdraget omfattar 4 arbetsveckor för Kim Forss, Andante.

Avrapportering

Arbetet avrapporteras senast 1 maj 1993. Rapporten skall vara skriven på engelska och levereras i manuskript (och på diskett) som möjliggör tryckning i SAU's rapportserie.

Annex 2. Evaluation documents.

Perspectives on multilateral assistance. A review by the Nordic UN Project. Stockholm 1990.

Bistånd genom internationella organisationer. Betänkande av kommittén för översyn av svenskt multilateralt utvecklingssamarbete. SOU 1991:48. Stockholm

Report of the Task Force on Improving Project Quality. 1994

Effective Implementation: Key to development impact. Report of the portfolio management task force (the Wapenhans report).

Evaluation of UNICEF. Synthesis report. 1992

Evaluation of UNFPA. Synthesis report. 1993

Evaluation of the World Food Program. Chr. Michelsen Institute. 1993

Evaluation of the use of WFP non-food resources.

Quo Vadis UNDP? A discussion paper on the UNDP. Chiasen AB, 1993.

ITC/UNCTAD/GATT. A review of issues regarding its future. Chiasen AB, 1994.

Managing for effective development. Report of the task force on portfolio management for the Inter-American Development Bank. 1993.

Effectiveness of Multilateral Agencies at Country Level. UNDP in Kenya, Nepal, Sudan and Thailand. DANIDA 1991.

Case study of 11 agencies in Kenya, Nepal, Sudan and Thailand. DANIDA 1991.

Rethinking Technical Cooperation. Elliot Berg and UNDP, 1993.

Annex 3. References:

- Argyris, C. and Schön, D. 1978. *Organizational Learning*. San Francisco: Jossey Bass.
- Berger, P. and Luckman, T. 1966. *The Social Construction of Reality*. New York: Doubleday.
- Bigelow, J. 1992. Developing managerial wisdom. *Journal of Management Inquiry*, 1:143-153.
- Berk, R.A. and Rossi, P.H. 1977. Doing good or worse: Evaluation research politically reexamined. In G.V. Glass (Ed) *Evaluation studies review annual* (Vol. 2, pp. 77 - 89). Beverly Hills, CA: Sage.
- Bovaird, A.G., Gregory, D. and Stevens, J.N. (4th draft - no date) *A study of multilateral agencies evaluation systems*. Volume 1.
- Brunsson, N. 1989. *The Organization of Hypocrisy. Talk, Decisions and Actions in Organizations*. New York: John Wiley
- Cracknell, B. 1991. *Evaluation Feedback. Project Appraisal*. Vol 8, No 2
- Douglas, M. 1987. *How Institutions Think*. London: Routledge & Kegan Paul.
- Dutton, J. 1986. Understanding Strategic Agenda Building and Its Implications for Managing Change. *Scandinavian Journal of Management Studies*. August 1986, Vol 3, Number 1, pp 3-24.
- Eccles, R.G. and Nohria, N. 1992. *Beyond the Hype: Rediscovering the Essence of Management*. Boston: Harvard Business School Press.
- Forss, K. 1985. *Planning and Evaluation in Aid Organizations*. Stockholm: EFI/IIB.
- Frost, J.P. et al 1991 *Reframing Organizational Culture*. Newbury Park: Sage
- Geranmayeh, A. 1992. *Organizational Learning Through Interactive Planning: Design of Learning Systems for Ideal-seeking Organizations*. Ph. D. Thesis. University of Pennsylvania.
- Gordenker, L. *International Aid and National Decisions*. London: Blackwell.
- Janowitz, M. 1959. Changing Patterns of Organizational Authority. *Administrative Sciences Quarterly*, 3: 473-493.
- Kelly, G. 1955. *The Psychology of Personal Constructs*. New York: Norton.

Langley, A. 1989. In Search of Rationality: The Purpose behind the Use of Formal Analysis in Organizations. *Administrative Sciences Quarterly*: Vol 34, No 4.

Leviton, L. C. and Hughes, E. F. X. 1981. Research on the utilization of evaluation: A review and synthesis. *Evaluation Review*, 5, 525 - 548.

Maclean, N. 1992. *Young Men and Fire*. Chicago: University of Chicago Press.

March, J. G. 1991. Exploration and Exploitation in Organizational Learning. *Organization Science*: Vol 2, No 1, pp 71-87.

Meacham, J.A. 1983. Wisdom and the context of knowledge. In Kuhn and Meachan (eds) *Contributions in human development*. Basel: Karger.

Meyer, J. and Scott, R.W. 1983. *Organizational Environments: Ritual and rationality*. Beverly Hills: Sage.

Mitroff, I. 1983. *Stakeholders of the Organizational Mind*. San Fransisco: Jossey- Bass.

Mosley P., Harrigan, J. and Toye, J. 1991. *Aid and Power: The World Bank and Policy-based lending*. London: Routledge.

OECD. 1986. *Methods and procedures in Aid Evaluation*. Paris.

Patton, M.Q. 1982. *Practical Evaluation*. Beverly Hills: Sage.

Perry, C. 1986. *Kognitiv Teori i Praktik och Teori*. Lund: Natur och Kultur.

Piaget, J. 1976. *The Essential Piaget*. Eds Howard Gruber and Jacques Voneche. New York: Basic Books.

Riksrevisionsverket. 1988. *Lär sig SIDA ?* Stockholm: Riksrevisionsverket.

Samset, K., Forss, K. and O. Hauglin. 1993. *Internal learning from Evaluations and Reviews*. Evaluation Report 1.93. Oslo: Ministry of Foreign Affairs.

Schein, E. 1985. *Organizational Cultures and Leadership: A Dynamic View*. London: Jossey-Bass.

Scott, W.R. 1987. *Organizations: Rational, Natural and Open Systems*. Englewood Cliffs, NJ: Prentice-Hall.

Starbuck, W. H. 1983. Organizations as Action generators. *American Sociological Review*., Vol 48 (February: 91-102)

Stokke, O. 1992. *Evaluating Development Assistance: Policies and Performance*. London: Frank Cass.

Waldrop, M.M. 1992. *Complexity: the emerging science at the edge of order and chaos*. New York: Touchstone.

Weick, K. E. and Roberts, K.E. (1993) *Collective mind in organizations*. *Administrative Sciences Quarterly*: Vol 38, 357-381

Weick, K. E. 1993. The collapse of sense-making in organizations. *Administrative Sciences Quarterly*: Vol 3, 628-651

Weinshall, T.D. 1993 *Societal Culture and Management*. berlin: de Gruyter.

Weiss, C. H. 1977. Research for policy's sake: The enlightenment function of social research. *Policy Analysis*, 3, 5331 - 545.

Williams, D. 1987. *The Specialized Agencies and the United Nations. The system in crisis*. New York: St Martins Press.