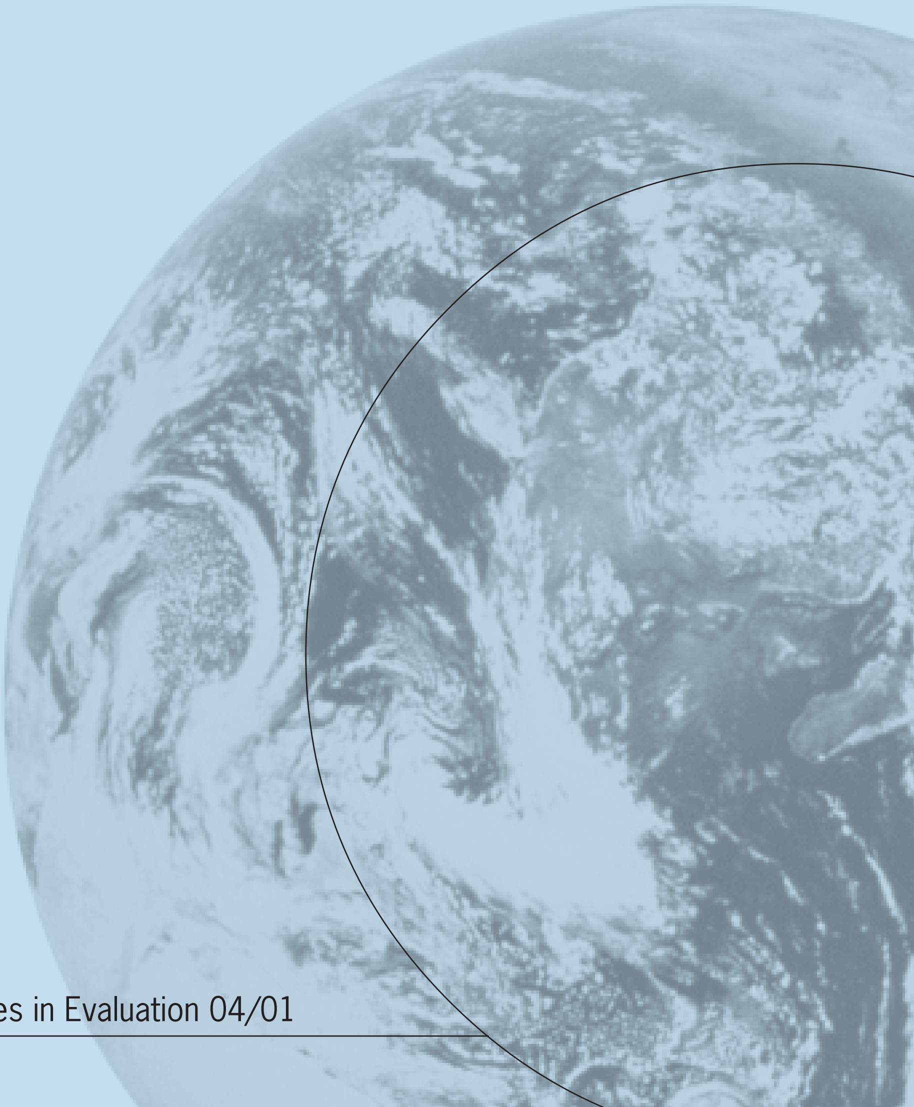
Stronger Evaluation Partnerships

The Way to Keep Practice Relevant

Gus Edgren



Sida Studies in Evaluation 04/01

Stronger Evaluation Partnerships

The Way to Keep Practice Relevant

Gus Edgren

Sida Studies in Evaluation 04/01

This report is published in *Sida Studies in Evaluation*, a series comprising methodologically oriented studies commissioned by Sida. A second series, *Sida Evaluations,* covers evaluations of Swedish development co-operation. Both series are administered by the Department for Evaluation and Internal Audit, an independent department reporting directly to Sida's Board of Directors.

This publication can be downloaded from:
www.sida.se/publications

Author: Gus Edgren.

The views and interpretations expressed in this report are the authors' and do not necessarily reflect those of the Swedish International Development Cooperation Agency, Sida.

Sida Studies in Evaluation 04/01

Commissioned by Sida, Department for Evaluation and Internal Audit

Registration No.: 2001-3920
Date of Final Report: April 2004
Printed by Edita Sverige AB, 2004
Art. no. Sida4080en

ISBN 91-586-8476-X
ISSN 1402—215X

SWEDISH INTERNATIONAL DEVELOPMENT COOPERATION AGENCY
Address: SE-105 25 Stockholm, Sweden. Office: Sveavägen 20, Stockholm
Telephone: +46 (0)8-698 50 00. Telefax: +46 (0)8-20 88 64
E-mail: sida@sida.se.
Website: http://www.sida.se

Table of Contents

Abstract

It is a recurrent complaint in reviews of evaluation practice that the recipient countries do not readily assume ownership of project evaluations, and that they often see them as a burden rather than as a useful tool. The changing modalities of development cooperation, with a growing share of aid transfers being channelled through multi-donor budget and sector support programmes rather than through single-donor projects are also raising challenges to the conventional practice of evaluation. The very relevance of assessing achieved results compared to the originally stated goals of these programmes is being questioned by some of their managers. The present paper suggests that strengthening the partnership around the evaluation would help making it more relevant and would also increase its outreach.

Levelling power imbalances: Donor-recipient partnerships are most often based on vast differences in power, and on different access to information. There are ways of levelling these differences somewhat, particularly by establishing stable and predictable systems that generate basic trust. It is not difficult to see why recipients show less interest in evaluation than donors: more often than not, it is the donor who sets the agenda and who appoints the evaluators. The donor often professes exaggerated needs for accountability, with very little attention paid to the recipient's accountability requirements. By engaging the major stakeholders of the project – direct beneficiaries as well as agency officials and people who have been indirectly affected – it will be possible to increase both relevance and outreach.

The paper considers four examples of activities that will affect evaluation practice:

(1) PRSP and macro approaches, (2) Sector-wide programme approaches (SWAP), (3) Stakeholder participation and (4) Evaluation capacity development (ECD).

PRSP and macro approaches: Through its comprehensive coverage of social and economic development policy and the broadening participation of social partners, the PRSP process provides unparalleled opportunities for evaluating both the effectiveness and the process of development cooperation in a national perspective. So far, evaluations have mostly been limited to the requirements of donor agencies (particularly for the HIPC facility), and interest has focussed on growth and the poverty headcount. Evaluations that have been made suggest that new systems are being superimposed on old ones, and that there is a reporting overload.

Sector-wide approaches (SWAP): SWAPs and budget support programmes involving large numbers of donor and recipient agencies have developed systems for joint monitoring of expenditure as well as impact indicators. Those systems are less complicated and demanding than multi-donor evaluations, and there is a tendency for the latter to become less frequent. In com-

plex partnerships of this kind, evaluation may more effectively focus on processes than on products. Sida is recommended to concentrate its evaluations on cases where partnerships are strong and constructive, and to review its participation in programmes where the partnership does not work.

Stakeholder participation: Ownership and outreach is strengthened by building partnerships with beneficiaries and other interested parties. People generally learn more from participating in an enquiry than from reading its report, and participants in the evaluation should be chosen both with regard to what they can contribute and what they can learn. It is important to involve the stakeholders from the beginning, at the stage when the questions of the study are formulated.

Evaluation capacity development: It is important to build national capacity for evaluation in the host country, and not primarily in order to evaluate donor-funded projects. An independent capacity for development study and policy analysis is an important factor in levelling the power and information imbalances of the donor-recipient partnership.

Acronyms

ECD	Evaluation Capacity Development
ESRF	Economic and Social Research Council (Tanzania)
CDF	Comprehensive Development Framework
DBS	Direct Budget Support
HIPC	Highly Indebted Poor Developing Country
IFI	International Financial Institutions
IMF	International Monetary Fund
MDG	Millennium Development Goals
M&E	Monitoring and evaluation
NGO	Non-government organisation
ODA	Overseas Development Assistance
OECD-DAC	Development Assistance Committee, Organisation for Economic Cooperation and Development
OED	Operational Evaluations Department (World Bank)
PAC	Public Affairs Centre (Bangalore)
PAF	Poverty Action Fund (Uganda)
PEAP	Poverty Eradication Action Plan (Uganda)
PER	Public Expenditure Review
PRA	Participatory Rural appraisal
PRS	Poverty Reduction Strategy
PRSP	Poverty Reduction Strategy Paper
Sida	Swedish International Development Assistance Agency
SOAS	School of African and Oriental Studies (London)
SWAP	Sector-Wide Approach
UDN	Uganda Debt Network
UNDP	United Nations Development Programme
WB	World Bank

1. Background

In studies of recipient ownership of development cooperation activities, the activity that always comes out as most donor-driven is evaluation (Sida, 2003). Project identification, management and stakeholder relationships are operations that the recipient partner will readily assume responsibility for, if offered the opportunity to do so. But the evaluation function is often seen as serving the donor's interest of financial control and accountability and it finds fewer takers on the recipient side, even when the partnership is 'genuine' in the sense that both partners exercise the same degree of control over project activities.

This lack of interest among aid recipients is often a donor dilemma more than a general problem of aid effectiveness. Aid recipients are interested in making the best possible use of external resources, but their information needs are different, and their perspective on alternative uses of funds are also mostly different from those of the donors. Factors outside the scope of the project will have to be taken into account, and the questions to be examined in an evaluation study might be very different from those formulated in the traditional evaluation practice of a donor. Developments in this direction are already seen in some countries, driven by new aid modalities and by partnerships where many donors pool their resources in support of the same programme. Other forms of feedback have become more common than proper evaluation of outcomes in relation to targets, and some programme managers have begun to question the relevance of conventional evaluation practice.

The present paper reviews some current practices of evaluation from the points of view of the donor-recipient relationship and attempts to examine how those practices might be affected, if their focus were to be chosen by recipients rather than by donors. We shall try to examine this question in the context of some emerging modalities for donor-recipient cooperation, such as various approaches to programme aid, participatory evaluation and evaluation capacity development. A key question is, to what extent stronger partnerships can make evaluation practice more relevant to the needs of both sides.

2. The partners' purposes and interests

Aid evaluation often serves a number of different purposes, some of which are closer than others to perceived donor interests. The following three purposes are of particular interest in this context:

1) *Accountability:* The donor agency has an interest in showing the providers of funding (government, audit bodies, parliament) that the money is properly accounted for and that the intended development results have been attained, or have a likelihood of being attained in a near future. The recipient agency has a similar interest in accountability at least in cases where good results may lead to continued resource flows.
2) *Implementation feedback:* Both partners normally have an interest in following the implementation of the project and in getting information that can guide their management towards a successful outcome. This purpose is served by day-to-day monitoring of activities, resource use and results, and it is also underpinned by recurrent evaluations of overall outcomes compared to original plans and targets. Feedback can be used to change the direction of the activities or the shape of the inputs to make implementation more effective.
3) *Strategic learning:* Evaluation will generate some knowledge that transcends the scope of the project. Methods and approaches will be tested that can be applied elsewhere if they work, or rejected if they don't. The partners may enjoy a collective learning experience which will enable them to work better together in a wider variety of areas. A series of systematic evaluation studies may provide the base for major revisions of policy on both sides.

It is pertinent to discuss a bit further the interests of donor and recipient in different purposes, or expected outcomes, of evaluation. As regards *accountability*, the donor agency is certainly accountable for the use of the funds to those who have provided them. But most of the agencies have rather wide latitude for defining how the accountability requirement should be met. Although their principals in parliament committees, ministries of finance or foreign affairs are not known to be avid readers of their evaluation reports, donor agencies more often than not prefer to err on the high rather than the low side, and many reports that are chiefly produced for the purpose of accountability remain virtually unread. A particularly serious case of wasted resources is when evaluation has been made a mandatory requirement for allocating funds for prolonging a project. Such a formal way of treating the accountability requirement is sure to reduce evaluation to little more than an administrative nuisance.

The recipients are also accountable for the funds they are using, particularly in a narrow administrative sense and to their own auditors. But politically, they may be accountable for different things, and to different principals. Political officials usually have wider responsibilities than what is covered by one project or one donor relationship, and they must manage each project so that the outcome satisfies the whole remit, not only the project. There are also cases where officials are accountable for loyalty to their superiors rather than for their integrity and effectiveness, which could lead to wide divergence in the interest of the partners in determining the remit.

There are also measurement problems connected with aid projects, insofar as it is often impossible for the recipient to assess their true opportunity costs. If the health ministry has managed to attract an aid project from a certain donor, its political principals may well think that the opportunity cost is zero, if the funds might otherwise have gone to other sectors, even to other countries. Funds that are untied and fully fungible, on the other hand, have several alternative uses and are therefore rated higher. For this reason, the ministry's accountability to Parliament or to the Cabinet may diverge sharply from that of the donor agencies, whose opportunity cost accounting is based on entirely different alternative uses of the funds. These differences will also be reflected in the partners' views of accountability.

In the case of *implementation feedback*, it is easier t o find common ground for the partners' interest in evaluation as well as monitoring. They both want to keep track of project outputs, in terms of quantity as well as quality, and they are usually prepared to adjust implementation to reach a higher rate of success in relation to agreed targets. But differences often arise in drawing operational conclusions from the feedback. The partners may have agreed on action in order to achieve certain specific results, but they may not have a commonality of interests as regards alternative lines of action, let alone the policy environment required to facilitate them. Beyond the bare essentials of operational data, the partners may find themselves in deep disagreement about what information is required, and for what purpose. These differences may be caused by factors on either side of the partnership:

Donor side

- Donor controls choice of evaluators
- Certain project inputs are regarded as 'givens', e.g. expertise and equipment
- Scope of study formulated by donor, excludes factors of interest to recipient agency
- Timing and methodology impractical to staff and work process of recipient agency
- Feedback is limited to selected groups of stake-holders

Recipient side

- Insufficient capacity for managing and responding to all donor-driven evaluations

- The recipient agency regards findings to be of limited use outside the project
- Lack of transparency, fear of criticising senior management
- Management systems that are unable to feed information back to decision-makers

These differences reflect not only different management cultures but also diverging interests based on the economic and political power structure on either side. Country situations differ so widely that it is impossible to generalise, but it may still be fair to suggest that operational data will be the least controversial and that differences increase the closer the evaluation remit moves into policies and the effectiveness of senior management (Sida, 2004, Chap. 2).

Similar differences between the interests of the partners occur in the case of *strategic learning* from evaluation. Donors often complain that recipient agencies do not make use of the knowledge generated by evaluations, while recipients complain that donors make them undertake too many project-specific evaluations, from which agency staff have great difficulties drawing any generally useful conclusions. However, important changes have emerged since the mid-nineties by the increasing use of sector-wide approaches (SWAP) and budget support, aid modalities that have given rise to new approaches to evaluation, in which the strategic learning elements are more important than the study of operational efficiency. There may still be a divergence between the interests of the donors and the recipient, but the configuration of these multi-donor programme evaluation exercises will make it easier to minimise some of the problems listed above. The increasing involvement of stakeholders at several stages of the project cycle has also been important, both by widening the circle of learners and by introducing a second "loop" to the learning curve by questioning the development process as a whole (Carlsson and Engel, 2002). These tendencies will be further examined in section 4 of this paper.

In conclusion, it is not difficult to explain why recipients find the common practice of evaluation less useful than the donors. Some of the reasons are related to factors like low capacity and lack of transparency, which are weaknesses that development cooperation seeks to remedy, but many of them are legitimate concerns of any government agency, in developed countries as well as in the developing world. Before we start discussing the nature and expressions of partnership, it is important to note that the topic and remit of evaluation should as far as possible be chosen to satisfy the interests and needs of both recipient and donor.

3. Deepening partnerships by levelling imbalances

From the late 90s on, current jargon refers to the aid relationship as a partnership and to donors and recipients as partners. To a large extent, this change of terminology is cosmetic, since the terms are indiscriminately used to cover all sorts of aid relationships regardless of power structure, transaction rules and relational dynamics. There are no standard definitions of partnerships in the aid industry, although it is generally recognised that the partners should try to achieve "mutually agreed objectives involving shared responsibility for outcomes, distinct accountabilities and reciprocal obligations" (Picciotto, 1998). When we are trying to direct the attention to a type of partnership that we describe as "genuine", we are not widening the definition but deepening it, by addressing two *structural* weaknesses in the conventional aid relationship, namely first, the power imbalance between donor and recipient, and second, the very narrow stakeholder involvement in the process as a whole.

Mutually agreed objectives: It is of fundamental importance to the outcome of a partnership undertaking, that the parties agree not only on operational objectives, but also on the broader development objectives. It is often easier to agree on activities that need be undertaken, like building a road or launching a training programme, than on the type of institutional or policy framework that these interventions should support, and what broader development goals they should be aiming at. The recipient is often prone to play down reservations against the donor's medium or long-term agenda in the interest of receiving the initial investments of the agreed project. But a collaboration agreement that tries to paper over deep-going differences over development objectives may degenerate into an acrimonious relationship when the next steps are to be taken. Genuine partnerships are based on a common vision of where development should lead in the longer term, and whom it should benefit.

Shared responsibilities for outcomes: In private business partnerships, the division of responsibilities for producing results is as important as the sharing of profits. Development cooperation agreements are mostly precise in regulating the inputs to a project or programme but they are often less precise as to what should happen if the expected outcomes are not produced. There is an element of power imbalance behind this vagueness, which will be dealt with below. But there are also tendencies in the current efforts to find more flexible forms of cooperation to postpone a more definitive division of responsibilities until a later stage, particularly in pilot projects or in using a 'process' approach of trial and error. In fact, mutual interaction through experimentation and problem-solving is an experience which could build genuine partnerships. The practical advantage with such approaches is that they do not require a long period of planning and preparation, and the step-by-step approach may also reduce the risk of making big planning mistakes. To avoid instead making big operational mistakes, however, it is necessary to support the process approach by constantly revising operational targets and calibrating them with long term development objectives. In this way, it will also be possible to use a process approach to evaluation (Forss and Rebien, 2002).

Distinct accountabilities: We have already mentioned some reasons why accountabilities are often unclear or lopsided in aid projects. The donor agency usually overemphasises its accountability to its own principal (government, tax payers, media) and finds it difficult to see itself as accountable to users and stakeholders on the recipient side. The recipient agency often looks at accountability in the strictly bureaucratic context of financial accounts and adherence to formal agreements. If a 'genuine' partnership is to be built, the partners must accept accountability also to the stakeholders of the project, be they users, beneficiaries or people who risk being negatively affected by the development activities. If there is a 'deficit' in the way the partners see their accountability in relation to any of the objectives or stakeholders of the project, this may lead to serious disagreement further down the road. And it will certainly not provide a good base for joint monitoring and evaluation of the relevant outcomes.

Reciprocal obligations: A good deal of the donors' renewed interest in the partnership concept seems to be explained by a feeling that conventional aid relationships were lacking rigour in keeping the partners to reciprocal obligations. There are worries, particularly among the NGOs, that the drive towards more clearly defined partnership obligations is motivated by attempts to provide a more compelling framework for the conditionalities of the IFIs (Fowler, 2000). The question of reciprocity is strongly affected by the power relationship between the parties, which will be dealt with below. It is also closely related to the role of stakeholders: a negotiated process involving a number of different stakeholder groups could produce an agreement on reciprocal obligations that is understood and accepted by all participants in the project, and hence more likely to be adhered to. What would undermine a partnership is first, if one of the parties is free to change its commitments at will in the course of the collaboration, and second, if there is uncertainty as to what will happen if contractual obligations are not fulfilled (See box.) In evaluating outcomes it will of course be difficult to establish responsibilities if reality on the ground differs too much from what has been formally agreed.

The give-and-take of partnerships explained by game theory

Robert Axelrod has explained the donor-recipient relationship through a simple game-playing model, where the donor and the recipient each have only two choices, to *cooperate* or to *defect* (e.g. by avoiding to fulfil their obligations). The basis of the partnership is that if both cooperate, both will benefit. But the temptation of either side is that if only the other partner continues to co-operate, there will still be benefits for the non-cooperating side. If defection becomes increasingly common, however, the game will deteriorate and the benefits will eventually disappear.

The game strategy that Axelrod finds will give both sides the best returns is called 'Tit for Tat'. It entails cooperating on the first move, and then doing what the other side does on its previous move. If this strategy is applied in a series of games, it will benefit the two players even more, since they can foresee each other's moves.

A very interesting finding which is relevant to all partnerships is that it is necessary for each partner to respond to provocation – the sooner one responds to a defection on the other side, the easier it is to restore mutual cooperation by establishing that defection does not pay.

Source: Axelrod, R., *Theoretical Foundations of Partnerships for Economic Development.*

The power imbalance: The biggest obstacle to genuine partnership is the imbalance in power between donor and recipient. The main base of this imbalance is resources: the donor is prepared to provide resources that the recipient wants, and even if he does not make all his conditions and motives explicit at the outset, the recipient is at a disadvantage in seeking to gain access to those resources. It is possible to neutralise part of this disadvantage in the short term by making the resources available to the recipient under an aid agreement. But the snag is that the recipient wants the donor to 'stay in the game' and to continue to provide resources even after the end of this agreement, and as also suggested by Axelrod's game theory, it takes a long series of such agreements to establish the kind of *mutual trust* required to compensate for the formal uncertainty. A number of proposals and innovations have been put forward for donors to pool their resources and make long-term commitments that would at least partially mitigate the power imbalance (Kanbur and Sandler, 1999), but power imbalance is a structural flaw in aid relationships which needs to be kept constantly under examination.

The asymmetry in resource endowment is also reflected in the *asymmetry of information*. The most important factor in this regard is the donor's exclusive knowledge of the scope and extension of his own resource commitment, as explained above. There are other information sources that are unequally available. The donors have an advantage in analytical capacity and access to global data sources, while recipients are most often limited to secondary sources. The information imbalance is sometimes as important as the inequality of resource endowment, not least since it may lead the weaker partner to make expensive mistakes. The recipient, on the other hand, is often better informed about the behaviour and interests of other donors, and he will sometimes try to use this advantage by playing one donor against the other. The tendency towards multi-donor cooperation in sector and budget support at macro level (see section 4.2) will hopefully mitigate the information asymmetry to some extent, even though it cannot entirely counterbalance the disparities in analytical capacity between the partners.

Stakeholder participation: Agreements based exclusively on the two negotiating government agencies will provide the base for a partnership, but it is a base that is narrow and vulnerable to changes in the institutional or political environment. It is generally accepted that partnerships will be stronger and more sustainable if they in some way involve stakeholders on either side, from NGOs and political support groups on the donor side to users and beneficiaries on the recipient side, including organisations representing civil society. What is not so easy to agree on, however, is which ones of those interests that should participate, and what influence they should have over the formulation, implementation and evaluation of the proposed activities. Some stakeholder groups are seen by the government as vested interests, that should be curbed or at least moderated by government in the public interest. Some of the civil society organisations may be dominated by political parties or movements that the government regards as disloyal political opponents. The closer stakeholder groups of any kind get involved in the decision-making process, the less control will the government be able to exercise over it.

There is no set piece for engaging stakeholders in a partnership. Sometimes, like in the case of a project to computerise a central government department, thorough stakeholder involvement may contribute rather little to strengthening the government-to-government partnership. In other cases, the participation of users, beneficiaries or civil society will make project formulation more relevant and implementation more effective. It may also help giving local political or material support to the project which is useful if it is threatened by political forces or by budget cuts. Where the government partners worry about the risk of losing control over a project, they can use a step-by-step approach beginning by informing the stakeholders, then gradually extending their involvement to decision-making. If this is done in a spirit of genuine partnership, the participation of stakeholders will soon take on a life of its own.

The practice of evaluation will also be greatly enriched in a partnership with participation of stakeholder groups. Stakeholders and beneficiaries will raise their own questions of accountability and effectiveness, and they are part of a collective learning experience that can be strengthened and better utilised through systematic evaluation methods. To achieve such results, however, evaluators will have to adjust the focus of their enquiry to take in what is of interest to stakeholders on the recipient side.

4. Examples of changing evaluation practice

The last ten years or so have seen important changes in the way development cooperation works in many countries, in particular in the way donors are co-ordinating their interventions, in trying to find more flexible modalities for linking resource transfers to outcomes, and in encouraging recipients to take leadership of these processes. These changes may still appear tentative and partial, but in some recipient countries they are likely to have significant impact on how aid transactions are conducted in the future. Evaluation practice is only slowly adjusting to these new trends, having some problems liberating itself from its past in conventional project aid. In the following, some examples will be given of areas where stronger partnerships could improve the managerial feedback as well as the collective learning experience.

4.1. PRSP and Macro Approaches

The Poverty Reduction Strategy Paper (PRSP) was approved as a new programme instrument by the boards of IMF and WB in December 1999. It was originally conceived as a document that would commit beneficiaries of debt relief from IMF and WB under the HPIC Initiative to a strategy of poverty reduction with growth. The strategy was supposed to be comprehensive in scope and inclusive as regards the process of formulation, involving a wide range of political and civil society stakeholders. The PRSP immediately became the centre of attention and development dialogue in the countries that pioneered it (Uganda for example), and a large number of poor countries drew up their own PRSP, even those lacking the prospect of debt relief. Like the comprehensive development plans in the 1970s, the PRSP is central to the dialogue between the government and its development partners.

Although serious efforts were made by all partners to make the PRSP comprehensive and poverty-focused and to include major stakeholders in the process, the first PRSP round was marked by numerous weaknesses. Many of the development strategies appeared to be traditional growth strategies with anti-poverty element grafted on to them, and the involvement of NGOs, civil society and academe was often superficial, leaving some stakeholders dissatisfied with the process (UNDP, 2003). But the exercise also demonstrated the potential strength of this new tool in focussing the dialogue and building partnerships. By raising the question of how all efforts of society with the help of donors can more effectively reduce poverty, the PRSP could not only set the agenda for development cooperation but also provide a framework for examining results and for collective learning from the experience of implementing the strategy.

The institution of PRSP with its combination of socio-economic analysis, programme design and collaborative framework would be an ideal instrument for broadening the perspective of aid evaluation. Development analysts have often criticised aid projects for being 'islands of development' in a sea of despair. The PRS nexus invites partners not only to cooperate to bring about results, but also to review the effectiveness of their joint efforts in a national perspective. Four dimensions of the PRS undertaking deserve particular attention:

- *Design of the strategy:* How effective is the overall design of the strategy with respect to its objective to reduce poverty within existing constraints and opportunities?
- *Implementation:* To what extent do the development outcomes correspond to what was foreseen by the strategy?
- *Cooperation process:* Have the government and its donor partners undertaken all efforts to engage civil society, the private sector and the academic world in the cooperation process?
- *External resources:* How effective have the donors of external resources been in supporting the strategy?

In the short period since the first launch of PRSPs, evaluation of outcomes has primarily been focused on the implementation dimension, and in particular on growth and poverty headcounts. This is quite understandable, since the PRSP is an operational document linked to the release of IMF and WB funds. But there is no reason why evaluation should stop there, since the objectives of the strategies entail not only far-reaching policy reforms but indeed transformation of entire societies. Hence, there is no reason why reporting on outcomes should be so limited and be mainly directed at external agencies (UNDP, op. cit.). A great number of partners support the effort in the interest of the much broader agenda of the PRS strategy as a whole. The whole range of issues mentioned above are of interest to the partners who support the strategy, and it is appropriate that those partners themselves be involved in the evaluation.

Comprehensive evaluation of Uganda's poverty reduction strategy

A comprehensive evaluation of monitoring and evaluation in Uganda's PRS exercise (called PEAP) has given examples of evaluation partnerships and of the potential uses. The evaluation, which was undertaken by WB-OED, reviewed both substantial outcomes and processes. It concludes that the availability of data has significantly improved as well as the analytical level, and that analysis is largely 'owned' by Ugandans. Sector working groups meet regularly, and many ministries have research units that carry out their own analysis of outcomes. Monitoring of expenditure has improved, and the budget system has become more transparent. Civil society organisations like the *Uganda*

Debt Network (UDN) coordinates civil society monitoring of PAF activity in the districts. UDN, an independent NGO often critical of government and the donor community, is also piloting a community- based monitoring and evaluation system, which mobilizes the public to verify expenditures and service delivery. Moreover, a client scorecard system, initially for piloting in the health sector, is currently under preparation to make budget allocation to the Districts known in the locations, since it was found that leakages of public funds were reduced when the public was better informed. Research centres attached to the University are engaged in policy and impact studies related to the anti-poverty programme. External support has been given for building national capacity for evaluation.

The OED evaluation finds a number of positive outcomes in improving management systems, accountability and feedback, but it criticises the donors for not allowing the new information systems to *replace* the older ones – in many cases both are continuing to function in parallel, increasing the workload where rationalisations would have been possible. The incentive systems in M&E are often tied to the formal conduct of reporting rather than to its usefulness, which deflects rather than strengthens attention to service delivery. The study recommends a thorough review of the plethora of reporting systems used by Uganda's external partners.

Source: Hauge, A. *The Role of Monitoring and Evaluation Capacity Development in Improving Government Performance*:*Lessons Learned from the Uganda Poverty Reduction Support Process,* WB-OED, August 2003.

An aid modality which is of particular relevance to macro policy reviews is *direct budget support (DBS)*, through which several donors pool their resources directly into the government budget. DBS has become increasingly common in Africa, mainly as an attempt to overcome the ownership and transaction cost problems of project aid. This aid form is unpopular among the public auditors in donor countries, since it makes it impossible to trace the movements of the donor dollar or *krona* from the Finance Ministry to the designated beneficiary, and it limits the discussion of effectiveness and efficiency to macro level indicators like economic growth and the distribution of public expenditure. To compensate for this loss of control in comparison with project aid, donors often condition their budget support on a vast array of public expenditure reviews (PER) that are supposed to ensure that aid money is not siphoned off for non-developmental purposes. In some countries like Tanzania, the institutional framework built around the PER has brought the donors right into the national budget process, which raises some interesting questions of accountability (Odén and Tinnes, 2003).

From the point of view of evaluating development results, DBS shifts the focus of enquiry from activities and sectors to the performance of the national economy and of the impact of the government budget as a whole. Most of the attention has so far been directed at the structure and timing of public expenditure, a form of input monitoring. Assessing the effects of outputs, however, is a task that requires economic and social research rather than ad hoc effectiveness studies of the type normally undertaken in project evaluation. The consequences of this requirement for national research capacities will be dealt with in section 4.4.

Evaluating a development strategy in terms of its consistency and its ability to reach the goals with expected policies and resource inputs would be of interest outside the narrow group of official partners. The academic community in the country could contribute a valuable dimension to this discussion, working in partnerships of their own choice with institutions and researchers outside the country. Not all governments welcome critical scrutiny of their strategies, but a large number among the first round participants in drafting PRSPs are aware of the advantages of such a discussion. When it comes to the implementation dimension, evaluation is constrained by a paucity of relevant and reliable data. But the current work to relate national information sources to the targets adopted for the Millennium Development Goals may lead to increased interest in measuring progress, as well as to a broader participation of stakeholders in validating and analysing those data sources.

The process of elaborating, implementing and following up the poverty reduction strategy is not only important for gaining support and understanding of it, but it is in itself an indicator of the extent to which various actors take part in making decisions about development. NGOs, the private sector and civil society spokespersons often have strong and well-considered views about their roles in the PRS process, and they should also be invited to take part in its monitoring. OED's evaluation of Uganda's PRS experience (see box) gives a number of positive examples of this could work, while at the same time warning of the risks that new information systems may be superimposed on old ones, thereby creating confusion and double work. The evaluation also expresses concern over the tendency of government incentive systems to reward the production of reports rather than the use of their feedback for improving services.

Recipient governments are mostly reluctant to openly criticise their donors, even when donor policies create serious problems. Nevertheless, a critical review of the donor side is necessary to take in the whole picture in evaluating the PRS efforts. In 1995, Gerry Helleiner proposed a system of independent performance monitoring of Tanzania's donors, which was introduced by the Government in 1997 and has produced very frank and pertinent advice to both sides at the Consultative Group meetings (ESRF, 2002 and Helleiner, 2000). This approach holds a lot of promise for recipient countries that maintain an open and constructive relationship to their donors. It is possible that OECD-DAC will contribute to its replication as an instrument for its attempts to harmonise donor procedures, policies and practices.

Robert Picciotto has gone a step further by suggesting that all policies of the donors should be taken into account in this evaluation, not only with regard to ODA (Picciotto, 2003). The abolition of protectionist trade policies would often be more important than ODA in supporting the poverty reduction efforts of a given low-income country. Millennium Development Goal number 8 on developing a global partnership for development may open up for such discussions, both globally and at the national level. But it is hard to find much interest in such critical analysis among governments or in the multilateral organisations, although OECD-DAC and WB have made attempts at developing indicators for monitoring donor behaviour (Holmgren and Soludo, 2002). More likely, the academic community would be the best starting point for such a discussion.

4.2. Sector-Wide Programme Approaches (SWAP)

Sida's East African Ownership Study (SOAS 2003) was convinced that programme approaches like SWAPs will make it easier for recipient governments to assume ownership than conventional project aid, provided that both sides of the partnerships shared the values and vision guiding the programme. The flexible links between funding and operations were cited as the main factor that would facilitate recipient ownership and level the base for partnership between donor and recipient. When recipient government agencies have full control of both organisation and budget without being constrained by a plethora of donor earmarkings, this would mark a step in the direction of more 'genuine' partnerships. From the point of view of aid evaluation, the following differences between project and programme aid are of particular interest:

- Institutions and policy environment are emphasised over specific activities
- Outputs are described in terms of broad aggregates and indicators of development impact rather than as the results of specific activities
- Resources are provided for the programme by several donors, without strict activity earmarking
- Responsibility for managing the programme is vested in a line organisation like a ministry rather than in a special project unit
- The programme uses the government's reporting systems rather than the ones of individual donors

In practice, the difference between SWAP and project aid is often less clear. What passes as SWAPs are in many cases hybrids of project and programme aid, with some elements of resource pooling and some of activity earmarking. The objective of replacing the donors' different reporting and accounting systems with a uniform government standard is rarely met, as evidenced in the box on Uganda's PEAP. And the reporting and technical support requirements that those partnerships impose on the implementing agencies are so vast, that the government's thinly stretched expertise is placed at a serious disadvantage in relation to the hordes of experts that the donors can provide for sector reviews and policy discussions (Ostrom et al, 2002). To make the job of evaluators even more complicated, different implementing government agencies may disagree on objectives or methods to a degree that makes evaluating outcomes highly controversial. An example is the education sector programme in Tanzania, which the SOAS study characterised as a "strong but fragmented ownership" (SOAS, op. cit, Tanzania country report). The brave new world of SWAPs does not offer the evaluators any easier task than that of conventional project aid.

The first feasibility test of a partnership for evaluating a SWAP would be if all potential partners feel the *need to evaluate* the programme and if so, what *purposes* do they want such an exercise to serve. Where partnerships are weak, unbalanced or fragmented, it may turn out that different partners have different objectives, interests and motives as exemplified in Section 2. The Uganda example shows how implementing ministries may accept too many different and overlapping studies that contribute very little to increased understanding of the development problems but adds to the administrative burden of an overstretched ministry.

SWAP evaluations are typically *multi-donor evaluations.* In a game where donor and recipient have equal power, multi-donor situations could be to the advantage of the recipient, who would be able to play one donor off against the other. Such situations might occur in countries like India or Brazil, but the power imbalance in poor, aid-dependent countries is very different. Bringing all the donors together around one evaluation agenda is sometimes tortuously difficult. Even the World Bank that commands so much financial clout has found it difficult to achieve genuine multi-donor partnerships for evaluating a large programme. Feinstein and Ingram (2003) find that the WB has participated in very few integrated multi-donor evaluation exercises, and even where such collaboration was tried, some donor agencies still had to report separately to their principals, using their own reporting format. Joint evaluations tend to reduce the transaction costs of recipients, since they take place at one time rather than several, but they certainly do not reduce the costs to the donors, because of all the special interests and reporting requirements that are maintained in the professed interest of accountability. The problems of finding common ground for evaluating multi-donor SWAPs could even become a serious disincentive to undertaking joint analytical studies.

Where implementing government agencies and their political principals have a strong interest in learning from the collaboration experience, it is important that the donors *let the government lead* in determining the focus of the enquiry. This is the only sustainable solution to the problem of escalating and conflicting donor demands for data and studies. Both donors and recipients must be aware that in terms of the game they are all involved in, they would all lose in the long run if they allow a single donor to make short-term gains at the expense of everyone else. From the point of view of strengthening the partnership it would be more constructive if the partners could *start with the requirements of the host government agency* and carefully work out what additional questions could be added without tipping the balance against the recipient's interests.

Since objectives and targets are formulated in a broad sector perspective, *success indicators* in a SWAP will often be defined at a high level of aggregation, like primary education enrolment at national level or reduction of the prevalence of malaria. These indicators are indeed more relevant for the analysis of development impact than most project-level data, but many donors see a problem of attribution in not being able to relate their own contribution – or even the external resource flow as a whole – to specific outcomes. A joint evaluation will assess the outcome of the country's achievements, not the effectiveness of individual inputs. Some of the major outcomes of SWAPs consist of policy changes and institutional development, and it is even difficult to measure them in conventional categories. An assessment of the government's progress against macro indicators would not require a proper evaluation – it can more easily be achieved through macro studies and review meetings without much analytical work regarding the effectiveness of specific inputs, policies and methods – so the annual sector review could eventually replace analytical evaluations as the standard source of feedback. The increasing attention that is being paid to Public Expenditure Reviews (PER) also contributes to shifting the focus of the development dialogue from assessing outputs to the monitoring of inputs and activities.

One way around the problem of assessing results in SWAPs is to evaluate *processes* rather than products. Instead of formulating their objectives in terms of quantitative outcomes, the partners could formulate principles and criteria for the change process, and evaluate it through a combination of self-evaluation by participants and independent consultants as it evolves. Since the expected outcomes of a SWAP are often defined as policy changes or institutional development, the process approach will often give more useful feedback than quantitative data and attempts to measure end results. As Forss and Rebien point out (op cit, 2002), even in a conventional results-based evaluation there may be more learning before and during the evaluation than after it. If those who are closely affected by the process take part in the evaluation, they are likely to learn more from the exercise than they would from reading a report. One might even be able to do away with reporting entirely, thereby avoiding the problems of different reporting formats for different donors.

A conclusion that can be drawn from this review of SWAPs is the necessity to apply *selectivity* to the choice of evaluation approach. Far from all SWAPs are built on 'genuine' partnerships, either seen from the point of view of power sharing or from the point of view of interest in learning from experience. Only where a group of like-minded donor- and recipient agencies have developed a sufficiently strong and well-balanced partnership around some explicit common objectives will it be possible to use joint evaluation effectively for collective learning purposes. Where the partnership is weak and fragmented, a concerned partner might do better trying to remedy the partnership before pushing for an evaluation.

4.3. Stakeholder participation in monitoring and evaluation

Participatory approaches to development cooperation would by association be assumed to generate strong partnership and hence be of interest from the point of view of evaluation. Both participation and partnership are slippery concepts, however, which are difficult to define operationally and, in addition, they are frequently misused in political rhetoric. Moreover, although participatory evaluation has been eloquently advocated by an impressive number of theoretical studies, it is hard to find much evidence of its practice (OECD-DAC, 1997). Trying to find the potential for partnership development in participatory evaluation may easily turn into a red herring expedition. We shall approach the subject with an appropriate degree of caution.

We have given a general definition of partnership based on common purpose and shared responsibilities, with a scale of deepening defined by measures of equality, transparency and shared responsibility for resource inputs as well as outcomes. In a similar way, participation can take different forms and degrees of intensity. Arnstein (1971) defined 8 degrees of participation, depending on the degree of power exercised by citizens over the activity. 'Partnership' was included as one category of 'citizen power', at third level after citizen control and delegated power. The descending scale included various forms of 'tokenisms' including consultation and information sharing, and ended with varieties of non-participation like outright manipulation. This taxonomy is particularly interesting since it provides an interface between participation and partnership: partnership is a form of participation that allows the citizen a certain degree of control over the activity in which she takes part.

The advantages in using participatory forms of evaluation are several. By involving at an early stage all those who would have a stake in the project – positive as well as negative, expected beneficiaries as well as those whose interest may be threatened – both the donor and the recipient agency will be better informed of how the stakeholders see their problems and how they would like to see them tackled by the project. By giving them responsibility for implementing project activities, they will ensure transparency and public support, and by inviting them to take part in evaluation and monitoring they will achieve a higher degree of feedback and learning than if this function was handled entirely by external consultants. Disadvantages of this approach are risks for delays (upsetting the timing of the evaluation cycle) and for letting the activity get bogged down by political controversies or vested interests. It is generally believed (although evidence is scarce) that the hard-won negotiated outcomes of the participatory approach are more sustainable in the long run.

Where partnerships can make a difference in this process is precisely by offering participants a degree of control ('empowerment') over activities that are supported by the project. This could be done through elected bodies and committees, or by contracting out the responsibility for certain functions to organisations, enterprises or elected individuals. These arrangements should be institutionalised within the framework of the project, so that they play a part in all functions, including that of evaluation. Such 'empowering' arrangements are particularly important for projects that are aiming to reach groups that are disadvantaged in their communities like women and marginalised farmers. Targeting in these cases is often done through Participatory Rural Appraisal (PRA), where the whole community will be involved in selecting the beneficiaries of the intervention (Chambers, 1992).

Advocates of participatory evaluation (e.g. Rebien, 1996) argue that participation also increases the possibility of *double-loop learning*. According to this theory, conventional evaluation is carried out in three steps: (1) gathering information on the intervention and its surroundings, (2) comparing achievements against stated goals and (3) initiating action to correct activities in order to be able to achieve stated goals. When beneficiaries and implementing staff participate in the evaluation, this will enable them to examine not only the activities but also goals and basic assumptions of the project, and hence to be able to criticise and adjust the very foundations of the intervention, which represents a second loop of the learning process. Narayan-Parker (1991) has given a very pedagogic summary of the way she perceives the differences between conventional and participatory evaluation:

	Conventional Evaluation	Participatory Evaluation
Who	External experts	Community people, project staff, facilitator
What	Predetermined indicators of success, principally cost and production outputs	People identify their own indicators of success, which may include production outputs
How	Focus on "scientific objectivity", distancing of evaluators from other participants, uniform complex procedures, delayed limited access to results	Self-evaluation, simple methods adapted to local culture, open immediate sharing of results through local involvement in evaluation process
When	Usually upon completion, sometimes also mid-term	Merging of monitoring and evaluation, hence frequent small evaluations
Why	Accountability, usually summative, to determine if funding continues	To empower local people to initiate, control and take corrective action

Participatory forms of development cooperation were propagated by all major donor agencies in the early 90s, stimulated by the political changes at the end of the Cold War and by the widely felt need to reform the aid relationship. In practice, the experiments and pilot projects very rarely engaged stakeholders at the project formulation level, or in evaluating the outcomes. Most often, they were aimed at involving stakeholders and beneficiaries in managing and maintaining resources that were made available by the donors, like water pumps, woodlots or rural roads. In many cases, the role of the stakeholders was limited to running and maintenance, but there were also instances of genuine partnership, with organisations or committees of stakeholders entering into more or less formal agreements with government agencies about sharing responsibilities, costs and benefits. Participatory projects were generally more time-consuming and costly than others, but in return they were more effective in reaching sustainable results. NGOs were generally believed to use more participatory approaches than government or multilateral agencies, but there was not much evidence to suggest that they were more effective, nor that the outcomes were sustainable (OECD-DAC, 1997).

When it comes to participatory forms of evaluation, practice is even thinner on the ground. A fundamental feature of participatory evaluation is that the *value system* and the standards against which performance will be measured, are chosen by the participants themselves, with external evaluators confined to a facilitating role. This approach is very rare, and it is doubtful whether it has ever been seriously applied to a major evaluation. A growing number of evaluations carry out prior *consultations* with stakeholder groups like implementing agency officials and beneficiaries. This is done through workshops in which the objectives and the measurement problems of the project are discussed with a stakeholder group before the evaluation is launched. But value systems and standards of measurement are generally not a subject for negotiation, since they have already been determined by the original design of the project. And the choice of participants is mostly limited to project 'insiders'. Whether or not such consultations can qualify as participatory is a question that depends on the degree of openness of the discussion and the ability of participants to formulate the questions to be answered by the evaluation.

Citizens' Report Cards as an Aide to Public Accountability in India

An interesting example of participatory M&E is given by the experience with citizens' report cards in some Indian cities, beginning in Bangalore in 1993. The project was implemented by the Public Affairs Centre (PAC) in Bangalore, a non-profit organisation, and was launched in response to the unsatisfactory quality of municipal services in the city. The PAC designed a series of interview schedules covering a range of Government agencies and invited random samples of citizens to rate the quality of services from those agencies. The findings were given large publicity and created a lot of public discussion. One immediate result of this discussion was the Swabhimana Initiative, a forum for discussion between the City of Bangalore and its citizens, initiated by a City official. The forum later developed into a partnership between the City and the citizens for solving the problems brought up by the report cards.

The use of citizens' report cards has later spread to other cities in India, like Chennai.

Source: Transparency International, *The Corruption Fighters' Tool Kit,* 2004.

If the evaluation concerns a project that is itself using a participatory approach, its focus will to a large extent be directed to non-quantifiable phenomena such as *attitudes, relationships and responsiveness* to other actors. The description of such behaviour is best left to the participants themselves, through some sort of self-evaluation or focus groups. This is also why participatory evaluation tends to become more of process monitoring than of passing judgement on a final product. In process evaluation, the participants themselves try to find solutions to the problems they discover, through a continuous series of *negotiations.* Again, if participants have sufficient influence on the outcome of those negotiations, they represent a base for building partnerships. On the other hand, this approach of self-evaluation and negotiation is less suitable for producing a comprehensive judgmental report at the end of the exercise. Descriptive reports of such processes are often difficult to present in a way that satisfies the agencies that commissioned the study.

From the point of view of determining the quality of partnerships, it is important *who chooses the participants, and how.* If all stakeholders are to be involved, they will be identified through the stakeholder analysis that precedes project formulation: direct and indirect beneficiaries, and groups that may be threatened by the proposed intervention. The forms in which they take part may have to be adapted to their different functions – those who are expected to sell equipment or services to the project do not always speak the same language as the villagers who are supposed to benefit from the outcome. But participants can be chosen through other criteria than their expected stake in the project. When the recipient government agency does the choosing (which is common in government-to-government partnerships), it will be more interested in involving local institutions and persons with an official status than in inviting personalities and organisations that are potentially controversial. From the point of view of the donor agency, however, it is not only more interesting but it will also serve the interest of effectiveness and sustainability to involve potential opponents or critics, and it would be worthwhile to test the strength of the partnership by asking for inclusiveness in the selection of participants.

4.4 Evaluation Capacity Development (ECD)

As mentioned at the outset, one reason for the recipient's limited interest in evaluations is a limited capacity to design, implement and interpret evaluation studies. From the very beginning of development cooperation, donors have been aware of this problem and have tried to help building evaluation capacity on the recipient side. At the beginning, those efforts were mostly limited to funding and staffing evaluation units in project organisations, or ministry units dedicated to monitoring a project funded by the same donor. In some cases, however, analytical capacity was developed at national level, often in research institutions attached to universities. Sida's support for the Economic and Social Research Foundation (ESRF) in Tanzania may be quoted as a successful example of such capacity development.

During the 90s, it became increasingly clear that the evaluation was a particularly unbalanced element in the development partnership. Evaluation Capacity Development (ECD) was launched as special line of activity by the World Bank and UNDP, the results of which were to be followed with special attention. The Bank's working group on M&E concluded in 1999 that the weaknesses were due to shortcomings both in the Bank and borrower agencies, including (World Bank, 2002):

- poor incentives to conduct good M&E
- diffused accountability because of unclear roles and responsibilities, both within the Bank and between the Bank and the borrowers
- weak capacity for M&E both within the Bank and in client countries

The Bank has considered a number of means for increasing the interest of its partners in ECD, including "persuasion and compulsion", and Bank management has brought pressure to bear on the Regions to step up the effort. In 2002 the Bank reported ECD activities in 21 countries.

Over the years, many developing countries have built national capacity for monitoring, and analytical studies, but their priorities for using this capacity do not always lie in evaluating donor-funded projects[1]. There are many reasons for this, some of which were mentioned in Section 2. Other reasons could be found in the budget systems and in attempts to make the donors provide 'additional' resources for evaluating their favourite projects. A study made for UNDP found that there was often an "uneasy" relationship between national evaluation systems and those used by the donors (Khan, 1993). Evaluation is an industry where procurement tying is still the dominant practice among bilateral donors, and its consequences are not conducive to developing national capacity. Fortunately, many donors have recently opened up for using local consultants, mostly in cooperation with expertise from the donor country. In some cases, institutions developed with support from ECD projects have been given preference for international evaluation contracts (OECD-DAC, 1996).

The increased attention from donors to national capacity for evaluation is welcome from the point of view of partnerships, and particularly when it is not tied to monitoring specific donor-funded projects. In poor developing countries, the capacity for social and economic research is presently being commandeered by donor agencies, to serve an agenda that is determined by those agencies rather than by national needs. It is important, as in the case of the Tanzanian ESRF, that national institutions be allowed to develop an identity and a capacity of their own, to become an independent voice in the development debate. Partnering with independent research institutions could create research networks and give a boost to evaluation driven by national interests and development priorities.

[1] Of a dozen HIPC countries that received World Bank support for ECD in 2002, only half included analysis to support the PRS process.

5. Conclusions and recommendations

This review of the evaluation partnership in a rapidly changing environment gives the impression that evaluation practice has been slow in responding to external changes, even when those changes offer opportunities for improving quality and outreach. It is by no means the first time this observation has been made. A Nordic programme manager in Tanzania recently expressed the view that conventional evaluation methods were becoming irrelevant in an environment where external assistance is increasingly transferred in the form of budget or sector programme support (Dahlgren, 2004). Administrative reviews of PRSPs and SWAPs seem to reduce the need for analytical evaluation, being more manageable than multi-donor evaluations and more immediately related to operational decision-making. Is the traditional approach of evaluating outcomes in relation to stated goals becoming a redundant practice?

While we have found numerous opportunities for improving evaluation practice that has been poorly utilised, we have also found simple and practical ways of making evaluation practice more relevant. It is quite obvious that more evaluations of the 'run-of-the-mill' type are undertaken than are really necessary for the sake of accountability and feedback. There are many indications that the burden of serving all the donors' requirements for reporting and enquiries is straining recipient administrations and thereby reducing their interest in evaluations. *Hence, what is needed is not more evaluations, but fewer, better and more adaptable ones*.

The concept of 'partnership' offers an approach to addressing this problem. Where donors and recipients have built a partnership that is aiming at something all of them want, and a partnership that is genuinely power-sharing, *the recipient agency should take the lead in setting the priorities for information gathering and feedback*. There is no other sustainable solution to the problem of escalating donor requirements for reports and data than letting the recipient be the judge. If one player in this game is allowed to impose requirements that the other players do not accept, this could wreck the whole game in the end by inducing non-cooperation, in particular from the recipient agency. Another conclusion that follows from this line of reasoning is that in some cases it may be better to withdraw from a degenerating relationship than to go on pretending it is a genuine partnership.

If the partnership is sound, it can be strengthened and deepened by engaging other groups of stakeholders than donor and recipient agencies. All new interventions should be preceded by a thorough stakeholder analysis, and the agency partners should discuss how to engage the stakeholders in formulating the intervention as well as in implementing and evaluating it. It is necessary to adapt the evaluation approach to the kind of partnerships that have been built around the project, but one essential requirement is that *stakeholders should take part both in defining the focus of the evaluation and in drawing the conclusions from it*. For Sida as a donor it is important that a broad range of stakeholders be invited to take part. There may be stakeholders within the recipient agency who want to limit this range, but a strong partnership can manage such differences.

In addition to these general observations, we should like to add a few recommendations regarding evaluation of some specific aid forms that have been mentioned in the paper.

- *The PRSP process:* Through its comprehensive coverage of social and economic development policy and the broadening participation of social partners, the PRSP process provides an unparalleled opportunity for evaluation to hold centre stage. So far, evaluations have mostly been limited to what is needed by donor agencies (particularly for the HIPC facility), but there are other aspects that would be of particular interests to parliaments, the NGO community or to academic research institutions in the country. Social consequences of a given growth pattern is one such aspect, the consequences of foreign aid flows on the national budget process may be another. Donors should assist national centres of policy analysis in building capacity for conducting such studies on a recurrent basis.
- *Sector Programme Support:* The problems of evaluating multi-donor SWAPs are partly technical and partly political. Technically, it is not more difficult to evaluate outcomes against targets at sector level than in a project, but the multiplicity of donor and recipient interests creates problems that may erode the partnership and even make it counterproductive. One such problem is attribution, which entails linking specific donor inputs to specific outcomes. A related problem is that of accountability, the need of which many agencies exaggerate far beyond the requirements of their principals (parliaments, government departments etc). Setting too high standards for evaluating SWAPs could lead to the eventual cessation of this practice, since annual sector reviews based on administrative reporting are so much easier to undertake (and are indeed necessary for monitoring progress). Sida is advised to concentrate its sector evaluation efforts on cases where strong and functional partnerships exist and to join like-minded donors in helping the recipient government take a lead in setting evaluation priorities.
- *Reviewing Incentive Systems:* Many studies of evaluation practice complain that incentive systems are stacked against evaluation in general, and against analytical approaches in particular. There are many factors behind this, both on donor and recipient side. The way around this problem as recommended here is to engage intended beneficiaries and other

stakeholders in the evaluation. But we would also recommend that the government partners review the system as a whole to create more general incentives for reporting that leads to policy change and to learning. At the moment, if there are any positive incentives at all, they reward the production of a report, not the presentation of new findings and ideas.

- *Engage Stakeholders in the Evaluation:* The original stakeholder analysis of the project would tell the partners commissioning the evaluation where to start. How far to go in inviting stakeholder groups to take part in designing and implementing the study and in learning from its results will be decided by the commissioning partners. People generally learn more from participating in an enquiry than from reading its report, and participants in the evaluation should be chosen both with regard to what they can contribute and what they can learn. In participatory studies, the process is often more important than the end product.
- *Evaluate the Donors' Contribution:* The success of the aid partnership depends on the contributions of both donor and recipient, but more often than not the evaluation focuses on the host government only. An yet, in PRS undertakings and in large sector programmes, the donors' contributions in terms of resources as well as their participation in the policy dialogue are often critical for the outcome of those processes. Recipient governments are understandably reluctant to criticise individual donors, but independent experts could be commissioned to do this. Sida should join like-minded donors in commissioning such studies to independent national institutions.
- *Develop National Capacity for Evaluation:* Some progress is being made with support from multilateral as well as bilateral donors in developing national capacity for evaluation. Some of these efforts are directed at evaluation units in agencies that implement large sector programmes, while an increasing share is going to independent research centres and university departments. It is important to help building such independent research capacity, and to offer these institutions a chance to take part in evaluations commissioned by the donors, in the country as well as outside.

References

Arnstein, S. R., *Eight Rungs on the Ladder of Citizen Participation,* in Cahn & Passett (eds), Citizen Participation: Affecting Community Change, Praeger, NY, 1971

Axelrod, R., *Theoretical Foundations of Partnerships for Economic Development,* prepared for World Bank Conference on Evaluation and Development – the Partnership Dimension, Washington DC, July 2001

Carlsson, C. and Engel, P., *Enhancing Learning through Evaluation: Approaches, Dilemmas and Some Possible Ways Forward,* (Background Papers), ECDPM, Maastricht, 2002

Chambers, R., *Rural Appraisal: Rapid, Relaxed and Participatory,* ISD Discussion Paper No 311, IDS Sussex 1992

Dahlgren, S., Report on Mission to Maputo, Lusaka and Dar es Salaam, Sida-UTV January 2004 (in Swedish)

Edgren, G., *Donorship, Ownership and Partnership, Issues arising from four Sida studies of donor-recipient relationships,* Sida Studies in Evaluation 2003:03, Stockholm 2003

ESRF, *Enhancing Aid Relationships in Tanzania: Report of the Independent Monitoring Group,* Report submitted to the Consultative Group for Tanzania, Dar es Salaam December 2002

Feinstein, O.N. and Ingram, G.K., *Lessons Learned from World Bank Experiences in Joint Evaluation,* Paper presented to the Workshop on Partners in Development Evaluation – Learning and Accountability, organised by France and OECD-DAC, Paris 2003

Forss, K. and Rebien, C., *Process Use of Evaluations. Types that Precede Lessons Learned and Feedback,* Evaluation Vol. 8(1), Sage Publications, London 2002

Fowler, A., *Questioning partnership: the reality of aid and NGO relations,* IDS Bulletin, 31(3), IDS Sussex 2000.

Hauge, A., *The Role of Monitoring and Evaluation Capacity Development in Improving Government Performance: Lessons Learned from the Uganda Poverty Reduction Support Process,* WB-OED, August 2003.

Helleiner, G., *Towards Balance in Aid Relationships. Donor Performance Monitoring in Low-Income Developing Countries,* Cooperation South, No. 2, 2000

Holmgren, T. and Soludo, C.C., *Implementation of Country-Led Partnership in the Context of the Comprehensive Development Framework,* World Bank August 2002.

Kanbur, R. and Sandler, T., *The Future of Development Assistance: Common Pools and International Public Goods,* ODC Policy Essay 25, Overseas development Council, washington 1999.

Khan, M.A., *Generic Issues in Monitoring and Evaluation: What Works and What Does Not?* CEO UNDP New York 1993

Narayan-Parker, D., *Participatory Evaluation: Tools for Managing Change in Water/Sanitation,* PROWWESS/UNDP/IBRD/UNICEF/WHO, New York 1991

Odén, B. and Tinnes, T., *New Aid Modalities and Donor Harmonisation,* Norad Report 8/2003, Oslo 2003.

OECD-DAC, *Evaluation of Programs Promoting Participatory Development and Good Governance,* DAC Expert Group on Evaluation, Paris 1997.

Ostrom, E., Gibson, C., Shivakumar, S. and Andersson, K., *Aid, Incentives and Sustainability, An Institutional Analysis of Development Cooperation,* Sida Studies in Evaluation, 2002

Picciotto, R., *The Logic of Partnership, A Development Perspective*, World Bank, Operations Evaluation Department, Photocopy, 1998.

Picciotto, R, *The Missing dimension of Development: Impact of Rich Countries' Policies on the Poor,* Paper presented to the Workshop on Partners in Development Evaluation – Learning and Accountability, organised by France and OECD-DAC, Paris 2003

Rebien, C. C., *How to Evaluate Development Aid, a Theory and Methodology in the Making – with a Special focus on Participatory Evaluation,* Handelshojskolen i Kobenhavn, Copenhagen 1996

Schaumburg-Muller, H, *Evaluation Capacity Building, Donor Support and Experiences,* Report for the DAC Expert Group on Evaluation, Paris 1996

Sida, *Evaluation Manual,* UTV, Stockholm 2004

Transparency International, *The Corruption Fighters' Tool Kit, Civil society Experiences and Emerging Strategies, www.transparency.org 2004*

UNDP, Human Development Report, New York xxxx

UNDP, *Evaluation of UNDP's Role in the PRSP Process,* Evaluation Office, New York, 2003

Weeks, J., et al. *Supporting Ownership. Swedish Development Cooperation with Kenya, Tanzania and Uganda.* Sida Evaluation 02/33, 2002

World Bank, *2002 Annual report on Evaluation Capacity Development (ECD),* OED, Washington, 2002

Sida Studies in Evaluation

96/1 *Evaluation and Participation – some lessons.*
Anders Rudqvist, Prudence Woodford-Berger
Department for Evaluation and Internal Audit

96/2 *Granskning av resultatanalyserna i Sidas landstrategiarbete.*
Göran Schill
Department for Evaluation and Internal Audit

96/3 *Developmental Relief? An Issues Paper and an Annotated Bibliography on Linking Relief and Development.*
Claes Lindahl
Department for Evaluation and Internal Audit

96/4 *The Environment and Sida's Evaluations.*
Tom Alberts, Jessica Andersson
Department for Evaluation and Internal Audit

97/1 *Using the Evaluation Tool. A survey of conventional wisdom and common practice at Sida.*
Jerker Carlsson, Kim Forss, Karin Metell, Lisa Segnestam, Tove Strömberg
Department for Evaluation and Internal Audit

97/2 *Poverty Reduction and Gender Equality. An Assessment of Sida's Country Reports and Evaluations in 1995–96.*
Eva Tobisson, Stefan de Vylder
Secretariat for Policy and Corporate Development.

98/1 *The Management of Disaster Relief Evaluations. Lessons from a Sida evaluation of the complex emergency in Cambodia.*
Claes Lindahl
Department for Evaluation and Internal Audit

98/2 *Uppföljande studie av Sidas resultatanalyser.*
Göran Schill
Department for Evaluation and Internal Audit

98/3 *Evaluating Gender Equality – Policy and Practice. An assessment of Sida's evaluations in 1997–1998.*
Lennart Peck
Department for Evaluation and Internal Audit

99/1 *Are Evaluations Useful? Cases from Swedish Development Cooperation.*
Jerker Carlsson, Maria Eriksson-Baaz, Ann Marie Fallenius, Eva Lövgren
Department for Evaluation and Internal Audit

99/2 *Managing and Conducting Evaluations. Design study for a Sida evaluation manual.*
Lennart Peck, Stefan Engström
Department for Evaluation and Internal Audit

99/3 *Understanding Regional Research Networks in Africa.*
Fredrik Söderbaum
Department for Evaluation and Internal Audit

99/4 *Managing the NGO Partnership. An assessment of stakeholder responses to an evaluation of development assistance through Swedish NGOs.*
Claes Lindahl, Elin Björkman, Petra Stark, Sundeep Waslekar, Kjell Öström
Department for Evaluation and Internal Audit

00/1 *Gender Equality and Women's Empowerment. A DAC review of agency experiences 1993–1998.*
Prudence Woodford-Berger
Department for Evaluation and Internal Audit

00/2 *Sida Documents in a Poverty Perspective. A review of how poverty is addressed in Sida's country strategy papers, assessment memoranda and evaluations.*
Lennart Peck, Charlotta Widmark
Department for Policy and Socio-Economic Analysis

00/3 *The Evaluability of Democracy and Human Rights Projects. A logframe-related assessment.*
Derek Poate, Roger Riddell
Department for Evaluation and Internal Audit

00/4 *Poverty Reduction, Sustainability and Learning. An evaluability assessment of seven area development projects.*
Anders Rudqvist, Ian Christoplos, Anna Liljelund
Department for Evaluation and Internal Audit

00/5 *Ownership in Focus? Discussion paper for a Planned Evaluation.*
Stefan Molund
Department for Evaluation and Internal Audit

01/01 *The Management of Results Information at Sida. Proposals for agency routines and priorities in the information age.*
Göran Schill
Department for Evaluation and Internal Audit

01/02 *HIV/AIDS-Related Support through Sida – A Base Study. Preparation for an evaluation of the implementation of the strategy "Investing for Future Generations – Sweden's response to HIV/AIDS".*
Lennart Peck, Karin Dahlström, Mikael Hammarskjöld, Lise Munck
Department for Evaluation and Internal Audit

02/01 *Aid, Incentives, and Sustainability. An Institutional Analysis of Development Cooperation. Main Report.*
Elinor Ostrom, Clark Gibson, Sujai Shivakumar, Krister Andersson
Department for Evaluation and Internal Audit

02/01:1 *Aid, Incentives, and Sustainability. An Institutional Analysis of Development Cooperation. Summary Report.*
Elinor Ostrom, Clark Gibson, Sujai Shivakumar, Krister Andersson
Department for Evaluation and Internal Audit

03/01 *Reflection on Experiences of Evaluating Gender Equality.*
Ted Freeman, Britha Mikkelsen
Department for Evaluation and Internal Audit

03/02 *Environmental Considerations in Sida's Evaluations Revised: A follow-up and analysis six years.*
Tom Alberts, Jessica Andersson, with assistance from:
Inger Ärnsfast, Susana Dougnac
Department for Evaluation and Internal Audit

03/03 *Donorship, Ownership and Partnership: Issues arising from four Sida studies of donor-recipient relations.*
Gus Edgren
Department for Evaluation and Internal Audit

03/04 *Institutional Perspectives on the Road and Forestry Sectors in Laos: Institutional Development and Sida Support in the 1990s.*
Pernilla Sjöquist Rafiqui
Department for Evaluation and Internal Audit

03/05 *Support for Private Sector Development: Summary and Synthesis of Three Sida Evaluations*
Anders Danielson
Department for Evaluation and Internal Audit

04/01 *Stronger Evaluation Partnerships. The Way to Keep Practice Relevant*
Gus Edgren
Department for Evaluation and Internal Audit

Stronger Evaluation Partnerships

The Way to Keep Practice Relevant

Sida

SWEDISH INTERNATIONAL DEVELOPMENT
COOPERATION AGENCY

Department for Evaluation and Internal Audit

Address: SE-105 25 Stockholm, Sweden.
Visiting address: Sveavägen 20, Stockholm.
Tel: +46 8 698 50 00. Fax: +46 8 698 56 15
E-mail: sida@sida.se
www.sida.se