

# Effective Practices in Conducting a Multi-Donor Evaluation



DEVELOPMENT ASSISTANCE COMMITTEE

O E C D

## FOREWORD

This publication provides guidance on how to plan and conduct a joint evaluation of development programmes when more than one donor agency is involved. The emphasis is on practical guidance, and it will serve as a useful tool for those who seek to promote joint evaluation and collaboration among development co-operation agencies. Such evaluations are becoming increasingly important at a time when the implementation of the partnership strategy is taking ever deeper root in the management cultures of agencies. With the enhanced role of partnership in the implementation of development programmes we can expect a growing demand for joint evaluation and for lessons learned from various modalities of jointly implemented development co-operation. In this light, the guidance provided in this report should be useful over the coming years.

Joint multi-donor evaluations provide opportunities but also create challenges. The collaborative nature of such evaluations requires special attention and handling. This guide sets out the key steps to be taken when planning and implementing multi-donor evaluations. It draws upon experiences encountered in various types of multi-donor evaluations, some of which were conducted by the DAC Working Party on Aid Evaluation, while others were joint evaluations of multilateral agencies such as UNICEF, or more recently the European Commission's aid programmes and UNCDF. Insights have also been gained through joint evaluations by the Nordic countries.

This publication is highly relevant for new directions of work in the Working Party, notably on promoting greater partnership in evaluations through more evaluations with partners and identification of useful practices for closer collaboration with partners in all stages of the evaluation process.

This work was initiated by the Working Party on Aid Evaluation and led by USAID. The report was written by Ms. Annette Binnendijk (consultant to USAID) who has done an excellent job of identifying and analysing useful experience. The DAC Working Party on Aid Evaluation has reviewed the report at several stages of its elaboration and approved its publication at its meeting in February 2000.

Special thanks should be given to USAID for leading this work and to Sida, Sweden for financial support, which enabled this to be a free-distribution publication. The publication was prepared by the Development Co-operation Directorate of the OECD under the direction of Mr. Hans Lundgren.

Niels Dabelstein  
Chairman, Working Party on Aid Evaluation

## TABLE OF CONTENTS

EFFECTIVE PRACTICES IN CONDUCTING A JOINT MULTI-DONOR EVALUATION ...	4
What Is a Joint Multi-Donor Evaluation? .....	4
Why Conduct a Joint Evaluation with Other Donors?.....	4
What Are Some Pros and Cons of Joint Evaluation? .....	5
Key Steps in Planning and Conducting a Joint Multi-Donor Evaluation .....	6
1. Deciding on a joint multi-donor evaluation .....	7
2. Deciding on evaluation partners.....	8
3. Planning the evaluation’s management structure .....	10
4. Other aspects of planning (preparing the scope of work) .....	12
5. Selecting the evaluation team .....	14
6. Preparing the evaluation team.....	15
7. Collecting, analyzing, and reporting results.....	16
8. Reviewing, communicating, and following-up on evaluation results .....	20
Annex 1: The Joint Multi-Donor Evaluations Reviewed .....	22
References.....	26

## **EFFECTIVE PRACTICES IN CONDUCTING A JOINT MULTI-DONOR EVALUATION**

This is a guide for donor agency managers involved in planning and implementing joint evaluations of development assistance programs with other donor partners.

It addresses what joint multi-donor evaluations are and why they are useful, discusses some of their strengths and limitations, and provides tips for conducting them effectively. The tips for effective practices are organized according to key steps in the evaluation process, emphasizing those aspects requiring special consideration given the collaborative nature of joint evaluations. It is based on a review of donor agencies' experiences and lessons with joint evaluations.<sup>1</sup> Annex 1 provides an overview of the joint multi-donor evaluations that were reviewed as background for this paper, organized into several key categories. References are provided in Annex 2 for those wishing to examine the specific evaluations in more detail.

### **What Is a Joint Multi-Donor Evaluation?**

These are evaluations of development assistance programs or activities conducted collaboratively by more than one donor agency. Thus, the focus here is specifically on *multi-donor collaboration* in evaluation, rather than on other possible types of evaluation collaboration, such as donor-recipient country or donor-NGO collaboration. Nor does it cover participatory evaluation with its specific techniques for bringing beneficiaries and other stakeholders into the evaluation process. To do justice to these other types of joint, collaborative, and participatory evaluations will require separate efforts. Nevertheless, integrating recipient country partners into joint multi-donor evaluation processes is discussed in this paper in several places. For simplicity, hereafter, the term *joint evaluation* is used interchangeably with *joint multi-donor evaluation*.

Joint evaluations among donors can vary considerably in terms of their characteristics. For example, they may vary by the number of donors participating in the joint evaluation (i.e., small versus large groups). Their approach to managing and implementing the evaluation process can differ considerably -- with varying degrees of donor participation at different stages. Their focus may be on individual projects, multilateral agency programs, sector-wide programs, co-financing arrangements, crosscutting or thematic concerns, etc. Their scope may be limited to a single country, may cover a region, or be worldwide. Their purpose may emphasize learning lessons for program improvement (managing-for-results) or may stress judging performance (accountability-for-results). Methodologies used can also differ widely, ranging from desk reviews of existing information sources, to primary data collection and fieldwork in developing countries. The nature and degree of recipient country participation in the evaluation process also varies. More will be said on these dimensions later.

### **Why Conduct a Joint Evaluation with Other Donors?**

The changing context of development assistance approaches during the 1990s has increased the importance and utility of joint multi-donor evaluations. Sector-wide assistance programs and national development programs are emerging that vary considerably from traditional assistance activities. These new approaches emphasize collaborative, multi-donor assistance efforts with shared common objectives, and sometimes employ co-financing schemes, rather than separate donor projects. They also stress greater participation -- and leadership roles -- for the recipient countries. Effective evaluations of such new assistance modes call for -- even require -- greater donor collaboration.

The partnership strategy, *Shaping the 21st Century: The Role of Development Co-operation* (OECD/DAC: May 1996), which has been broadly endorsed by the development community, challenges donors to make joint evaluations not an occasional effort but a routine approach to conducting evaluations. It calls for collaboration in evaluation not only among donors but with recipient countries' participation as well. The Development Assistance Committee (DAC) Senior Level Meeting in January 1998 identified joint monitoring and evaluation as one of the important points for implementing the partnership vision (see Box 1).

Several donor agencies also have specific policies that promote the use of joint evaluations.<sup>2</sup>

### **Box 1: DAC Places Priority on Joint Evaluation**

One of the "11 points" for Strengthening Partnerships for Development identified at the January 1998 Forum of Development Partners and the DAC Senior Level Meeting states:

"The practices of joint monitoring and evaluation of development programmes by donor and recipient partners should be further developed and applied, with a view to learning together the lessons of achievements and failures."

### **What Are Some Pros and Cons of Joint Evaluation?**

Experience reveals a number of strengths and weaknesses of joint multi-donor evaluations, compared to those sponsored by a single donor agency.

Some often-cited advantages of joint evaluations include:<sup>3</sup>

- Reducing the burden of multiple, separate donor evaluation efforts on recipient country institutions, including overlapping team visits and duplicative data collection efforts.
- Providing an opportunity for donor agencies to learn and share their evaluation processes and techniques among themselves.
- Enabling a greater diversity of perspectives and talents on the joint evaluation team.
- Permitting cost-sharing of the evaluation's expenses.
- Enabling a broader scope and number of evaluation questions to be addressed, given extra (jointly-shared) resources.
- Promoting higher quality, more credible evaluations with broader ownership of findings and greater influence on decisions by the participating agencies.
- Fostering a greater consensus among donors and host country agencies on program and country priorities and needs, thus stimulating improved donor co-ordination of programs.

Moreover, joint evaluations will have a distinct advantage as donors continue to shift their assistance strategies away from isolated projects and towards national, sector-wide programs that stress collaboration and harmonization of multi-donor assistance efforts. In some cases (e.g., the Special Program of Assistance for Africa), donors have been moving towards common implementation arrangements and joint co-financing schemes. As donors increasingly focus on achieving more significant development goals that can only be achieved through partnerships, more joint multi-donor evaluations will be needed. These joint evaluations will also require increased participation and leadership roles by the recipient countries themselves.

Despite the large and growing number of advantages of joint evaluations, a recent *Review of the DAC Principles on Aid Evaluation* found joint evaluation work among donors to be "weak". While there are numerous examples of successful joint evaluations, they are not yet common practice. The review concluded that generally speaking there is "little enthusiasm" for joint evaluations.<sup>4</sup> Why this gap between policy advocating joint evaluations -- and actual practice? Why aren't more joint evaluations taking place?

Reasons may be due to a variety of obstacles and difficulties, such as those encountered in:

- Identifying other donors interested in the same evaluation topic.
- Developing comprehensive, yet manageable, terms of reference that accommodate each participating donors' particular issues and interests.
- Selecting teams that are mutually acceptable to all participants.
- Coordinating schedules and travel logistics.
- Agreeing on cost-sharing and contracting arrangements.
- Developing joint management structures and communication processes that work effectively.
- Reaching agreement on reporting formats, findings and recommendations.

Differences among donor agencies' operational processes and organizational cultures can also create obstacles to smooth collaboration efforts. For example, joint evaluations can be complicated by differences among donor agencies'

- Evaluation policies and procedures, methods, and reporting requirements.
- Administrative policies and procedures -- especially financial and contracting mechanisms.
- Evaluation priorities, agendas, schedules, and issues -- and with limited time and resources getting one's "own" work completed usually has priority over "joint" efforts.

Clearly, joint evaluations require greater lead-time and effort to organize, schedule, and otherwise coordinate among the participating donor agencies. The costs of joint evaluations, in terms of overall resources, personnel, and management time, are often considerable.

This guide to more effective practices in planning and conducting joint evaluations aims to provide tips for dealing with and lessening difficulties encountered, while enhancing the benefits.

### **Key Steps in Planning and Conducting a Joint Multi-Donor Evaluation**

Below are some practical tips for donor agency managers involved in planning and managing a joint multi-donor evaluation. There are also suggestions for teams conducting the joint evaluation. These guidelines address effective practices -- based on the lessons of experience -- and are organized according to key stages of the joint evaluation (i.e., in a step-by-step fashion). While arranged as if they were sequential steps, in reality some processes may be more iterative in nature.

## ***1. Deciding on a joint multi-donor evaluation***

The first question an evaluation manager must address is whether a joint evaluation approach is more appropriate and useful in a given situation, versus the more traditional single-agency evaluation approach. What are some of the factors to be considered in this decision?

***Focus.*** An evaluation's focus should be a key influence on the decision of whether to conduct an evaluation jointly or singly. If the focus is on a project or program funded by only one donor, it is probably best evaluated by that donor agency alone.<sup>5</sup> If the focus is on a jointly funded project or program, a joint evaluation by the co-sponsoring donors is more appropriate. If the focus of the evaluation is on a broad sector development goal in a recipient country, which is influenced by the collaborative program efforts of a number of donors, then a joint evaluation effort by the involved donors is advantageous and should be considered. In this case, it is also particularly important to encourage participation by the recipient country in the joint evaluation effort.

Similarly, evaluation of a multilateral or regional development agency can best be done in a joint context by interested donor countries (e.g. those contributing substantial funding to the agency being evaluated).

***Purpose.*** Whether the primary aim of an evaluation is to learn lessons for improving operations or to judge performance for accountability purposes may -- to some extent -- influence the decision of whether a joint evaluation is to be considered. However, there are certainly examples where joint evaluations have successfully fulfilled both purposes.

If a donor agency is interested in assessing accountability for performance of its own program and wishes to single out its own individual contributions to results achieved (i.e., attributing results to specific agency activities), then a joint evaluation approach is probably inappropriate.

If instead, the purpose is to assess the combined contributions of a number of donors' programs on a broader (e.g., sector-wide) development goal which they share in common, then joint evaluation becomes much more suitable. Still, making judgements concerning performance and accountability issues can result in heated debates and controversies -- participating donors may even be unable to reach agreement on findings. Prior agreements on processes for resolving potential conflicts may be useful.

Assessments of performance or accountability may also call for special precautions to preserve independence and objectivity of the evaluation, which may be more complicated, but not impossible -- in jointly conducted evaluations. For example, sponsoring donor agencies and participating recipient country partners may be asked to agree up front to respect the independence of the evaluation team and their findings. If independence is effectively protected, joint evaluations can actually have greater credibility than single donor evaluations.

When the key evaluation purpose is drawing lessons and implications for operational decisions, collaborative approaches such as joint evaluations tend to be especially useful. Lessons based on the collective wisdom of numerous donors (and recipient partners) are likely to be more valid and reliable than those based on a single donor's experience. Moreover, broader participation tends to increase confidence in and ownership of evaluation findings, making follow-up action on recommendations more likely by all partners. Generally speaking, drawing lessons for program improvements is less likely to result in situations of conflict and controversy (than judgements on performance), and may thus be "easier" subjects for joint evaluations.

***Comprehensiveness of the scope of work.*** The complexity or comprehensiveness of the evaluation questions found in the scope of work (or terms of reference), may be very relevant to the decision of conducting an evaluation jointly.

If the evaluation questions to be addressed by an evaluation are very broad, a joint evaluation can be an advantage, in the sense that the workload and costs can be shared among the participating donors. For example, if a scope covers all democratization programs, one participating donor might cover legal system support, while another handles decentralization, etc. Similarly, if the scope is worldwide, one donor might focus on Africa, the next on Asia, etc.

A joint evaluation's scope of work often tends to be quite comprehensive for another reason as well -- each of the participating donors and partners may have their own issues and questions they want addressed.

Conversely, if a donor wishes an evaluation to remain very focused -- say, on a few questions raised by its own management -- and wants to remain exclusively in control of the scope, then joint evaluations are probably inappropriate.

***Practical (cost, timing) considerations.*** If a donor is in a big hurry to get an evaluation completed, joint evaluations may be disadvantageous. The collaborative nature of joint evaluations tends to increase their complexity -- which often increases the total level of effort and the length of time needed as well. Joint evaluations tend to increase the management burden, since frequent meetings or other forms of communication and collaboration among participants are required at every key stage. Differences in donor agencies' policies and procedures complicate any joint effort. On the other hand, some very complex joint evaluations have been completed quickly, so it is possible. For example, the joint evaluation of international emergency assistance to Rwanda, despite its comprehensive scope-of-work accommodating special interests of 38 participating member agencies, was completed in only a little over a year.<sup>6</sup>

However, another practical consideration -- cost -- may favor undertaking a joint evaluation, since it provides opportunities for cost sharing. The joint evaluation approach provides small donors with the option of participating in comprehensive types of evaluations than might otherwise be beyond their capacity to conduct.

Ultimately, the decision of whether to conduct an evaluation singly, or to seek other donors' participation, must be based on a careful weighing of all the likely advantages and disadvantages in a given case.

When deciding to bring in other donor partners, careful consideration should be given to how it may affect the whole evaluation process. A truly collaborative joint evaluation means sharing decisions concerning the evaluation's planning and management, its scope of work, the team composition, the methodology, the reporting format and findings, and many other aspects. If a donor agency wishes to maintain tight control of an evaluation, better not do it jointly.

## ***2. Deciding on evaluation partners***

***Other donors' participation.*** Which other donors might be most appropriate and interested partners for a joint evaluation? The answer may be found most naturally in the evaluation's focus. For example, if the focus is on a jointly funded project, the most appropriate partners are the co-financing donor agencies. If the focus is on achievement of a broader development goal in a recipient country,

appropriate evaluation partners might be those with major program activities contributing towards that sector goal. In order to facilitate joint evaluations of this type, Danida has prepared a matrix indicating for each recipient country and sector, the donors that have major development programs.<sup>7</sup> In cases where sector-wide investment programs are financed by multi-donor co-financing schemes, the participating donors would make natural partners for a joint evaluation of that sector program. Likewise, if the focus is on a multilateral or regional development agency, those donor countries who contribute financing to the institution may be most interested in participating in its evaluation.

Other factors in selecting other donors for partners in joint evaluation work may be relevant as well. Selecting those with similar development philosophies, organizational cultures, evaluation procedures and techniques, regional affiliations and proximity, etc. may make working together easier. For example, such similarities have made the Nordic donor countries frequent collaborators in joint evaluations. Another issue may be limiting the total number of donors participating to a "manageable" number.

Once appropriate donors are identified that have a likely stake in an evaluation topic, the next step is to contact them and see if they are interested in participating. In some cases, there may already be an appropriate donor consortium or group where the issue of a joint evaluation can be raised and expressions of interest can be easily solicited. For example, one such group which has officially sponsored numerous joint evaluations and has informally encouraged many more, is the DAC Working Party on Aid Evaluation. The group's long tradition of co-operation, shared vision on evaluation principles, and personal relationships built over the years has fostered numerous joint efforts.

***Recipient country participation.*** At this stage, also consider broadening participation in the joint evaluation to include recipient country partners as well. For example, in a joint multi-donor evaluation focused on the sector development efforts of a particular recipient country, seek the participation of appropriate representatives from that country's government ministries and agencies, NGOs, or civil society organizations. Even in cases where the evaluation's scope is regional or worldwide, there are ways to ensure at least selective participation from recipient countries. For example, an advisory group comprised of selected recipient country representatives might be formed to consult with the team or to review drafts. Or nationals from several different recipient countries might be hired to serve on the evaluation teams in various capacities. Box 2 provides some suggestions for roles recipient country partners might take in a joint evaluation.

The DAC partnership strategy strongly encourages such participation. Moreover, participation by recipient country nationals often bring to the evaluation effort special expertise regarding country context and program implementation issues. Another advantage is that participation in the evaluation process also often brings about ownership of the evaluation's findings and recommendations, thus making follow-up on actions by the host country institutions more likely.

On the other hand, there may be some disadvantages in certain cases. For example, there may be legitimate concerns about the evaluation's objectivity and independence, if the recipient country participants have strong biases or "stakes" in the evaluation's outcome. Moreover, recipient countries may not be able to help finance or sponsor the evaluation, and thus their participation may mean higher costs for the donor participants.

### ***3. Planning the evaluation's management structure***

Deciding on how the joint evaluation is to be effectively managed, and determining the roles of all donors participating is a key aspect of successful planning. The number of participants influences the complexity of the management structure. Joint evaluations can vary significantly in terms of the number of donors participating, from small groups of 2-6 donors to very large groups of 20 or more. Participation by a donor may be limited to a sponsorship role (i.e., providing financing and broad guidance), or may include management responsibilities (i.e., becoming actively involved in planning and implementing the evaluation).

In cases where only a few donors are involved, the management structure is usually quite simple. For example, the participants may decide to share equally in all management decisions, meeting frequently to ensure a smooth evaluation process. Perhaps one of them will take a special lead or coordinating role. Alternatively, the participants may decide to delegate management to just one among themselves, who then conducts the evaluation on their behalf, while the rest limit their role to sponsorship. Sponsors usually share the evaluation's costs and review key products -- such as the scope of work and the final reports.

In other cases, a relatively large group (either a formal or informal association) of donors decides to sponsor a joint evaluation. The larger group's role is usually limited to:

- Providing broad guidance and direction to the effort.
- Sharing the evaluation's costs.
- Contributing other appropriate inputs (such as sharing their agency evaluation reports, files, databases, policy papers, etc, and responding to interviews or questionnaires).
- Reviewing key evaluation documents at critical stages (such as the evaluation's scope of work and final reports).

#### **Box 2: Some Approaches for Integrating Recipient Country Partners in Joint Evaluations**

Consider the following possibilities:

- Invite recipient country representatives to participate on the evaluation's management group.
- Select recipient country nationals to be members of the evaluation team.
- Form an advisory panel comprised of recipient country nationals to advise the evaluation team and/or review findings.
- Involve recipient country nationals in the evaluation process via other roles -- as key informants, focus group members, researchers, interviewers, country/regional experts, etc.

Day-to-day management is typically delegated to either one member or to a few members of the larger group. When a few are actively involved, they typically form a management "working group" or "steering committee" which is designated responsible for most aspects of planning and managing the evaluation.

Often the small management group is self-appointed, comprised of representatives of those donors with special interests in the evaluation and thus willing to volunteer substantial efforts. One member is usually selected as chairperson of the group. Individual members share the workload -- often volunteering to undertake specific tasks for the group, such as drafting the evaluation scope of work, preparing a desk study, conducting fieldwork for a country case study, etc. In larger studies with several sub-components, often different members will take the lead on each component -- conducting

them concurrently. One member then takes responsibility for pulling the component studies together into a synthesis report.

There are pros and cons to conducting several sub-components concurrently by different donor agencies as part of a management group, versus conducting a unified study by just one member. The sub-component model may suffer from theme overlaps (duplications of effort) as well as missing gaps (issues that may fall between the cracks). Communications among teams working on the different components may be problematic. On the other hand, this approach has significant advantages of breaking large studies with diverse and complex issues into manageable components. Teams working concurrently rather than a single team working sequentially may save time. Importantly, it affords greater opportunity for sharing the workload and for participation and leadership roles for more members.

In joint evaluations sponsored by a large group of donors, the approach of delegating management responsibilities to a smaller working group or committee has generally worked well. Size of the working group is probably best limited to no more than 5 or 6 donor members. Below are some suggestions and considerations for establishing an effective management working group structure:

- *Consider dividing comprehensive scopes of work into components for which individual working group members will be responsible.* Where studies break down into sub-components, having different members volunteer to take the lead in each is a practical way of dividing the workload and getting broader participation. This approach has advantages of breaking large studies with diverse and complex issues into manageable components that can be conducted simultaneously, thus speeding the process.
- *Appoint one of the members as chairperson of the working group, with key leadership and coordinating roles.* Designating a working group member as chairperson is essential -- for coordinating the overall study process, chairing working group meetings, reporting progress to the larger sponsoring group, overseeing quality control, and keeping the study on-track. Another key role of the chairperson is to facilitate coordination and communication among teams working on different study components, thus avoiding overlaps, preventing gaps, and ensuring greater consensus when it comes time to synthesize overall study findings and recommendations.<sup>8</sup>
- *Consider assigning other key tasks to individual members.* For example:
  - Designate one group member responsible for drafting the evaluation's overall scope of work, which other members can then review and revise. Sometimes this elaboration takes the form of developing more detailed design or approach papers for each sub-component by the responsible lead members. An alternative that has also worked well is to have a few representatives from the working group work intensively together for a short period to develop the scope.
  - Make one member responsible for preparing the overall evaluation synthesis report that draws on and summarizes various sub-component studies.
  - Have one member provide overall administrative support for the evaluation effort, possibly including tasks such as contracting with consultants, keeping track of overall budget and schedule, handling fieldwork logistics, preparing progress reports, finalizing documents for publication, etc.
- *Select appropriate organizational units to manage the evaluation.* It may be helpful to make the same functional units (e.g., all evaluation units or all program units) responsible for managing the evaluation within each of the donor agencies participating in the

management group. If there is a need for a high level of independence and objectivity, the evaluation unit is probably the more appropriate choice.

- *Ensure good communication among working group members.* Meetings should take place regularly at key points in the evaluation process, including representatives of the agencies participating in the management group, as well as the core members of the evaluation team. Moreover, make use of frequent, informal communications via phone conversations, conference calls, and Internet email (for sending notes, holding virtual meetings, sharing documents, etc.).

#### **4. Other aspects of planning (preparing the scope of work)**

A key lesson from experience is this: giving adequate time and attention to the planning of joint evaluations can be critical to their success. One key aspect of planning -- deciding on the management structure -- has already been discussed. Other aspects of planning include:

1. Identifying the evaluation's focus, purpose and audience,
2. Clarifying the key evaluation questions to be answered,
3. Identifying the methods to be used to answer the questions, and
4. Planning various other aspects of the evaluation process (e.g., team selection, schedule, reporting requirements, administrative support, and budget).

To help ensure adequate planning, it is very useful to prepare a detailed scope of work. The evaluation's scope of work (also frequently called terms of reference) is the key planning document. Not only is the scope of work a guide that the evaluation team must follow, but a document that all participants -- including the managing and sponsoring donor agencies -- can agree upon and use to help ensure a smooth evaluation process and a quality product.

Some suggestions for what to include in a joint evaluation's scope of work are elaborated on below:

- ***Evaluation focus, purpose, and audience.*** Begin by clarifying the evaluation's focus. Is this an evaluation of a particular project, institution, sector assistance program, etc? Is it examining one donor's efforts, two, three, etc? Is it concentrated on one developing country, a whole region, or global? The purpose also needs to be stated. Is it to judge performance of an assistance effort, to draw lessons for making improvements, or a mix of both? What aspects of performance -- impact, cost-effectiveness, sustainability, etc. -- will be assessed? The intended audiences and uses of the evaluation findings can also usefully be articulated. In the case of joint evaluations, the audiences may be quite varied, including the sponsoring group of donors, the institutions and implementing agencies being evaluated, the recipient country officials, etc. Intended uses may be many -- for example, to make funding decisions, to guide improvements in operations of implementing agencies, to influence donor agencies' policies, procedures and strategies, etc. If the intended use involves a specific time constraint (e.g., input into a conference), give the date.
- ***Evaluation questions.*** A major challenge in preparing a scope of work for a joint evaluation lies in balancing the need for including the special interests and issues of the participating member agencies, with the need to keep a clear focus and manageability. Some joint evaluations have suffered from scopes that were too broad and complex, given their resource and timing constraints. On the other hand, one of the advantages of joint evaluations is that costs can be shared, enabling more comprehensive scopes to be accommodated.

- **Methods.** Good scopes of work should identify the basic methods to be used (e.g., document reviews, interviews, questionnaires, case studies, rapid appraisal methods, sample surveys, etc.) and key phases in which they will be implemented. Methods should be selected based on their capacities to credibly answer the evaluation questions posed.<sup>9</sup> Details, such as development of the specific instruments, are typically left to the evaluation team. Sources of information that will be made available to the evaluation team (e.g., documents, performance monitoring data, project files, etc.) should be identified. Remember that implementing data collection methods in joint evaluations often will require more time and effort than in single donor evaluations. For example, if the method requires interviewing key informants in all participating donor agencies, this will be far more time-consuming than if representatives of only one donor agency were interviewed. Fieldwork involving site visits to the programs of numerous donor agencies will require more effort than visits to a single agency's programs would take.
- **Management structure.** Outline the management structure agreed upon, and the specific roles and responsibilities of the sponsoring group members, and of those members handling the day-to-day management functions. Plan how the management group will be coordinated and channels of communication.
- **Team selection process.** Identify the planned team size, composition, and skills mix. State any agreed-on criteria and procedures to be used for selecting team members. Outline the roles and responsibilities of team members. Clarify team relationships and lines of communication with the management group.
- **Schedule.** Scopes should cover the evaluation's duration, phasing and timing considerations. Prepare a timetable for the key phases and establish due dates for submitting reports. This is essential for keeping a joint evaluation on track. Schedules should be realistic, with consideration given to the comprehensiveness of the evaluation topics, related to the size of the team and other resource constraints. Take care not to set unrealistic schedules that underestimate the level of effort required -- this can result in limited analysis and weaker findings.
- **Report requirements.** State when various draft interim and final reports are due and who is responsible. Provide adequate details on their format, contents, length and language requirements. Requirements for executive summaries are especially useful when there are numerous component studies that will later need to be summarized in a synthesis report. Discuss report review procedures, with particular attention to arrangements for negotiating possible conflicts over findings and/or for protecting the evaluation team's independence. Finally, it is useful to provide information and dates for planned briefings, conferences, or other means of communicating and disseminating evaluation findings.
- **Administrative support.** Scopes should provide information on administrative and logistical support services available to the team (e.g., translators, interviewers/data collectors, drivers, office space, equipment, vehicles, travel arrangements, etc.). Logistics can be especially complicated in joint evaluations, with team members often coming from different agencies and continents. Each donor agency has their own policies and procedures for travel arrangements, for contracting services, etc. Coordination of all these administrative aspects can be daunting. Some joint evaluations have simplified things by assigning all administrative aspects to a single donor agency or by contracting out these functions to a single firm. If the joint evaluation has been neatly divided into separate sub-components, the usual approach has been to let each participating agency handle the administrative support and contracting for their own component.

- **Budget.** The scope of work usually estimates the evaluation's budget requirements and discusses the cost-sharing arrangement agreed to by the sponsoring donors. It may also address contracting procedures (e.g., for hiring team consultants and for obtaining administrative support services) and responsibilities of participating agencies.

The task of drafting a coherent and manageable scope of work may effectively be accomplished by delegating the task to one donor agency that agrees or volunteers to take the lead. This often happens in practice. Sometimes it is drafted jointly. For example, in one case (the joint evaluation of emergency assistance to Rwanda), a small core group representing the key donors were highly successful in drafting a comprehensive scope of work in a very short time, by working intensively together for a week.<sup>10</sup>

Regardless of who takes the lead in drafting the initial scope, it must undergo review and approval by the rest of the management group and usually by the broader sponsoring donor group as well. This is needed to ensure reasonable interests and issues of participants are included and thus facilitate their "ownership" of the evaluation. Such reviews also often improve the quality of the scope, although there is danger that they may result in adding on too many special questions, topics or components -- and ending up with an unmanageable and unrealistic scope, given the resources and timeframe allotted. If this happens, the level of effort that can be devoted to any particular topic diminishes and quality will suffer. Considerable facilitation skill may be needed to gain a general group agreement (if not a complete consensus) on the scope, while avoiding the pitfall of over-burdening it with every possible suggestion for expansion or add-on.

### ***5. Selecting the evaluation team***

The appropriate size of the joint evaluation team, of course, depends on the total level of effort as well as on the timeframe allowed. Often with comprehensive scopes of work, there may be several evaluation teams involved (e.g., each working on different components or conducting fieldwork in different countries or regions). When this is the case, it is effective practice to have either a core team and/or a team coordinator that facilitates communication and coordination among the various teams and that helps ensure continuity of effort. For example, the core team may conduct a document review in phase 1, lead a number of expanded country teams in a fieldwork phase 2, and then prepare a synthesis report in phase 3 -- thus helping to ensure continuity throughout the evaluation process. One approach is to have a small core team of competent generalists, and then bring in specialists (with technical, sector, country, or other specific skills) when needed.

Joint evaluation teams are often comprised of consultants who are nationals from the participating donor countries, although experience indicates it is often useful to include some donor agency officials on the team as well as consultants. This practice facilitates team communications with the management group. Having such officials participating on the team -- and thus being thoroughly familiar with the substance -- is especially useful if there arises a need to "sell" controversial team findings to the management group or to the broader sponsoring group of donor agencies.

Teams may also have members who are representatives from the implementing agencies or recipient country organizations being evaluated. This may be particularly useful if the evaluation's purpose is to learn lessons for program improvements, although potential conflicts-of-interest may arise casting doubts on credibility if the purpose is more of an accountability function (i.e., of judging performance). This issue may be overcome by giving them a special "observer" or "informant" role that protects the overall team's objectivity, or by having them serve on a special advisory panel with which the evaluation team may consult. Selecting recipient country nationals who are knowledgeable,

but who are not official government or implementing agency representatives, is another way of increasing team participation while minimizing possible vested interests.

Team selection criteria should be based on relevant qualifications and skills that team members collectively should have to conduct the particular evaluation effectively, such as areas of technical or sector competence, language skills, in-country work experience, and evaluation methods/ data collection skills. In addition, joint evaluations may involve other criteria, such as selecting team members that are nationals from the participating donor countries or are from selected recipient countries. In cases where maintaining the independence and objectivity of the evaluation is a special concern, team members may be screened to ensure they have no evident conflicts-of-interest (i.e. no particular stake in the evaluation's outcome).

Selecting a team leader requires special attention. Especially with the large and diverse teams typical of joint evaluations, the team leader must have not only evaluation and subject-matter expertise, but also strong management, communication, and writing skills. The team leader will have to manage the team, handle internal team disputes, negotiate with the management team representatives, communicate with high-level host country officials, and often pull together various components of the study into a final report.

Selecting a competent and objective team is important for obtaining a high-quality and credible evaluation. Because of this, it has been suggested that the management group members should jointly review and approve all candidate team members. One approach that has sometimes been used is for the group members to develop a short-list of candidates -- via a process of discussing the strengths and weaknesses of various individuals, agreeing on those who would be most suitable, and then limiting the bidding/contracting process to those individuals.

## ***6. Preparing the evaluation team***

Holding a team planning workshop at the beginning of a joint evaluation will help the team get off to a good start. Evaluation team members as well as representatives from the donor management group should participate, if possible. The purpose of the workshop should be to build an effective team that shares common understandings of the evaluation's purpose and plans, and prepares them as much as possible for the work ahead. Typically, they review all aspects of the scope of work, clarify roles and responsibilities of the team members (and their relationships with the management group), and agree on how the team will function together.<sup>11</sup> A suggested workshop agenda is outlined in Box 3.

The teams of joint evaluations often benefit from considerable diversity, such as multi-nationalities and cultural backgrounds; different language proficiencies, disciplines, and skills; and varying approaches to evaluation. While such diversity can enrich the evaluation's findings, differences may simultaneously pose a challenge. Language barriers may be problematic for inter-team communications. So may differences in perspectives and opinions, if they result in conflicts rather than compromises. In this context, team planning meetings become especially useful, because they prepare teams to work together more collaboratively and effectively.

## 7. *Collecting, analyzing, and reporting results*

Joint multi-donor evaluations have drawn on a variety of evaluation methodologies for gathering empirical information. The evaluation studies sponsored by the DAC Working Party on Aid Evaluation, for example, have usually relied largely on desk reviews of evaluation reports and other program and policy documents of member donor agencies. Occasionally this document review approach has been supplemented by information from questionnaires and interviews with donor agency officials, and/or with fieldwork in developing countries.

Other joint multi-donor evaluations have usefully combined several methods -- often in a phased approach. For example, several joint evaluations of multilateral agencies typically began with a desk study reviewing documents and interviewing staff at the multilateral agency's headquarters.<sup>12</sup> Results from this phase then fed into a second phase consisting of fieldwork in selected developing countries, focusing on the performance of the multilateral agency in these country contexts. Country reports were drafted which would be used in the final phase of preparing a synthesis report with overall findings, conclusions and recommendations.

Some of the lessons from joint multi-donor evaluation experience with the methods and processes used in these data collection phases are discussed below.

### *Phase 1: Document reviews and supplemental interviews*

**Document reviews.** This approach involves a systematic review by the evaluation team of evaluation reports and/or other documents (e.g., project records, performance monitoring information, etc.) submitted by donor agencies in order to answer specific evaluation questions.

Evaluation synthesis studies have proven very useful for summarizing the experiences, findings and lessons of the donor community in a particular program, sector, cross-cutting or thematic area of

### **BOX 3: A Suggested Agenda for a Team Planning Workshop**

***Evaluation purpose and audience.*** The team should gain a clear, shared understanding of the evaluation's purpose, questions, audience, and uses.

***Programs being evaluated.*** Hold a session familiarizing the team with the assistance programs, organizations, or activities to be evaluated. Have knowledgeable people brief the team about the programs. Set aside time for preliminary document reviews. Advise them about key sources of information.

***Evaluation plans (scope of work).*** Have the team review the scope of work together and ask questions, in order to develop a common understanding of the tasks ahead. They should jointly review data collection and analysis methods, expectations for the evaluation report, and plans for briefings.

***Team workstyles, roles, and workplan.*** Teams -- especially diverse teams -- should discuss their individual work preferences and agree on how they will work together (e.g., styles of decision-making and negotiating disagreements, work hours/holidays). They also must decide on procedures for communicating (e.g., regular meetings) and dividing work among themselves. They should jointly develop a workplan and schedule of tasks.

***Relationship to the management group.*** Clarify the team's relationship to the donor management group. What communication and reporting channels will there be? How will the independence of the team's evaluation findings be protected?

interest. A big advantage of document reviews is its low cost and practicality. Findings are typically world-wide, rather than being based on only a handful of country case studies that may or may not be representative of broader experience.

Problems may occur, however, if the evaluation documents provided to the team have not been carefully screened for their quality or relevance, or when they are submitted late. Teams may then have to sift through large volumes of mostly irrelevant materials in too short a time period. Moreover, since existing documents have been prepared for other purposes, they may not specifically address the joint evaluation's key questions and extracting needed information is not always possible, easy, or successful. Another problem may be that existing evaluation reports may not cover newer program areas and approaches of interest. A final limitation of document reviews is that the methodology does not easily lend itself to participatory approaches directly involving recipient country officials, program stakeholders or beneficiaries.

While not all these limitations can be easily overcome, document reviews can be made more efficient and effective by ensuring the teams get documents sufficiently early, and that these documents have been carefully selected based on their quality and relevance to the joint evaluation's key questions/issues. If language is an issue, translation of key documents should be planned in advance. Combining document reviews with additional methods is often a good practice, if time and resources permit.

***Questionnaires and/or interviews with agency officials.*** Because document reviews may not fully answer the joint evaluation's specific issues and may be somewhat out-of-date, they can be usefully supplemented by other information-gathering techniques, such as sending questionnaires to be completed by donor agency officials, or by holding interviews with them.

Success of the questionnaire approach depends on the response rate, its timeliness and on the completeness and accuracy of the questions answered. Some questions may require extra work on the part of the responding donor agency offices, and may or may not be fully completed.

Interviews or face-to-face meetings with donor agency officials may get better responses, but there are additional costs, such as allocating team time to conducting the interviews and associated travel costs.

### *Phase 2: Fieldwork in developing countries*

Many joint evaluations have used desk studies as a first phase to help them better frame the issues and get first impressions on findings, followed by a fieldwork phase in selected developing countries. While adding to cost and level of effort, a fieldwork phase will also certainly add to the timeliness, quality, and credibility of the evaluation's findings. The fieldwork enables the team to "ground-truth" the preliminary desk study findings and to gather information first-hand on the evaluation questions (e.g., how assistance programs are performing) in specific developing country contexts.

***Selecting countries and programs for fieldwork.*** When selecting developing countries for fieldwork case studies, it is a useful practice to devise selection criteria to help ensure important aspects are covered. For example, if the scope of the evaluation is worldwide, countries may be selected to represent all the key regions. Also, countries and programs may be selected because they receive assistance from the donors conducting the joint evaluation. If the evaluation is a comprehensive assessment of a broad sector, selection criteria might ensure that examples of all key program strategies or approaches for achieving sector goals are included, so that their comparative merits can

be evaluated. The exact criteria to be used, of course, should vary depending on the focus, scope or other aspects of the specific joint evaluation.

**Selecting sources and methods for gathering information in the field.** Once countries and programs have been identified, joint evaluation teams usually gather information from sources such as government officials, implementing agency managers and field staff, local experts and key informants, and various other program stakeholders and beneficiaries.

All too often, teams do not directly gather information and views from the intended program beneficiaries. While it requires extra cost and time, gathering information directly from beneficiaries is often very enlightening. They sometimes provide surprisingly different perspectives on a program's performance, problems, and solutions than using other sources alone would suggest.

There are a variety of rapid appraisal methods that teams can use to gather information systematically from these sources. While relatively low-cost and fast, these methods still provide a structured approach to gathering useful information from beneficiary and other stakeholder groups. Techniques, such as direct site observations, focus group discussions, community group interviews, key informant interviews, participatory workshops, and mini-surveys, can be employed by joint evaluation teams.

In cases where it is important to gather quantitative information on exactly "what" results are being achieved by a program (e.g., what percentage of an intended beneficiary farmer group are using a new agricultural technology, or how their incomes are affected), a sample survey approach might be required. However, if qualitative information is wanted, such as an in-depth understanding of "why" a program is not benefiting people as intended, then a rapid appraisal method may be a better choice. In many cases, especially where expensive surveys are involved, the methods to be used have probably already been determined and budgeted in the evaluation's scope of work.

Selection of methods may also depend partially on the evaluation's purpose. If it is primarily to learn lessons or make recommendations for program improvements, it is usually appropriate to use participatory approaches and rapid appraisal methods. If, on the other hand, the primary purpose is to provide an objective and convincing assessment of performance, it may be better to use more rigorous methods and less participatory evaluation processes (i.e., less direct involvement by implementing agency staff, recipient government officials, or other potentially biased stakeholders).

**Recording information.** Regardless of which data collection approach or approaches are selected, it is important for team members to document or record the information they gather systematically. Each method has its own data collection instrument that is used for this purpose. For example, surveys employ structured questionnaires, site observation techniques use observation forms; focus groups use loosely structured interview guides to record information, etc. This helps ensure that different team members, who may be gathering information independently, follow a common format. If two team members are present during an interview or focus group, it is useful for one to do the talking or facilitating, while the other concentrates on recording and note taking. Laptop computers are particularly useful for purposes of recording information in the field, sharing the information among team members, and facilitating analysis of the information.

**Communicating.** While it makes sense for team members in the field to split up occasionally to do different tasks, they should set aside regular times for team meetings -- to share information, experiences and views, to review progress, and to decide on next steps. Especially when teams are highly diverse, frequent meetings can help everyone to appreciate the value of different perspectives. With continuous interaction, reaching team consensus on evaluation findings will be easier. If one of

the purposes of the joint evaluation is to share and learn from each other's evaluation expertise and methods, these meetings can contribute to this process as well.

If a number of teams are concurrently conducting fieldwork in different locations (e.g., for different country case studies or for other sub-component studies) as part of a larger joint evaluation effort, there is also the challenge of communicating across these various team efforts. Sometimes this is a special role assigned to a team coordinator, or it may be left up to members of the joint evaluation core team (who are usually leading the individual country teams). Emails, conference calls, and planned workshops at critical points in the process may be used to facilitate cross-team sharing of preliminary findings and conclusions, or to discuss common problems, gaps, or other issues facing the teams.<sup>13</sup>

Finally, the teams in the field will need to arrange for communication with the donor management group, so that issues arising concerning interpretations of the scope of work can be clarified, progress and preliminary findings can be reported, and problems requiring guidance can be resolved.

***Preparing evaluation reports in the field.*** Teams preparing country case studies will find it useful, once their data gathering phase is over, to hold a workshop, in which all the team members participate and are actively involved in the analysis and reporting of results. It is useful to divide this process into three sequential (separate, but related) steps of discussing and agreeing on (1) the key findings (the empirical facts gathered), (2) the conclusions (the team's interpretations and judgements in light of the findings) and (3) recommendations (proposed actions based on the conclusions). This approach is usually to be favoured over one in which the team leader works individually with different members on different parts of a report. A workshop approach typically results in a better, more integrated product that is endorsed by the whole team. It usually takes better advantage of the differing perspectives of a diverse team. Moreover, tensions which may arise between a team leader and individual team members concerning the nature of the findings and conclusions can be managed more effectively if the whole team is brought together to discuss them. It also helps to agree ahead of time on a process for resolving disagreements among team members concerning the evaluation's results.

Country reports (or sub-component study reports) should follow the format outlined in the scope of work. Typically, it is useful to distinguish between findings, conclusions, and recommendations. Executive summaries are often required, not only because they help the readers to focus on key results, but because such summaries may help facilitate the synthesis process later.

***Sharing the draft report for comment in the field.*** Once the team has a preliminary draft report, it is good practice to share the draft for comment with those organizations and key individuals in the field that have been especially involved in the evaluation. Sharing the draft widely -- and allowing adequate time for comments to be received and incorporated -- is important for confirming the accuracy of the findings, for flushing out additional information and views, and for ensuring a more credible report. If time is short, the team might consider holding a "stakeholder" workshop, at which the report's findings and conclusions may be presented, discussed, and commented on by the workshop participants.

### *Phase 3: Preparing the synthesis report*

In complex joint evaluations with multiple sub-components or country case studies, a final phase is typically the preparation of an overall report that synthesizes findings from the various sub-studies. Generally this is the task of the core team, although it may also be assigned to a single individual, such as a team coordinator.

In a number of joint multi-donor evaluations, the synthesis report has relied heavily on the executive summaries of the sub-studies as an effective way to organize and present results; e.g., with summaries serving as individual chapters for the synthesis report. This approach is usually least controversial and most favored by the teams that prepared the sub-studies.

Another approach is for those preparing the synthesis report to go "beyond" the contents of the sub-studies to arrive at overall findings, conclusions and recommendations that may not appear in the earlier reports. This approach has the benefit of building synergies and becoming more than the mere sum of its parts. On the other hand, such an approach may well be challenged and contested by those who served on the sub-teams. Such controversy may be minimized if the synthesis report is prepared by a core team that also participated in the sub-study work (as opposed to a team or individual that was not involved in the various component efforts). Ownership may also be built via a participatory workshop approach to framing the final report's overall findings, conclusions, and recommendations, to which all the various sub-team members are invited.

### ***8. Reviewing, communicating, and following-up on evaluation results***

Final steps in an effective joint multi-donor evaluation involve:

- Reviewing and revising the evaluation report's results by the management group and by the sponsoring donor agencies.
- Publishing and distributing the reports to their intended audiences and communicating results by other means.
- Following-up (e.g., monitoring whether or not the recommended actions are being taken).

***Review procedures.*** Evaluation reports submitted by the joint evaluation team typically undergo reviews by the management group and also often by the larger sponsoring group of donor agencies as well. If communication between the team and the management group has been good all along, then the report's findings should come as no big surprise.

Such reviews and discussions of the draft report will help build a sense of ownership by these groups. The reviews and consequent revisions also generally improve the quality of the final report. If the evaluation's independence is an issue, then any revisions may be subject to previously agreed-on restrictions, such as not changing the substance of the team's findings, conclusions, or recommendations. If an evaluation's results are controversial, the review process can be a tricky affair, especially if procedures for resolving disagreements haven't been worked out ahead of time. It can be helpful in such circumstances if a respected official from one of the participating agencies, who has been intimately involved with the evaluation, to present the results to the group in a convincing manner. It also helps if the team's conclusions and recommendations are closely tied to and supported by the empirical evidence (findings).

Whereas in the past, reviews could take a considerable amount of time, today the Internet is making the process easier and faster. The draft report can be emailed to reviewers around the world and comments emailed back. Sometimes, however, especially with an initial draft or potentially controversial report, it may be better to review the report in a meeting or workshop setting, so there is an opportunity for discussion among reviewers and a forum for settling disagreements.

***Communicating results.*** Reports on country case studies or other sub-component studies are usually published by the responsible lead donor agency. Often one of the donor agencies in the management

group offers to publish the final synthesis report, or it may already have been covered in the evaluation's budget, e.g., under an administrative contract.

Distribution traditionally involves mailing copies of the published reports to its intended audiences. Joint multi-donor evaluation reports typically have very broad distributions, including officials of the sponsoring donor agencies, participating recipient country officials, program managers, other stakeholders and interested parties. Recently, some of these reports are also being made accessible to broader audiences within the development community -- via the Internet.

In addition to published reports, a joint evaluation's results can be effectively communicated orally in settings such as international conferences or workshops. It is especially important on such occasions to invite relevant operational managers from the donor agencies, recipient country institutions, NGOs, and other organizations and decision-makers towards whom the evaluation's recommendations are directed. An advantage of such international forums is that participants can discuss the evaluation's findings, conclusions and recommendations among themselves, and agree to take appropriate actions. This can promote collaborative actions and greater program coordination among donors and other partners.

Other marketing strategies for "selling" a joint evaluation report's results within each the sponsoring donor organizations might be considered. Distributing short report highlights, bulletins, or agency notices may help bring management attention to the study's key points. Oral briefings and seminars for managers may also help focus attention within an agency on the joint evaluation.

***Follow-up.*** Because joint multi-donor evaluations usually raise broad system-wide issues and recommendations that relate to a diverse range of organizations, compliance with recommendations cannot be compelled. Monitoring of follow up actions taken by these organizations can be a useful practice -- and may even provide a level of accountability that would otherwise be missing. A well-structured monitoring effort may influence agencies to account for their response --or lack of it -- to the evaluation report. One such example is the Joint Evaluation Follow-up Monitoring and Facilitation Network (JEFF) that was established after the Rwanda evaluation to institutionalize post-evaluation monitoring.<sup>14</sup>

Another aspect of follow-up is how to ensure that donors will undertake future joint evaluations when the need arises -- especially in critical international programs or sector-wide assistance efforts. Most joint evaluations have been initiated spontaneously on the suggestion of one donor agency or another. What sort of "system" might be established that would raise a "red flag" whenever an evaluation is needed? Establishing international units responsible for performance monitoring and evaluation in key areas of international program assistance may be one answer.<sup>15</sup> Routine, on-going monitoring of results achieved by international assistance efforts could alert agencies when performance is off-track, or less than expected, and signal the need for another in-depth evaluation to uncover problems and recommend solutions.

## **Annex 1: The Joint Multi-Donor Evaluations Reviewed**

This annex provides a brief overview of the joint multi-donor evaluations that were reviewed as background for this paper, and from which the effective practices were drawn. These joint evaluations have been grouped into three broad categories (based primarily according to who has sponsored them). Some of their characteristics are discussed below. Specific references for each of the joint multi-donor evaluations that were reviewed are provided in Annex 2.

### **1. Joint evaluations by the Working Party on Aid Evaluation**

These joint evaluations are conducted as part of the Working Party's official work program for the DAC. All Working Party members must agree on a topic for it to become part of the work program. These evaluations typically involve synthesizing members' evaluation findings, experiences, and lessons learned in a particular program, sector, cross-cutting or thematic area of interest to the group. Recommendations and best practices are often drawn to guide members' future assistance strategies. The Working Party synthesis work is often done in close cooperation with other relevant DAC working parties and expert groups, providing cross-fertilization between the groups.

Some recent examples of completed evaluation synthesis studies sponsored by the Working Party that were reviewed include the assessments of Women In Development (WID), Participatory Development and Good Governance (PD/GG), Evaluation Capacity Building (ECB), and NGO performance. The PD/GG study focuses on sector/ programmatic areas, whereas the WID, ECB, and NGO studies examine cross-cutting or thematic topics of interest to the Working Party members. The scope of all of these studies is worldwide.

Although world-wide in scope, these evaluations are relatively low-cost since they generally involve desk reviews of donor agencies' documents or questionnaire responses, rather than primary fieldwork in developing countries. Participation by developing country nationals (e.g., government officials, technical experts, beneficiaries) in these evaluations has been relatively limited as a result. These studies all used document reviews as a primary source of information, although some (WID, ECB) also relied on donor responses to structured questionnaires, and one (WID) conducted selected field trips to interview donor agency staff.

Management of these studies has typically involved establishment of a smaller working group or steering committee comprised of any Working Party member willing to play an active role. One member is usually selected as chairperson. Individual members of the working group volunteer to undertake specific tasks. For example, often a single member will prepare the draft study design (also called terms of reference, approach paper, etc.). In larger studies with several sub-components, often different members will each take the lead on a component, and conduct them concurrently. One member then takes the responsibility for pulling the component studies together into a synthesis report. This approach was taken in the cases of the WID and the PD/GG evaluations. In other studies, for example in the ECB and NGO studies, a single member has taken the lead responsibility for conducting the whole study.

Participation by the larger Working Party on Aid Evaluation membership in the evaluations is most intensive at two critical stages -- review, discussion and approval of the draft study design and of the final report. Members also participate by submitting appropriate documents and by responding to questionnaires. The working group chairperson generally provides study progress reports to the WPE at the biannual meetings.

## **2. Other joint evaluations by large groups of donors**

Like the Working Party evaluation studies, this category of joint evaluation is sponsored by and for the benefit of a relatively large group of donors. Similarly, they often focus on comprehensive, sector-wide assistance programs or activities of multiple donor agencies -- either in a single country, a region, or worldwide. However, these evaluations have not been part of the formal Working Party work plan (although the collaboration and coordination for some may have been facilitated by the existence of the Working Party). Another difference is that these evaluations typically go beyond desk review methods to include fieldwork in developing countries.

Examples in this category that were reviewed include the joint evaluation of emergency assistance to Rwanda, the Special Program of Assistance for Africa (SPA), and European Union Food Aid. The Rwanda evaluation focused on emergency assistance activities of a large number of donor agencies in a single country. The SPA evaluation's focus was on multi-donor coordination mechanisms for quick-disbursing aid to the reform and structural adjustment processes in the Sub-Saharan African region. The EU Food Aid evaluation assessed the performance of program food aid worldwide, including both actions of EU Member States and Community actions.

These evaluations were sponsored by and for relatively large groups of donors (either formal or informal associations of donors). The role of these larger groups of donors in the evaluation process was typically limited to a few key stages, such as review and approval of the design approach (including decisions on evaluation focus/scope, cost-sharing, selection of contractors, etc.), and of the synthesis report findings and recommendations. Day to day management was typically delegated to a smaller management group comprised of those few donor agencies taking lead responsibilities in the study, or to one organization designated to manage the study. The Rwanda evaluation is a case where several donors each took lead roles in preparing study sub-components, whereas the SPA evaluation is a case where a single agency (World Bank) took the primary responsibility for conducting the whole study on behalf of the broader group of donors. In the EU Food Aid study, the Netherlands commissioned the Overseas Development Institute (ODI) to serve as the core evaluation team (responsible for the desk study, the synthesis report, and overall coordination), while other EU Member States financed individual country field studies.

Compared to Working Party evaluations, the joint evaluations reviewed in this category involved more direct fieldwork in developing countries, and a broader range of methods, including not only document reviews but also interviews and data collection in developing countries. Greater participation by developing country nationals typically took place, by including them on the evaluation teams where appropriate, by using them as key informants, or including them on advisory panels.

## **3. Joint evaluations by small groups of donors**

Joint evaluations have also been initiated by relatively small groups of two or more donors that share a common interest in a particular program, sector, thematic, or other aspect of assistance activities -- often within one country or region. They may focus on a specific jointly implemented or co-financed activity, the assistance programs of a single agency, or on the multiple assistance activities of the sponsoring donor agencies within a particular country and sector. Some examples of joint evaluations by small groups of donors that were reviewed for this paper are discussed below.

### **(a) Joint evaluations of multilateral agencies**

During the 1990s, a number of joint evaluations were conducted by small groups of donors focused on assessing the program performance of specific multilateral agencies. Examples include the joint evaluations of the United Nations Children's Fund (UNICEF), the United Nations Population Fund (UNFPA), and the World Food Program (WFP).

Typically, these joint evaluations have been co-sponsored by 3-4 donor agencies for the purpose of providing them and their governments (often facing policy decisions on future funding of these multilateral agencies) with an assessment of the multilateral agency's relevance, effectiveness, efficiency, and/or impacts. A second purpose was often to provide the multilateral agency with lessons to improve its operations.

The usual approach to managing these evaluations was to make the evaluation departments of the sponsoring donor agencies jointly responsible, with representatives from each holding meetings with the evaluation team regularly or at key points as a study progressed. The terms of reference were generally prepared jointly by the donor agencies. Sometimes one of the donor agencies would be designated as overall coordinator, providing various sorts of administrative and logistical support for the evaluation effort. Independent consultants or consulting firms -- agreed to by the sponsoring donor agencies -- comprised the evaluation team. The sponsoring donor agencies agreed to the independence of the teams, acknowledging that the findings would be those of the team authors and not necessarily reflect their agencies' views.

The evaluations usually consisted of several phases, beginning with a desk study reviewing documents and interviewing staff at the multilateral agency headquarters. Results from this phase would then feed into a second phase consisting of fieldwork in selected developing country case studies, focusing on the performance of the multilateral agencies' programs in these country contexts. Country reports would be drafted which would then be used in preparing a final synthesis report with overall findings, conclusions, and lessons/ recommendations. The study usually had a small core evaluation team responsible for the overall study, and then added other technical or sector specialists to the individual country teams as needed. Typically the team members were nationals of the sponsoring donor countries, although in some studies local consultants from the developing countries were hired to advise and support the country teams.

### **(b) Joint evaluations by the Nordic countries**

The Nordic countries have cooperated extensively in development programs and in evaluation activities for years. As their evaluation efforts increasingly focus on main development goals and on comprehensive sector reviews, the impetus for collaborative evaluations has increased even further. Examples of joint efforts by Nordic countries include evaluations of the Beira Port Transport System (BPTS) program in Mozambique, of the Nordic Africa Institute (NAI), and of the United Nations Research Institute for Social Development (UNRISD).

As exemplified in these three examples, joint Nordic evaluations are usually undertaken to review the performance of development programs or institutions to which the Nordic donor countries have significantly contributed. Drawing lessons based on past experience and making recommendations for future Nordic country involvement is the primary purpose of the evaluations.

The typical approach taken in these evaluations is to have one of the participating Nordic donors take responsibility for managing and coordinating the study on behalf of the others. This donor drafts the terms of reference (TOR) which is then reviewed and agreed to by the others. The funding of the

evaluations are shared by all the Nordic countries. Usually one consulting firm or institute is selected to conduct the evaluation. A small team (3-6) comprised of a team leader and appropriate specialists are selected based on suggestions by the Nordic donors. At least in one case, the TOR called for team membership to include recipient country nationals.

Methods in two of the cases (NIA and UNRISD) were limited to desk reviews involving document reviews and interviews with personnel from the institution being evaluated and from Nordic donor agencies. In a third case (BPTS program), a desk study involving document reviews and interviews was followed by a fieldwork phase in Beira, Mozambique to visit and assess selected projects funded by the Nordic donor countries. In all cases, draft reports were submitted for review and comment to the institutions/programs being evaluated and to the Nordic donor agencies. Final evaluation reports considered these comments but remained the independent work of the team authors (and thus do not necessarily reflect the views of the sponsoring donor agencies nor of the institutions being evaluated).

These joint evaluations have benefited from the long history of collaboration in development assistance and evaluation activities among the Nordic countries, leading to shared development goals, strategies and perspectives. These commonalities make agreement on terms of reference and scopes more easily achieved.

## References

### 1. Joint evaluations by the Working Party on Aid Evaluation

DAC Working Party on Aid Evaluation, *Lessons Learned on the Evaluation and Synthesis Process from the Women in Development Assessment Study*, Note by the Chair of the Working Group on the WID Assessment, DCD/DAC/EV(95)14, September 1995.

DAC Working Party on Aid Evaluation, *Assessment of DAC Member's WID Policies and Programs, Main Report*.

Danida, *Issues in Aid Evaluation: Evaluation Capacity Building, Donor Support and Experience*, prepared for the DAC Working Party on Aid Evaluation, 1996.

Finland Ministry of Foreign Affairs, *Searching for Impact and Methods: NGO Evaluation Synthesis*, prepared for the DAC Working Party on Aid Evaluation, 1997.

OECD, *Evaluation of Programs Promoting Participatory Development and Good Governance: Synthesis Report*, 1997.

### 2. Other joint evaluations by large groups of donors

“An Anatomy of an Evaluation Process”, Evaluation of European Aid Managed by the Commission, 1995-1999. Submitted by France on behalf of the Steering Group for the Evaluation of European Aid.

Dabelstein, N. (Danida), “Evaluating the International Humanitarian System: Rationale, Process and Management of the Joint Evaluation of the International Response to the Rwanda Genocide” in *Disasters*, Volume 20, Number 4.

European Commission (1999), *Joint Multi-Donor Evaluation of European Union Aid Managed by the Commission 1995-1999*, European Commission, Brussels, Belgium.

World Bank, Operations Evaluation Department, *The Special Program of Assistance for Africa (SPA): Independent Evaluation of the SPA as a Mechanism to Promote Adjustment and Development in Sub-Saharan Africa*, 1997.

### 3. Joint evaluations by small groups of donors

#### (a) Joint evaluations of multilateral agencies

CIDA, Audit and Evaluation Division, *Review of Lessons from the Process of Evaluation of Multilateral Organizations*, December 1993.

*Joint Multi-Donor Evaluation of the United Nations Capital Development Fund (UNCDF)*, commissioned by Belgium, Denmark, France, Japan, Norway, Sweden, Switzerland and the Netherlands, 1999.

*Evaluation of the United Nations Population Fund (UNFPA): Synthesis Report*, 1993.

Evaluation of UNICEF: *Strategic Choices for UNICEF: Synthesis Report*, 1992. Undertaken by AIDAB, CIDA, DANIDA and SDC.

*Evaluation of the World Food Program: Main Report*, 1994, undertaken by Canada (Canadian International Development Agency), Netherlands (Ministry of Foreign Affairs) and Norway (Royal Ministry of Foreign Affairs).

Monique Kamphuis, “The Multi-donor Evaluation of UNICEF: Methodology Used”, in *UNICEF Evaluation Newsletter*, No. 14, May 1993.

#### (b) Joint evaluations by the Nordic countries

Evaluation of United Nations Research Institute for Social Development (UNRISD), February 1997.

Evaluation of the Nordic Assistance to the BPTS Program, 1995.

Evaluation of the Nordic Africa Institute, 1997.

## Endnotes

1. Donor Experience with Joint Evaluations: A Typology and Lessons Learned (May 1998), prepared by Annette Binnendijk, USAID PPC/CDIE, for the Development Assistance Committee, Working Party on Aid Evaluation.
2. For example, USAID's policy guidance strongly encourages the use of participatory and collaborative evaluations with other donors, partners, and stakeholders. See USAID Automated Directives System (ADS), section 203.5.6a.
3. For more discussion on the advantages and limitations of joint evaluations, see DAC Working Party on Aid Evaluation, Synthesis of Discussion on Joint Evaluations, a note submitted by the Delegation of Canada, March 1992 [DCD/DAC/EV(92)2] and Continued Progress in Joint Evaluation, a note submitted by the Delegation of Japan, September 1991 [DCD/DAC/EV(91)10].
4. Review of the DAC Principles for Evaluation of Development Assistance: Final Report, Conclusions and Recommendations (1998), prepared by consultant Haven North for the Working Party on Aid Evaluation (see page 8).
5. An exception may be if a donor agency thinks it could benefit from having other donors participate in the evaluation of one (or more) of its own programs. For example, if it wishes to include more diverse perspectives on the evaluation team, wants to learn first-hand about other donors' evaluation techniques, or feels the evaluation's findings will be more credible with other donor participation.
6. See Niels Dabelstein, "Evaluating the International Humanitarian System: Rationale, Process and Management of the Joint Evaluation of the International Response to the Rwanda Genocide", in *Disasters*, Volume 20, Number 4.
7. Sector Matrix, submitted by Danida to the DAC Working Party on Aid Evaluation, 31st Meeting, 27-28 January 1999.
8. The donor agency that chairs the working group often also provides overall administrative support, takes the lead in developing the evaluation's scope of work, and synthesizes the study's overall findings, although others might also undertake these tasks, thus spreading the workload.
9. Different evaluation methods (case studies, sample surveys, analyses of existing data, etc.) have distinct features that make them either more or less appropriate for answering a particular type of question. For example, to answer the question "What percentage of farmers in a country have obtained credit from donor assistance programs?", a sample survey would be appropriate. If the question is "Why aren't farmers using the credit programs?", use of a rapid appraisal method, such as holding focus groups of farmers, would be more appropriate. In practice, a combination of methods may be selected, either to improve the persuasiveness of the findings, or to answer different evaluation questions.
10. See Niels Dabelstein, *op. cit.*
11. Holding a team planning workshop for a joint evaluation may be complicated if members are coming from many different parts of the world. If a workshop for all team members is not feasible, separate workshops for the core team or for sub-teams working on individual sub-components or on country case studies, is an alternative.
12. For example, see the joint evaluations of UNFPA (1993), UNICEF (1992), and WFP (1994).
13. Given that joint evaluation teams may be quite large, with complex structures spread out over several continents, holding cross-team meetings at critical stages may require advanced planning and budgeting. Nevertheless, it is useful to have all members get together at key stages of the evaluation process, if possible, in order to take advantage of the full team's diversity of perspectives and expertise. If cost or other factors prohibit such meetings, other forms of communication among different teams should be considered, such as teleconferences, or "virtual" meetings via the Internet.
14. See Niels Dabelstein, *op. cit.*
15. For example, lessons might be drawn from the systems set up by the Special Program of Assistance for Africa.