

Joint Evaluations: Recent Experiences, Lessons Learned and Options for the Future

The DAC Network on Development Evaluation has long been in the lead of promoting joint evaluations as a tool for increased participation and ownership, rationalisation of the process of evaluation, reduced transaction costs for partner countries, improved quality of the work undertaken and increased weight and legitimacy of the evaluation. In 2004, the Network commissioned this study on joint evaluations which would focus on recent experiences with joint evaluations, new and evolving issues and the partner country perspective. The report was written by Dr Horst Breier, consultant, and presented for discussion at the third meeting of the Evaluation Network on June 2-3 2005. The Network is now producing a short publication - *Guidance for Conducting Effective Joint Evaluations* - which will be completed in 2006. This working paper contains detailed information collected from members and complements the Guidance publication which is directed to the wider development community.



The Network on Development Evaluation is a subsidiary body of the Development Assistance Committee (DAC) at the OECD. Its purpose is to increase the effectiveness of international development programmes by supporting robust, informed and independent evaluation. The Network is a unique body, bringing together 30 bilateral donors and multilateral development agencies: Australia, Austria, Belgium, Canada, Denmark, European Commission, Finland, France, Germany, Greece, Ireland, Italy, Japan, Luxembourg, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, United States, World Bank, Asian Development Bank, African Development Bank, Inter-American Development Bank, European Bank for Reconstruction and Development, UNDP, and the IMF.

For further information on the work of the DAC Evaluation Network, please visit the website www.oecd.org/dac/evaluationnetwork or email dacevaluation.contact@oecd.org

TABLE OF CONTENTS

JOINT EVALUATIONS: RECENT EXPERIENCES, LESSONS LEARNED AND OPTIONS FOR THE FUTURE	
ABBREVIATIONS AND ACRONYMS	5
SUMMARY NOTE AND ISSUES FOR CONSIDERATION BY THE DAC	8
INTRODUCTION AND BACKGROUND	12
CHAPTER 1: JOINT EVALUATIONS REVISITED: RECENT EXPERIENCE AND NEW EVIDENCE	15
1.1 Toward a revised typology: How joint is “jointly”?	15
1.2 Gauging the magnitude of the subject: Big or small?	18
1.3 Justifying the efforts: Why complicate life?.....	20
Overarching policy reasons	22
Evaluation strategy motives	22
Developmental motives.....	23
Learning motives	23
Managerial, administrative and financial motives	24
1.4 Talking money: Are transaction costs worth it?	26
1.5 The political economy: Do big fish eat small?	30
CHAPTER 2: KEY STEPS IN PLANNING AND CONDUCTING JOINT EVALUATIONS: AN UPDATE.....	38
2.1 Upstream planning requirements for joint evaluations	38
Making principles and policies transparent.....	39
Clarifying purpose, objectives, focus and scope.....	40
Establishing a set of ground rules	41
2.2 Setting up and organising joint evaluation work	44
Creating a governance and management structure	45
The terms of reference	47
Costing and budgeting for joint evaluations.....	48
Bidding and contracting for joint evaluations	50
Coping with the legal issues involved in joint evaluations	52
Using modern communications technology	53
2.3 Implementing joint evaluations	55
Steering committees.....	55
Management groups.....	57
Consultants	59
Quality assurance.....	62
Field work	63
Crises	64
2.4 Following up on joint evaluations.....	66
CHAPTER 3: OPTIONS FOR THE FUTURE – A CRUCIAL CHALLENGE FOR THE DAC	69

3.1	Improving the existing practice of multi-partner evaluations.....	70
3.2	Enhancing developing country involvement and ownership.....	71
3.3	Focusing multi-partner evaluation work in the DAC.....	72
ANNEX 1 JOINT EVALUATIONS BY FOCUS/SCOPE SINCE 1990		74
ANNEX 2 TERMS OF REFERENCE - JOINT EVALUATIONS: RECENT EXPERIENCES, LESSONS LEARNED AND OPTIONS FOR THE FUTURE		93
	Background and Objectives	93
	Scope of the Study	94
	Approach and Methodology	95
	Timing.....	95
	Management structure.....	96
	Budget and Finance.....	96
ANNEX 3 REPORT OF THE WORKSHOP ON JOINT EVALUATIONS CHALLENGING THE CONVENTIONAL WISDOM - THE VIEW FROM DEVELOPING COUNTRY PARTNERS.....		97
	Introduction.....	97
	Rationale.....	97
	Context and Background	97
	Workshop Programme.....	103
	Participants List.....	104
ANNEX 4 – LIST OF PEOPLE MET		105
ANNEX 5 - BIBLIOGRAPHY AND REFERENCES		111

Boxes

Box 1. The Rwanda Evaluation - Ten Years After	14
Box 2. Joint Evaluations: Perceived Benefits and Obstacles Identified	21
Box 3. World Bank Partnerships in Joint Country Assistance Evaluations.....	25
Box 4. Joint Evaluation of the Netherlands’s Mixed Credit Programme in China	32
Box 5. The South African Joint Evaluation Model	34
Box 6. 16 Golden Rules for Consultants	60

ACKNOWLEDGEMENTS

This report has been prepared by the author on the basis of an extensive review and analysis of literature and written material, including official documents, evaluation reports, manuals, and similar papers. More importantly, however, this report owes a lot, if not most of what it can contribute to the debate on joint evaluations, to the many discussions with evaluators from bilateral as well as multilateral aid agencies, with other aid officials, with representatives from developing country governments and the research and NGO communities, and with many consultants who have been involved in joint evaluation work. All these discussions were characterised by a great deal of interest in the subject matter, and by candour and frankness. Therefore, the author would like to express his gratitude and feeling of indebtedness to all those who took the time to meet and share their knowledge, experiences and opinions with him.

A special word of thanks goes to Niels Dabelstein of DANIDA, the untiring advocate of joint work in evaluation, who strongly supported the launching of this study. His firm belief in the importance of joint evaluations has been encouragement and challenge for the author. Thanks are also due to Hans Lundgren of the Development Co-operation Directorate of the OECD who provided backstopping, encouragement and useful observations and comments and who took part in some of the missions to capitals of member countries and headquarters of international organisations. And finally, the author owes a lot to the support of Sebastian Ling in the OECD Secretariat who did the final editing of this report, and to Michelle Weston, colleague of Hans Lundgren, who put the draft report into an attractive and reader-friendly format.

The views expressed in this report do not reflect the official position of any of the DAC members or of any of the other organisations, institutions or persons contacted in the course of this work. These views are only and exclusively those of the author, and so are the mistakes.

ABBREVIATIONS AND ACRONYMS

AFD	Agence Française de Développement
AfDB	African Development Bank
ALA	Asia and Latin America Aid Programme of the EU
ALNAP	Active Learning Network for Accountability and Performance in Humanitarian Action
AusAID	Australian Agency for International Development
CCA	Common country assessment
CDF	Comprehensive Development Framework
CIDA	Canadian International Development Agency
CODE	Committee on Development Effectiveness (of the World Bank)
DAC	Development Assistance Committee
DRG	Development Research Group (of the World Bank)
EBRD	European Bank for Reconstruction and Development
ECG	Evaluation Cooperation Group (of the International Financing Institutions)
ECHO	European Community Humanitarian Office
EIB	European Investment Bank
EU	European Union
EUHES	European Union Heads of Evaluation Services
GBS	General Budget Support
GEF	Global Environment Facility
IADB	Inter-American Development Bank
IBRD	International Bank for Reconstruction and Development
ICB	International Competitive Bidding
ICPD	International Conference on Population and Development
ICRC	International Committee of the Red Cross
ICVA	International Council of Voluntary Agencies
IDC	International Development Cooperation (of the National Treasury of South Africa)
IDP	Internally Displaced Person
IFAD	International Fund for Agricultural Development
IFC	International Finance Corporation

IFI	International Financing Institution
IFRC	International Federation of Red Cross and Red Crescent Societies
IMF	International Monetary Fund
IPDET	International Programme of Development Evaluation Training
IOM	International Organization for Migration
IPPF	International Planned Parenthood Federation
IsDB	Islamic Development Bank
ITC	International Trade Centre
JBIC	Japan Bank for International Cooperation
JEFF	Joint Evaluation Follow-up Monitoring and Facilitation Network
JICA	Japan International Cooperation Agency
KfW	Kreditanstalt für Wiederaufbau
MDG	Millennium Development Goal
MEDA	Mediterranean aid programme of the EU
MIGA	Multilateral Investment Guarantee Agency
NCSTE	National Centre for Science and Technology Evaluation (China)
NGO	Non-Governmental Organisation
NORAD	Norwegian Agency for Development Cooperation
NZAID	New Zealand Agency for International Development
OECD	Organisation for Economic Co-operation and Development
OED	Operations Evaluation Department
PRIO	Peace Research Institute Oslo
PROAGRI	Agriculture Sector Programme in Mozambique
PRSP	Poverty Reduction Strategy Paper
SDC	Swiss Development Cooperation
Sida	Swedish International Development Cooperation Agency
SPA	Special Programme of Assistance to Africa
ToR	Terms of reference
UN	United Nations
UNAIDS	Joint United Nations Programme on HIV/AIDS
UNCDF	United Nations Capital Development Fund
UNDAF	United Nations Development Assistance Framework
UN/DHA	United Nations Department of Humanitarian Affairs
UNDP	United Nations Development Programme

UNECA	United Nations Economic Commission for Africa
UNEP	United Nations Environment Programme
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNHCHR	United Nations High Commissioner for Human Rights
UNHCR	United Nations High Commissioner for Refugees
UNICEF	United Nations Children's Fund
UNOCHA	United Nations Office of the Coordination of Humanitarian Affairs
UNPFA	United Nations Population Fund
UNRISD	United Nations Research Institute for Social Development
USAID	United States Agency for International Development
WFP	World Food Programme
WHO	World Health Organization
WID	Women in Development

SUMMARY NOTE AND ISSUES FOR CONSIDERATION BY THE DAC

1. Joint evaluations are an evolving and dynamic area of development cooperation. They have been on the international development agenda since the early 1990s but have been of increasing frequency and significance in recent years. The Development Assistance Committee (DAC) at the OECD has been in the vanguard of promoting the idea of more joint evaluation work as part of its broader agenda on enhancing donor coordination and cooperation. In 2000, the DAC Evaluation Network published the booklet *Effective Practices in Conducting a Joint Multi-Donor Evaluation*. In 2004, the Network commissioned this new report to update the existing guidance, review and analyse more recent experiences, and to include the perspective of developing country partners.
2. The aim of this report is to build understanding of joint evaluations: what we mean by the term, what the benefits and challenges are, and how the benefits can be maximised and the challenges minimised or overcome. The report also puts forward long-term and strategic recommendations on joint evaluations.
3. The first section of the report proposes a new typology for joint evaluations. This typology, based on the degree and mode of jointness, has three overall categories: (1) Classic multi-partner evaluations; in which participation is open to all stakeholders and all participate on equal terms; (2) Qualified multi-partner evaluations; in which participation is only open to a limited number of potential partners; and (3) Hybrid multi-partner evaluations; encompassing a range of different and more complex ways of joint working. A greater degree of clarity and understanding of these various modes of evaluation partnerships (and acknowledging the complex hybrid forms) will help reduce confusion and misunderstanding when partners attempt to work together on a joint evaluation.
4. The report also lists and reviews the joint evaluations that have been undertaken since 1990. This work demonstrates that joint evaluations are a dynamic area of development cooperation. Their frequency has increased since 1990, with particularly rapid growth (in numbers, participants and scope) over recent years. However, the review does not indicate a systematic pattern to explain why certain evaluations are undertaken jointly and at particular times.
5. Joint evaluations have the potential to bring strong benefits to all partners and stakeholders. They offer opportunities to harmonise and align the overall processes of evaluation, to build participation and ownership, to share the burden of work involved, to increase the acceptance and legitimacy of findings and recommendations, for mutual capacity building and learning between the partners, and to reduce the overall number of evaluations undertaken – thereby reducing transaction costs and administrative demands on aid recipient countries. However, joint evaluations also generate their own particular challenges and difficulties: the various partners in the evaluation may have different approaches, political objectives or even hidden agendas - and building consensus and agreement between the partners can be both expensive and time consuming. Development agencies have therefore taken different approaches to joint evaluations – with some prioritising joint working and others remaining more focussed on their own independent evaluation activities. Denmark, the Netherlands and Norway as well as Canada appear to be the most committed of the DAC donors to this mode of work.
6. There are other interests and motives that could persuade a donor to become part of a joint evaluation. The first is that the evaluation may be addressing a subject matter that is of priority interest to

the donor – for example the evaluation of general budget support has attracted a wide range of partners because this aid modality is of such focal interest to members of the international development community. Agencies may also join an evaluation because they are committed to increasing harmonisation and alignment of their work programmes. Some agencies join multi-partner evaluations because their own capacity is too limited to meet all the evaluation needs of the agency.

7. The main disincentive to participate in a joint evaluation is the perceived cost. There is no doubt that joint evaluations can be expensive. The direct costs for a large and complex joint evaluation, especially if it includes a number of case studies in developing countries, can easily reach over one million Euros. However, to look at the sheer volume of expenditure alone is misleading. Cost must be correlated with the number of partners that are contributing funds. The World Bank states that, “joint evaluations neither increase nor reduce financial costs for donors”. However, alongside the direct costs of a joint evaluation are the indirect costs - such as staff time and travel. Psychologically, the indirect costs are often predominant in shaping the perception of whether or not a joint evaluation is seen as ‘heavy’ and expensive.

8. In recent years, the international debate on development cooperation has focused on questions such as ownership, harmonisation, alignment, and mutual accountability. Cooperation between donors and partner countries on joint evaluations is one way of working towards these goals. Accordingly, representatives of partner countries participated in this study, both through individual consultations and through a workshop in Nairobi on 20-21 April 2005. The workshop captured prevailing developing country perceptions of evaluation and, specifically, of joint evaluations. The emerging picture has a number of distinct features:

- A strong feeling of frustration with the present state of affairs, especially as regards the level of partner country participation in evaluation work;
- A growing awareness among developing country representatives of the need for them to play a proactive role in setting and implementing the evaluation agenda in their countries;
- A clear understanding of the opportunities and benefits, as well as of the problems and challenges, of evaluation work and of carrying it out jointly with donors;
- A clear interest in learning from evaluation models and success stories in other developing countries; and
- A number of concrete proposals and steps that should be taken to strengthen ownership in the area of evaluation.

9. Participants from South Africa, Tanzania and Vietnam informed the workshop about their national approaches to strengthening ownership of monitoring, review and evaluation. South Africa, for example, has initiated a series of evaluations of individual donor programmes, carried out jointly between the National Treasury and the respective donor. Tanzania, on the other hand, has decided to turn a number of performance assessments and review processes into joint activities between the Government, the donors and the Independent Monitoring Group. Vietnam has a clear government strategy to provide the legal preconditions necessary to establish more participation and ownership in evaluation. Overall, the Nairobi workshop indicates a growing dynamism towards more national leadership and ownership of joint evaluations. Developing countries must now take on this challenge of re-balancing the political economy of joint evaluations in their favour.

10. Chapter 2 is the most substantial part of this report. It distils lessons learned from joint evaluations to improve the planning, organisation, management, implementation and follow-up of future multi-partner evaluations. The chapter looks at four key areas: (1) upstream planning requirements; (2) setting up and organising joint evaluations; (3) implementing joint evaluations; and (4) following up on joint evaluations. A variety of detailed and practical recommendations are put forward, to help evaluation managers overcome the challenges of joint working and maximise the potential benefits.

11. The international community has agreed the importance of monitoring the various indicators that enable us to measure progress towards the MDGs and the Paris Declaration on Aid Effectiveness. Moreover, new modes of assistance such as sector-wide approaches, general budget support and other collaborative multi-donor programmes are creating a growing need for joint work in monitoring and evaluating their implementation. Thus, one might expect a mushrooming of new multi-partner evaluations focused on the new aid modalities. However, apart from the ongoing GBS evaluation led by the UK and a few other examples including the UNDAF evaluation of UNEG members, the evaluation of the Comprehensive Development Framework and evaluation work on the PRS process, there is only limited progress towards taking up this challenge.

12. Without the work and efforts of the DAC, and especially of its Evaluation Network, the idea of joint evaluations would not be so firmly rooted in development thinking and practice. The DAC remains the obvious forum for donors to carry the debate on joint evaluations forward. Members of the DAC Evaluation Network are strongly urged to take on this challenge. The report identifies a range of specific issues for joint evaluations that will need to be addressed in the DAC in order to maintain evaluation's important position and focus within the overall international development agenda:

- The new development paradigms – MDGs, PRSPs, harmonisation, alignment, aid effectiveness, and so on – provide a strong case for joint evaluation work. However, some donor evaluation units have been hesitant to take this agenda forward. Therefore, the Evaluation Network as well as the DAC itself must take up and move forward this agenda. A number of questions must be urgently addressed: Will there be more joint evaluations in future – in response to the needs arising from the new aid paradigms? What would this imply for traditional evaluation work? Or will the number of joint evaluations remain stable, but with a new focus on different subjects and with enhanced emphasis and thrust?
- Is there a need for a DAC role in identifying priority areas and subjects and coordinating joint evaluations? If so, at what level of the DAC should this be taken forward? Who would make the proposals - and who would approve them? Would such a role for the DAC entail the risk of politicising decisions on joint evaluations? How could any risk of impinging on the independence and impartiality of evaluation units be avoided?
- How can a more broad-based constituency in support of joint evaluations be built among a wider range of DAC members - to ensure that the burden of joint work is shared more equitably among the DAC members?
- Should the DAC continue to deal with the whole range of questions connected to joint evaluation work? Or should it concentrate on fewer subjects, for instance on those linked to the new development paradigms, and perhaps a few other subjects of crucial importance (such as quality standards or the evaluation of development effectiveness), and leave the rest to individual members or other donor groupings?
- How can the risk of duplication of evaluation work, including in joint evaluations, between different donor groupings (DAC, ECG, UNEG, EU, Nordics, and Utstein etc) be reduced? What

role should the DAC play in better networking between these groups - with a view to attaining synergies and value added?

- Should the DAC play a role in encouraging members to use multi-partner evaluations to experiment with new forms of evaluation work, such as impact and ex post evaluations, longitudinal studies, and others?
- Should the Development Cooperation Directorate of OECD, on behalf of the DAC, play a focal role in collecting data on joint evaluations, maintain an inventory of them, provide information on lessons learned and good practice, and become an institutional memory for the donor community for joint evaluations. This would need funding by DAC members.
- Should the DAC agree to the compilation and publication of another, relatively short manual on how to organise and run a multi-partner evaluation, based on the lessons learned and good practices contained in this report?
- Should the DAC commission short technical papers on specific issues related to joint evaluation work - for example on the legal questions involved in establishing financing pools; on minimum requirements for consultant contracts; on different options for bidding procedures for consultancy services; on the assessment of bidding proposals in an effective and transparent fashion; etc?
- Finally, there is the question of the DAC Development Evaluation Network assuming a more proactive role in the planning and implementation of meta evaluations. Under DAC guidance and supervision, these evaluations would bring together dispersed evaluation knowledge, validate it and feed it into planned or ongoing international processes. Meta evaluations would also help to identify areas of sub-optimal evaluation coverage - perhaps, for example, in the proliferation of individual country strategy and programme evaluations; which are largely disconnected from each other and risk losing sight of the common aid effort.

13. During the extensive consultations for the preparation of this report, one question surfaced constantly: Is there a future for joint evaluation work - and where will this future lie? The answer is yes - there is an important future for multi-partner and joint evaluations. However, for this future to be realised DAC donors and the evaluation community must be responsive to the new modalities in development cooperation, more participatory and open to developing country ownership, and more accountable in its role and purpose as a crucial element in the global effort to fight poverty and realise the MDGs.

INTRODUCTION AND BACKGROUND

14. Joint evaluations have been on the international development agenda since the early 1990s. The Development Assistance Committee (DAC) at the OECD has been in the vanguard of promoting the idea of more joint evaluation work as part of its broader agenda on enhancing donor coordination and cooperation. The *DAC Principles for Evaluation of Development Assistance*, adopted by DAC Ministers for development cooperation and heads of aid agencies in 1991, state that, “joint donor evaluations should be promoted in order to improve understanding of each others’ procedures and approaches and to reduce the administrative burden on recipients”. The principles also underline the importance of involving the aid recipients as fully as possible.

15. Although there was agreement at the international policy level to promote joint evaluations, progress in implementing and delivering joint evaluations, at the level of aid agency headquarters, evaluation units and country offices has been uneven among the various agencies. Some agencies showed reluctance to participate in joint evaluations, probably because they were cautious of the staffing and resource implications that a more proactive stance might entail. Other agencies, however, were more forthcoming, and these more enthusiastic organisations have carried much of the burden of joint evaluation work by providing funding and making available staff resources to organise and lead the joint evaluations.

16. A core of knowledge about joint evaluations, and experience with them, was gradually built up within the evaluation units of aid agencies as well as in international fora such as the DAC. Moreover, some of the joint evaluations undertaken in the 1990s became flagship examples of the importance, relevance and usefulness of joint work. These contributed significantly to a broader acceptance of the real value of joint evaluations. Flagship initiatives included the tripartite evaluation of the World Food Programme by Canada, the Netherlands and Norway (1994), the evaluation of EU Food Aid approved by the Council of Development Ministers (1997), and, of course, the much celebrated Rwanda evaluation; *The International Response to Conflict and Genocide: Lessons from the Rwanda Experience* (published, in five volumes, in 1996). In a small number of other cases however, including the somewhat notorious evaluation of EU aid in the second half of the 1990s, joint evaluation work proved more problematic and brought grist to the mills of those sceptical about the approach.

17. In 1998, the *Review of the DAC Principles for Evaluation of Development Assistance*, commissioned by the DAC Working Party on Aid Evaluation (Now: DAC Network on Development Evaluation), was published. Regarding joint evaluations, the report concluded that the 16 members who had participated in joint evaluations, “found them highly – or, more often occasionally – satisfactory” (page 55). Furthermore, it was pointed out that joint evaluations “have proven to be satisfactory as they allow first-hand learning from each other, give greater results, facilitate feedback, mobilise knowledge, improve follow-up and save resources” (ibidem). On the other hand, respondents also voiced their reasons for concern, namely “higher costs, since [joint evaluations] require more time and resources to assure co-ordination and foster mutual understanding. Hidden agendas, different approaches, too general and diplomatic conclusions as they have to combine different interests, increased complexity and delays and different political objectives, also work against effective joint evaluations” (ibidem). In summary, although a stronger consensus in support of joint evaluations had emerged in the 1990s, the somewhat ambiguous attitude toward the various benefits and challenges of joint evaluations had not been completely abandoned among the DAC community.

18. The DAC Evaluation Network continued to lead the debate on joint evaluations into the 21st century. Two consecutive chairs, Niels Dabelstein from Denmark and Rob van den Berg from the Netherlands, as well as the members of their respective Bureaus, were strongly committed to the idea of joint evaluations. The difficulties of joint working were not disputed, but these were considered challenges that needed to be taken up and addressed in order to overcome them. It was in this spirit that the Network approved the proposal to produce a publication providing guidance on how to plan and conduct a joint donor evaluation. The emphasis was on practical guidance so that the publication could serve as a useful tool for agencies planning and delivering joint evaluations. Annette Binnendijk, a consultant to USAID, was commissioned to do the study, which was published in 2000 under the title, *Effective Practices in Conducting a Joint Multi-Donor Evaluation* in the DAC Evaluation and Aid Effectiveness Series.

19. Since 2000, a significant number of joint evaluations have been undertaken, including: *Joint Evaluation of the Road Sub-Sector Programme Ghana* (2000) initiated by Denmark; *Toward Country-led Development* (2003) a World Bank initiated evaluation of the Comprehensive Development Framework; *Local Solutions to Global Challenges: Towards Effective Partnership in Basic Education* (2003) initiated by the Netherlands; *Addressing the Reproductive Health Needs and Rights of Young People since ICPD – The Contribution of UNFPA and IPPF*, led by Germany (2004). Other joint evaluations are still ongoing or only recently finished. These include the evaluations of IFAD, of the *Enabling Development Policy of WFP*, of *Assistance to Internally Displaced Persons*, of the *International Trade Centre*, of the *Triple C Concept in EU Development Co-operation policy*, and of *General Budget Support*, led by the UK.

20. The ongoing evaluation of General Budget Support (GBS) is the first major joint effort to address the challenge of evaluating this new aid modality, which does not enable donors to easily disaggregate and evaluate their own individual contributions. The GBS evaluation is therefore an important, if not crucial, attempt to respond jointly to the challenge of demonstrating the results of this new parameter in international development co-operation. Other parameters that are radically changing development co-operation include the Millennium Development Goals (MDGs), the Rome Declaration on Harmonisation and the Paris Declaration on Aid Effectiveness, and, at the country-level, Poverty Reduction Strategy Papers (PRSPs) and Sector Wide Approaches (SWAs).

21. In this context of the new parameters in development co-operation and the relatively large number of joint evaluations initiated in recent years, with some of them experimenting with exciting approaches to governance and process, the DAC Evaluation Network decided that there was an acute need to review and analyse recent experiences with joint evaluations. This report therefore aims to supplement the *Effective Practices in Conducting a Joint Multi-Donor Evaluation* by identifying good practices along with emerging issues and new challenges in joint evaluations and options for the future.

22. The following chapters present: an overview and analysis of recent experiences and new evidence from joint evaluation work (Chapter 1); an update of effective practices in conducting joint evaluations (Chapter 2); and a review of emerging trends and options for joint evaluation work in the future, including some key issues for discussion within the DAC (Chapter 3). A look at some of the issues involved in the typology of joint evaluations, an annotated overview of joint evaluations carried out between 1990 and the present day, and a number of boxes illustrating key features and issues of joint evaluation work are also included.

23. This report is different from previous work on the subject in that it endeavours for the first time to include the views and observations of partners in the South. As an important component of this study, a workshop with partner country representatives was held in Nairobi on 20-21 April 2005. The findings from the workshop have been incorporated in this report, and the workshop proceedings are attached at Annex 3.

24. This report is primarily addressed to the donor community, as represented in the DAC, which commissioned it. It is hoped, however, that it will also prove to be of use to the broader development community, both in the North and in the South, including the multilateral development system, newly emerging donor governments, the academic and research community, evaluators and consultants.

Box 1. The Rwanda Evaluation - Ten Years After

In late 1994, on the initiative of the evaluation department of Danida, representatives of bilateral donors, UN agencies, and international NGOs agreed to sponsor a multi-partner evaluation - *The International Response to Conflict and Genocide: Lessons from the Rwanda Experience*. Commencing in January 1995, the evaluation was undertaken over a 15-month period by an international team with 52 consultants and researchers. The team produced four studies along with a synthesis report covering all phases and aspects of the crisis.

Ten years after the genocide and eight years after publication of the Joint Evaluation, Danida commissioned a study to assess the follow-up of the recommendations. This study comes to the following overall conclusion:

“The critical test is whether reports and policy prescriptions, explicitly attributed to the Joint Evaluation, get translated into practice. The assessment has revealed a number of areas where the Joint Evaluation had a positive influence and impact. It has also revealed recommendations that were not implemented that remain valid and warrant further efforts to implement them. Even allowing for the achievements in the humanitarian sector in relation to accountability, standards and greater professionalism, the views of interlocutors, the literature and the examination of the Darfur case on the central issue of the prevention and suppression of genocide and massive human rights abuses are on balance pessimistic. Several interlocutors proposed that massive public interest mobilization campaigns would be required to put sufficient pressure on decision makers in key countries to get action on an issue like genocide prevention and intervention. The successful global campaign against landmines demonstrated what can be achieved by such campaigns.”

Quote by John Eriksson and John Borton from the Journal *Den Ny Verden*, Copenhagen, 2004

CHAPTER 1: JOINT EVALUATIONS REVISITED: RECENT EXPERIENCE AND NEW EVIDENCE

1.1 Toward a revised typology: How joint is “jointly”?

25. It is difficult to define the term joint evaluation in one single, comprehensive definition that would satisfy all stakeholders. Conventional wisdom might suggest that joint evaluations are efforts undertaken by a group of donors, working together in a systematic and targeted manner, to obtain evidence of the achievements and failures of development co-operation activities and/or to assess the quality of - mostly multilateral - institutions. Although this kind of definition has, to date, prevailed in discussions, including those in the DAC and its Evaluation Network, it is a somewhat over-simplified definition, which is also challenged by partner countries as witnessed at the Nairobi workshop.

26. This kind of definition is found in the DAC publication *Effective Practices in Conducting a Joint Multi-Donor Evaluation* (page 7) which was published in 2000: “These [multi-donor evaluations] are evaluations of development assistance programs or activities conducted collaboratively by more than one donor agency.” However, the *DAC Glossary of Key Terms in Evaluation and Results Based Management*, published two years later, defines joint evaluations as “an evaluation to which different donor agencies and/or partners participate” (page 26). The explicit reference to partners refers back to the 1991 DAC Principles for Evaluation of Development Assistance but also signifies, to some extent, an opening up of the definition to reflect the present focus on partnership and ownership in development thinking.

27. The research carried out for this study, especially the discussions with evaluation officials and practitioners identified a strong interest in further clarification of what we mean by the term joint evaluation. Similarly, the DAC Evaluation Network expressed a strong interest in the development of a more detailed typology of joint evaluations. Greater differentiation in our use of the term is needed to reflect the increasing number and growing variety of joint evaluations (e.g. global, regional, national, sector, thematic, etc.). Furthermore, the evaluation community is facing additional challenges in evaluating the new partnership aid modalities such as budget support and SWAps. Joint evaluations may offer a suitable way of working together to evaluate these more joined-up and partnership-based aid delivery mechanisms, but we must first understand what we really mean by the term joint evaluation.

28. This study therefore looks at different options for a typology that would help to sharpen our understanding of the range and variety of activities covered by the term joint evaluation, and facilitate better assessment of their potential utility. A good starting point for this work is Binnendijk’s *Effective Practices in Conducting a Joint Multi-Donor Evaluation*, which lists a range of criteria that could be used to categorize joint evaluations. These include:

- Number of participating donors, that is the multitude of actors involved
- Approach to management and kind of implementation
- Focus (projects, programmes, sector-wide, cross-cutting themes, etc.)
- Scope (single country, region, worldwide)
- Purpose (lesson learning, accountability)
- Methodologies (desk work, field studies)
- Partner participation

29. Binnendijk decided to use the first criterion - the number of participating donors - for the purposes of her study. She breaks this down into three sub-categories:

- 1) Joint evaluations undertaken by the DAC Working Party on Aid Evaluation (now: DAC Network on Development Evaluation). These are usually desk-based meta-evaluations, which distil and share good practice for the use of all members of the DAC. Only a very limited number of these meta-evaluations have been undertaken.
- 2) Joint evaluations undertaken by a large group of donors.
- 3) Joint evaluations undertaken by a small group of donors (typically three or four).

30. This typology remains useful and pragmatic; the three categories are well understood and it is relatively easy to attribute all joint evaluations to one or another category. However, this typology looks solely at the number of donors involved, and does not reflect either the level of partner involvement or the varying focus, scope, purpose, methodology or approach (the other potential criteria listed in *Effective Practices in Conducting a Joint Multi-Donor Evaluation*) of joint evaluations. A more complex and detailed typology is called for if we are to understand and reflect the various key aspects of different joint evaluations.

31. This study has taken the approach that focus and scope are particularly appropriate and useful criteria in creating an overall summary and grouping of the totality of joint evaluations that have been undertaken, as per the overview presented in Annex 1. However, this study also proposes a typology of joint evaluations based on the mode of how actors actually work together (How jointly is 'joint'?). These partnership modes are of a varied nature; ranging from the full participation of a wide range of actors to more restricted forms of participation. This variety may help to explain some of the confusion that can occur in discussions on joint evaluations. A greater degree of clarity and understanding of the various modes of partnerships in evaluations (and acknowledging complex hybrid forms) will help reduce confusion and misunderstanding when partners attempt to work together on a joint evaluation.

32. The following table suggests a relatively simple typology for joint evaluations, based on the degree and mode of jointness:

TYPE OF EVALUATION	MODE OF WORK/EXAMPLES
1. Classic multi-partner	Participation is open to all stakeholders. All partners participate and contribute actively and on equal terms. Examples include the Rwanda evaluation, the tripartite evaluation of WFP, the UNFPA/IPPF evaluation, the GBS evaluation, and so on.
2. Qualified multi-partner	Participation is open to those who qualify in the sense that there may be restrictions or the need for "entry tickets" - such as membership of a certain grouping (e.g. EU, Nordics, UNEG, ECG, Utstein) or a strong stake in the subject matter of the evaluation (e.g. active participation within a SWAp that is being evaluated). Examples include the various EU aid evaluations, the evaluation of the road Sub-sector in Ghana, the Basic Education Evaluation, the ITC evaluation.
3. Hybrid multi-partner	This category includes a wide range of more complex ways of joint working. For example: (1) Work and responsibility may be delegated to one or more agencies while other actors take a 'silent partnership' role; or (2) Some parts of an evaluation may be undertaken jointly while other parts are delivered separately; or (3) Various levels of linkage may be established between separate but parallel and inter-related evaluations; or (4) The joint activities focus on agreeing a common framework - but responsibility for implementation of the evaluation is devolved to different partners

33. The typology presented above replaces the somewhat narrow term “multi-donor” with the more inclusive term “multi-partner”. This term leaves it open as to whether partners in any of the three evaluation categories are donors only, or donors and aid recipients. This is in line with the view vigorously expressed at the Nairobi workshop that multi-partner evaluations can be either: donor–donor or donor–recipient or recipient–recipient partnerships. However, participants at the Nairobi workshop were also determined to reserve the term “joint evaluation” for those evaluation activities which are carried out jointly by donor(s) and recipient(s) – and to not include donor-donor evaluations. Although this demand may need to be discussed further, it could help to add some precision to the debate. Generically, we would talk of the different forms of multi-partner evaluations indicated above. Within this categorisation, there would be two possible differentiations consisting of “joint evaluations” (donor – recipient) and of “multi-donor evaluations” (donor – donor).

34. It should be mentioned that some consideration was given to the possibility of including, under the hybrid forms of joint evaluations, the trust fund arrangements between bilateral and multilateral agencies in the area of evaluation. However, the idea was not progressed because a trust fund does not automatically lead to joint evaluation work. Funds are used for a range of purposes, including the regular evaluation activities of the recipient institution, capacity development in the trustee institution (or of developing country partners), training (such as the annual IPDET course at Carleton University in Ottawa, Canada) and also, on certain occasions, for joint evaluations. A notable example of a trust fund being used to support joint work is the CDF evaluation; which was supported by the trust funds established in OED by the Netherlands, Norway and Switzerland¹, in addition to a special trust fund set up for the CDF evaluation.

35. One of the striking findings of this typological review is that there are very few examples of “classic” multi-partner evaluations. Many more evaluations belong to either the qualified multi-partner type or to one of the many hybrid forms of partnership. Consequently, there is little sign of a functioning marketplace open to everyone for promoting and agreeing multi-partner evaluations – a space where ideas for such evaluations could be flagged and tested and potential partners identified. This has significant implications for the ways and means existent to promote multi-partner evaluations. Accordingly, various efforts undertaken in fora such as the DAC Evaluation Network to create a more dynamic marketplace for brokering joint evaluations and for bringing together organisations that could work together, have been of limited success. Most multi-partner evaluations are of a qualified, that is restricted nature - and are therefore not open to everyone. Consequently, large international bodies are probably not the most efficient platforms for promoting concrete cases of evaluation work and for bringing the right partners together.

36. Instead, most multi-partner evaluations are generated by individual actors negotiating through their bilateral or group channels to solicit support and find partners. It may therefore be useful to look into the possibilities of networking more systematically between different groups of actors, such as the DAC, EU, UNEG, ECG, and the Nordics when it comes to multi-partner evaluation work. Such networking would also need to imply a better division of labour between the various groups and the development of a common framework and of ways to identify priority evaluation subjects that would be of interest to more than one group and therefore become a good candidate for joint work.

37. This short discussion of issues related to typology is intended to contribute to more analytical rigour in discussing joint evaluations and organising the debate around them. It should help bring some

1. In 2003, Switzerland carried out an assessment of the SDC–OED Partnership Programme. The results of this assessment are rather mixed. Objectives of such partnerships are not always clearly defined, so that it is difficult to measure the benefits achieved. Also, the transaction costs can be quite high. Therefore, the consultant strongly recommends “to reorient and reshape the partnership [in phase II] in order to reach a more balanced degree of satisfaction on both sides”.

precision to policy discussions, especially with our partners in the South, and to better understanding the various modes of partnership. Acknowledging the existence of many complex hybrid forms of multi-partner evaluations will contribute to reducing confusion and misunderstanding when partners attempt to work together. Perhaps, this typology can also help streamline and better target international efforts to deliver joint evaluations in a more efficient manner. The DAC Evaluation Network will remain the largest donor forum to discuss joint evaluations. It may decide, however, to concentrate its future work on certain categories of joint evaluations only, and to leave the others to smaller groups of donors, recipient countries, partner country level aid arrangements, or individual actors.

1.2 Gauging the magnitude of the subject: Big or small?

38. Early on in the research for this study, the question was raised as to how many joint evaluations have actually been undertaken. No satisfactory answer, to this seemingly innocent question, was readily available. However, it is essential to know how big a subject we are dealing with in order to assess its importance. Therefore, one of the challenges for this study has been to find an answer to that question - how many joint evaluations are we really talking about? – and to provide an overview and summary of the joint evaluations that have been undertaken to date. It was decided to fix the time period under review from 1990 up to 2004-05. A very large share of all the joint evaluations ever undertaken is included within this timeframe. It was also decided to include all types of joint evaluations within the summary.

39. Previously, no study has seriously endeavoured to keep track of joint evaluations undertaken, or to collect key information such as the participating countries and agencies, thrust, scope, costs, and so on. This study, however, attempts to present a summary and overview of the joint evaluations that have been undertaken in Annex 1.

40. It should be noted that the search for the needed information was greatly facilitated by a range of informants at the level of agency headquarters and central evaluation units, but it became much more difficult to obtain information on joint evaluations at the country level. Many agencies have decentralised much of their development activities - aid missions, field offices and embassies are often no longer required to obtain prior approval for joint evaluations that are implemented at the country level, or to report on them to headquarters. Therefore, knowledge of joint evaluation initiatives is increasingly dispersed, and Annex 1 should by no means be considered as comprehensive or final. However, it was decided that the time and resources needed to undertake extensive country-level research could not be justified for this report.

41. The DAC Evaluation Network may wish to consider whether or not remedial action is needed and warranted to address this lack of comprehensive, reliable, and easily accessible information on joint evaluations. Some initial ideas of how to address this situation are included in Chapter 3: *Options for the Future*.

42. The organising principle of the table in Annex 1 is a set of seven categories, under each of which the relevant evaluations are listed chronologically.² The seven categories have been identified on the basis of focus and scope of the evaluation. They are:

- Global Policy Evaluations
- Global Impact and Effectiveness Evaluations

2. There are cases of evaluations that could be subsumed into different categories. One example is the Peace Building Study of the so-called Utstein countries which has been included in category 1 *Global Policy Evaluations*, but could equally well be subsumed under category 3 *Thematic and Sector Evaluations*.

- Thematic and Sector Evaluations
- Institutional Evaluations
- Country Strategy and Country Programme Evaluations
- Specific Project and Programme Evaluations
- Joint Evaluations with Partner Countries

43. So far, 53 joint evaluations have been identified for the period 1990-2005. Some observers may view this as a large number of evaluations, others as a more modest one. As a share of the overall total of evaluation work, the number of joint evaluations carried out by development agencies is comparatively small, but nevertheless important and significant. Moreover, while participation in a joint evaluation may be a big burden in terms of staff resources and funding for one agency, another agency may find it much easier to absorb the same level of involvement. So, the perception of magnitude is also a reflection of the capacity and constraints of individual agencies and their evaluation units.

44. The distribution of joint evaluations across the different categories is relatively even, with one exception: Joint evaluations with partner countries ranks top with 12 entries. The other categories come out as follows: Country strategy and Country Programme Evaluations, Global Policy Evaluations, and Institutional Evaluations - 8 entries each; Thematic and Sector evaluations and Specific Project and Programme evaluations - 6 entries each; the category of Global Impact and Effectiveness Evaluations has a slightly lower rate of occurrence - 5 entries.

45. We should not interpret too much into these numbers as they represent a relatively small statistical unit of reference. However, there are some conclusions which can be drawn from Annex 1:

- 1) The table highlights the varied and challenging nature of joint evaluations, and the many different actors working on the different types of evaluations.
- 2) The table shows that joint evaluations are a dynamic area of development cooperation. The frequency of joint evaluations has increased since 1990; particularly rapid growth (in numbers, participants and scope) can be observed over recent years. This is most likely a result of both the new paradigms for development cooperation - with greater emphasis on partnerships and joint work) - and the strong interest that many members of the DAC Network on Development Evaluation have consistently shown in the topic over recent years.
- 3) The table also indicates a growing number of experiments in the field of joint evaluations. These experiments focus primarily on the processes of joint evaluations and on improving the ways agencies work together. In this context, a strong emphasis is placed on finding ways and means to reduce (in priority order): the transaction costs of joint evaluations (especially with regard to staff absorption), the length of time needed to produce results, and the resource needs. Initial findings in these areas are reflected in Chapter 2 on Lessons Learned.
- 4) The table shows a very positive trend in the increasing number of evaluations with (full or at least fuller) partner involvement. This applies especially to evaluations launched within the last three or four years. This statement, however, does not leave room for complacency - a great deal of further improvement is needed when it comes to partner involvement (as shown in section 1.5 of this chapter, where the political economy of joint evaluations is looked at more closely in the light of the outcome of the Nairobi workshop with developing country representatives).

- 5) It is not possible to recognise a pattern in the table, which would help to explain why particular joint evaluations are undertaken at a certain time. The table instead suggests a high degree of *ad hoc* joint evaluations with random selection of topics. A random approach does have merits, allowing scope for individual initiative and usually resulting in a very serious commitment by those who take part in the evaluations. On the other hand, a potential drawback of this approach can be that urgently needed joint evaluation work is not progressed, falling between the cracks because there is no agreed mechanism to identify and prioritise, in a systematic and transparent manner, the different opportunities for joint evaluations. Such a selection mechanism would be particularly helpful when resources are limited and tough decisions need to be taken about which potential joint evaluations should be progressed.
- 6) Likewise, the table does not provide strong clues with regard to future trends in the orientation of joint evaluations (perhaps with one exception: the World Bank's striving toward more country and programme evaluations in partnership with recipient country governments or regional development banks).
- 7) It is perhaps surprising that the table does not indicate more interest among donors in joining their forces to evaluate the new and innovative aid delivery mechanisms, such as SWAps and GBS (the very few examples of such evaluations listed in the table include the ongoing GBS evaluation and the PROAGRI evaluation in Mozambique).
- 8) Finally, the table shows that different DAC members are displaying different levels of commitment to joint evaluation work. Denmark, the Netherlands and Norway appear to be the most committed - with very significant contributions to joint evaluation work. These three countries are also top aid performers in terms of ODA/GNP ratio. The same is true for Canada. A second and larger group consists of the European Commission, Germany, Japan, Sweden, Switzerland, the United Kingdom and the United States. These countries contribute to a number of joint evaluations, also in a lead role from time to time, but not as extensively as the first group. The rest of the DAC members, roughly half of the total membership, have more limited participation in joint evaluations; taking an active role occasionally or rarely. Among the multilateral institutions that are observers in the DAC different levels of commitment are also evident. The lion's share of contributing to joint evaluations is undertaken by the World Bank and UNDP.

1.3 Justifying the efforts: Why complicate life?

46. The question of why DAC members would or would not want to become part of joint evaluations has a long tradition of debate in the DAC Evaluation Network. There is, for example, a DAC document prepared by Canada for the meeting of the DAC Expert Group on Aid Evaluation in March 1992³ which includes sections on "Perceived Benefits of Joint Evaluations" and "Obstacles Identified" (Box 2 below). Similarly, the Review of the DAC Principles for Evaluation of Development Assistance of 1998 devotes quite some room to this issue, as do room documents of more recent origin prepared for the meetings of the Evaluation Network in 2003.⁴

3. *Synthesis of Discussions on Joint Evaluations*, Note by the Delegation of Canada, DCD/DAC/EV(92)2 (March 1992).

4. See Room Document 2: *Note on Joint Evaluations*, prepared by Niels Dabelstein, Denmark, and Room Document 3: *Lessons Learned from World Bank Experience in Joint Evaluation*, prepared by Osvaldo Feinstein and Gregory K. Ingram. Both documents were presented at the 37th Meeting of the DAC Working Party on Aid Evaluation, 27 – 28 March 2003.

Box 2. Joint Evaluations: Perceived Benefits and Obstacles Identified

Based on general comments of DAC members, Canada listed the following perceived benefits of conducting joint evaluations in 1992 which may provide a rationale and a level of expectations for members:

1. Joint evaluations result in an important sharing of experiences and increased learning. They provide an opportunity to critically analyse and enhance donor country's evaluation techniques.
2. Joint evaluations appear to have a greater influence on the recipient or executing agency.
3. Joint evaluations are able to bring more diverse talent to the table (e.g. more "eyes and lenses"). They permit a wider scope for the study in terms of financial and human resources.
4. On balance, real savings can be realized through joint evaluations, although the country playing the lead role may incur higher expenses.
5. Joint evaluations, as opposed to separate or concurrent evaluations, reduce the burden placed on the recipient or executing agency by minimizing time in interviews and meetings.
6. Joint evaluations have a tendency to be of higher quality and excellence, because of the vetting process, the composition of the team, and the broader range of political and aid related interests being considered.
7. Joint ventures are of particular interest and importance to smaller donors who are often less able to mount major country reviews which analyze the broad impacts on macro issues from programs such as structural development.
8. Joint evaluations can become an important means to improve the coordination of scarce aid resources in a recipient country.

Obstacles encountered that were listed by Canada, include:

1. Developing a comprehensive, yet manageable terms of reference to accommodate each country's preoccupations and interests was the most frequently mentioned obstacle.
2. The process of achieving a common understanding of the evaluation issues, the critical indicators, and the final recommendations requires time and team interaction to develop openness and trust.
3. Administrative accommodations should be made to respect the existing procedures and legal requirements of each donor involved. Financial arrangements and the selection of consultants are particular problematic areas.
4. Political sensitivities of all parties, including the recipient, are very complex and should be given careful study and consideration at the outset.
5. Final agreement on report findings can become a major obstacle including how divergent views are reported.
6. A large joint evaluation delegation may intimidate a recipient resulting in less than expected results.

47. Sida of Sweden is one of the few DAC donors to indicate in written form some of its policy thinking on joint evaluations in its Evaluation Manual *Looking Back, Moving Forward*, published in 2004. The manual underlines the utility of joint evaluations as "a suitable format for assessing sector-wide approaches and other programmes where contributions of different participating organisations cannot or should not be separated from each other" (page 16). It emphasises that joint evaluations are "likely to have a wider and, in some cases, more powerful impact than an evaluation commissioned by a single organisation" (ibidem). However, Sida is somewhat more hesitant, when it comes to the accountability role

of joint evaluations: “Joint evaluations are used for learning as well as for accountability, although they are perhaps especially useful when the evaluation purpose is to identify good practice and draw useful lessons for the future” (ibidem).

48. Although the above references to the strengths and weaknesses of joint evaluations and to the motivations of DAC members for joining or not joining such activities are anecdotal, they provide a good starting point for a more systematic presentation of the potential benefits of joint evaluation work as well as of the obstacles that those interested in joining a multi-partner evaluation may have to reckon with. The presentation is divided into five categories of reasons for joint evaluation work, namely reasons of (a) overarching policy, (b) evaluation strategy, (c) developmental requirements, (d) learning, and (e) management, administration and finance. The sixth section deals with the obstacles which also need to be taken into account when considering benefits and challenges of joint evaluation work.

Overarching policy reasons

- The DAC agenda on harmonisation, alignment and development effectiveness is calling for more joint efforts of donors and recipients. This is leading to peer pressure among and on donors to do more evaluation work jointly.
- Demonstrating development effectiveness in working towards the MDGs has become a central objective - and challenge - for policy makers and heads of aid agencies. Consequently, they are calling upon their evaluation services to do more joint evaluation work in order to show the results of the common aid effort. Some of these calls have even taken on the form of clear instructions to management and staff to do more joint work.⁵
- The need to evaluate together can also result from public pressure - in the media, in the academic and research community, or in parliaments. Pressure can also originate in recipient country governments and in international organisations.
- Lastly, corporate governance decisions such as mission and values statements may encompass strong stipulations with regard to the desirability of joint work. As a consequence, management and staff are constantly required to look for possibilities of realizing these aspirations.

Evaluation strategy motives

- Analyses, findings, conclusions and recommendations of joint evaluations are based on broader knowledge, a wider range of inputs and contributions, joint rather than individual scrutiny and quality assessment, and multi-partner commitment. Therefore, these evaluations usually carry more weight, credibility and legitimacy and are less easy to ignore.
- Closely connected with the preceding argument, joint evaluations are well suited to promote advocacy for change, especially if some of the “good boys” of development cooperation participate in the evaluation.
- Joint evaluations should be the preferred mode of work if there are issues to be taken up in an evaluation that are too sensitive for one donor alone to tackle.

5. This, for instance, has happened at a recent meeting of heads of agencies of the Nordic Plus group of countries.

- As a rule, Meta evaluations will always require a multi-partner approach as the preferred mode of implementation.
- Similarly, in cases of evaluating multi-partner financed projects, programmes and other development cooperation activities, joint evaluations are the preferred mode of work. The same applies to the evaluation of the work of multilateral institutions.
- Joint evaluations contribute in a very significant way to making evaluation into a transparent and less threatening process.
- Finally, a joint evaluation can be a useful option to evaluate important but controversial development issues. GBS is a case in point. A joint evaluation allows broad participation of many stakeholders, including sceptics.

Developmental motives

- The new modes of aid that emphasise joint donor-recipient efforts and basket or other forms of co-financing require the use of joint evaluations to look at results, outcome and impact. It is next to impossible to perceive meaningful evaluations in this area of cooperation that would allow for a single-donor approach.
- Joint evaluations are a powerful tool for working towards more ownership and participation of developing countries in aid evaluation.
- Joint evaluations are one way of contributing to coordination and harmonisation in the field of evaluation, both among donors and between donors and recipients.
- Joint evaluations help to avoid the danger of conveying to partner countries too many different and often conflicting evaluation messages, which are competing for attention and action and are often hard to reconcile.
- Joint evaluations contribute significantly to rationalising the development process and to making it transparent.
- Joint evaluations have the potential to reduce the transaction costs for developing countries – because these are increased if evaluation activities of different donors are undertaken separately and at different times.

Learning motives

- Joint evaluations are among the most effective tools for evaluation capacity building, for donors as well as for developing countries. Through working together, partners in joint evaluations come to understand better the different perspectives, mandates, approaches and cultures of each of the institutions involved. Partners also have the chance to compare different approaches to the planning and design of evaluations, to the selection of methodologies, and to implementation, including the role of consultants, the clearing and adoption of reports and dissemination and follow-up.
- Joint evaluations are an efficient way of working towards identifying and distilling lessons learned and good practice. They should therefore be given full consideration in all cases where lesson learning is the main focus of the planned evaluation.

Managerial, administrative and financial motives

- Joint evaluations can be of great help in preparing the ground for and informing management decisions, for instance on the funding of specific development activities. Findings from joint evaluations tend to be more readily accepted by management and decision makers, especially with regard to multilateral development work.
- Joint evaluations can be a way of redressing a lack of sufficient evaluation capacity within an agency. Contributing as a silent or low-profile partner to a multi-partner evaluation, can enable the obtaining of insights from that work that otherwise would be difficult and more expensive to get.
- Funding a share of the overall costs of a multi-partner evaluation can help to economize when evaluation funds are scarce. In addition, this may help to disburse funds more quickly – which can be useful in cases where the funds available for evaluation are more generous than the staffing situation.
- High profile contributions to joint evaluations are one way of showing an agency's willingness to assume responsibility for international cooperation. They are also instrumental in demonstrating an agency's determination to accept an international leadership role.

49. This is a long list of strong arguments in favour of joint evaluation work. But there are also obstacles and challenges, which need to be kept in mind when discussing the potential benefits of joining multi-partner evaluation work. These obstacles and challenges include the following:

- Joint evaluations are usually complex and complicated undertakings, with a high degree of coordination needs and the potential for disagreement and conflict, especially during the initial phase of a joint evaluation.
- Joint evaluations require significant inputs of time, negotiating skills and willingness to compromise in order to establish a common understanding of the purpose and objectives of the evaluation, decide on a common framework for action, determine appropriate governance structures, agree on procurement modalities, select consultants, and so on.
- There are also further practical and ethical issues that need to be resolved, such as stakeholder and beneficiary participation, quality standards, independence and impartiality, to mention just a few. They need to be harmonised to the extent possible among partners – and this task can lead to difficult, protracted debates among the evaluation partners.
- Partners in joint evaluations often have one official and another actual agenda. These different agendas can adversely affect the smooth running of the evaluation. The same applies if political objectives differ widely among the partners.
- Partners in joint evaluations may also pursue certain parochial or national interests, for instance in the selection of consultants who come from a particular country or background.
- Methodologies, too, can become a bone of contention. Qualitative versus quantitative methods, the role of consultants in drawing up terms of reference, indicators, evaluation matrices, and many other questions contain a lot of potential for conflict and debate. Often, the underlying issues are core differences in the approach to evaluation, based on questions such as - Is

evaluation a science or an art, or perhaps both? Is the core of a joint evaluation the process or the product? What should be an appropriate balance between accountability and learning?

- Compared with single donor evaluations, joint evaluations often require longer timelines for design and implementation. There are cases of serious delays in producing intermediate or final results. This is not always well received in other departments of the agency; which may be urgently waiting for the evaluation outcome.
- The above points, especially the potentially longer timeline, can lead to higher costs for joint evaluations than for single agency evaluations.
- Finally, joint evaluations put significant burdens on partners, especially on those who assume lead functions. This may constitute a heavy strain on the evaluation resources of an agency - in terms of financial commitment as well as in staff resources, as will be further explained in the following section.

Box 3. World Bank Partnerships in Joint Country Assistance Evaluations

The World Bank Operations Evaluation Department (OED) has made partnership a cornerstone of its values statement. Staff incentives and the department's organization have been aligned with its partnership strategy.

The evaluation partnership was launched in early 2000. Since then, six country assistance evaluations (out of 70 altogether since 1995) were prepared in collaboration with multilateral development banks, including the African Development Bank (Lesotho and Rwanda), Islamic Development Bank (Jordan and Tunisia), European Bank for Reconstruction and Development (Kazakhstan), and Inter-American Development Bank (Peru). In addition, three evaluations were undertaken in partnership with governments and local institutions (Burkina Faso, Tanzania, and Eritrea).

Partnership models take different forms. So far, four types of OED collaboration have been identified: (I) part joint / part parallel country evaluations - with some sections being prepared jointly but with each multilateral development bank assessing its own program; (II) parallel evaluations; (III) incorporating the evaluation findings of other donors into OED's report; and (IV) evaluations undertaken in partnership with governments and local institutions.

In the OED Working Paper "*Partnership in Joint Country Assistance Evaluations: A Review of World Bank Experience*", published in 2005, the author Fareed M. A. Hassan takes stock of past experience and looks in some detail at the benefits and costs of joint country assistance evaluations as well as at the lessons that have been learned. His summing up of benefits and costs is helpful and provides food for thought in the ongoing debate on joint evaluations.

Benefits

- Partnerships have led to significant evaluation capacity development.
- Joint work has led to the sharing of perspectives, lessons learned, methodologies, and better understanding of differing mandates of institutions; this, in turn, facilitated the development of evaluation capacity among partnering institutions.
- The substantial use of local consultants contributed also to a significant amount of local evaluation capacity development and government ownership of the evaluation findings.
- Joint evaluations have promoted discussions of evaluation methods and have encouraged the use of some common evaluation standards, consistent with the broader agenda of harmonisation.

- Joint work has been effective in identifying key constraints and gaps in donor assistance.
- Finally, partnerships have been effective in lowering the transaction costs to recipients of development assistance as the burden of multiple, separate evaluation efforts on recipient country institutions has been reduced.

Costs

- Joint country evaluations may take more time to prepare than anticipated and may cost more. This is due to greater time devoted to coordination, exchange of work programs and plans, joint drafting of common chapters, as well as time taken for comment and review of documents by joint evaluation stakeholders.
- Exchanges with partners reveal that differences in organizational cultures, mandates, and methodologies can impose significant constraint/delays to joint evaluations.
- Delays in the completion of joint evaluations adversely affect their timeliness and reduce their value in providing timely lessons to newly designed country assistance strategies. However, with more experience of partners these costs (both time and financial) may decline.

Lessons Learned

- Flexibility is needed to accommodate the special circumstances of each evaluation/partner, given differences in mandates, organizational cultures, evaluation methods, and work programs.
- When evaluation capacity development is part of the joint evaluation, it should be recognized as a separate objective and planned for in terms of time and cost.
- Despite the importance attached to joint evaluations, their actual number has been rather limited. There is, therefore, a need to actively involve management at donor institutions so that protocols for partnership can be established at highest level. The Multilateral Development Banks Evaluation Cooperation Group (ECG) can facilitate the identification of opportunities for partnerships among donors, as the major purposes of the ECG are to strengthen cooperation among evaluators and harmonize evaluation methodology in its member institutions.

1.4 Talking money: Are transaction costs worth it?

50. Joint evaluations are expensive, both in terms of direct and indirect costs. This is the prevailing perception of aid officials and evaluation managers when they refer to this kind of evaluation work. However, there are also a number of potential benefits – if these did not exist, the frequency of joint evaluations would certainly be much lower.

51. We need to understand better the cost side of a joint evaluation. First of all, it is necessary to disaggregate and distinguish between the various cost components, especially between direct and indirect cost. Furthermore, the specific role of a donor who decides to join a multi-partner evaluation must be taken into account. Will this donor play a prominent, perhaps even a lead role in the exercise? Or will the donor be satisfied to take a back seat and keep a low profile? And thirdly, the focus and scope of the evaluation as well as the number of partners are important.

52. The direct cost of a joint evaluation is usually expressed in the evaluation budget. The cost for a large and complex joint evaluation, especially if it includes a number of case studies in developing countries, can easily run to over one million Euros. Recent multi-partner evaluations of the *Enabling Development Policy of WFP* and of *Addressing the Reproductive Health Needs and Rights of Young*

People since ICPD –The Contribution of UNFPA and IPPF entailed a budget of about one million Euros each. Others, like the Basic Education evaluation, the CDF evaluation, and the ongoing evaluation of General Budget Support required much higher budgets, receiving well over two million Euros in the case of the GBS evaluation. Sums of this magnitude may look quite intimidating to some agencies, especially if their evaluation budgets are relatively small.

53. However, to look at the sheer volume of expenditure alone is misleading. It must be properly correlated with the number of partners that are willing to share the bill, either on equal terms or through a pledging process in which some partners shoulder more of the burden than others. It also has to be linked to the evaluation budgets of each agency. In some cases, total expenditure for a joint evaluation and the individual agency's share in it may look large; in other cases it may be rather more average. The World Bank stated at a meeting of the DAC Working Party on Aid Evaluation in March 2003,⁶ that, according to their analysis "joint evaluations neither increase nor reduce financial costs for donors." According to OED, the share of financial costs for joint evaluations differs little from what its costs would have been for a stand-alone evaluation of the World Bank component of the joint activity.

54. This assessment may not be correct for all donors. For some, the outlay on a joint evaluation may be higher than what they would normally spend on an evaluation. For others, it may actually become cheaper to participate in a joint evaluation rather than trying to independently organise a similar evaluation.

55. A complicating factor in discussing direct cost is that, due to rules and regulations or legal requirements, not all partners are always able to contribute to a pooled funding arrangement. Instead, these donors may offer to provide a contribution in kind, such as a consultant or the financing of a specific component of the overall exercise. In other cases, donors may want to retain an element of independence in a joint evaluation by providing some pool funding but also some separate funding for a specific component of the evaluation, the implementation of which remains under their control. In most of these cases, these arrangements are sub-optimal because they complicate the financing and administration of a joint evaluation, create "patchwork" approaches that are difficult to integrate into the overall work process, influence adversely the degree of common ownership of partners in a joint evaluation, and may even be seen as a restriction on the independence of the evaluation.

56. The impression that the costs of joint evaluations are particularly high, and that they are not always fully under control, is due to the recurrent need in many joint evaluations to increase the budget during the evaluation process. This can be for different reasons. It may be the result of new requests from the steering or management groups for work to be done by the consultants which was not originally foreseen; or of new items added to the budget, such as dissemination and follow up work; or simply of a budgeting process at the beginning which was not sufficiently thorough and did not include all potential expenditure items. As a result of increasing joint evaluation budgets, other evaluation staff may express reproach that joint evaluations appear to be treated more generously than those carried out by the agency individually.

57. Alongside the direct costs of a joint evaluation are the indirect costs. The magnitude of indirect costs is more difficult to assess than that of direct costs. Psychologically, however, the indirect costs are predominant in shaping the picture of whether or not a joint evaluation is perceived as 'heavy' and expensive.

58. Indirect costs usually encompass three main categories of expenditure: (i) the staff time of the evaluation unit (and perhaps also of other branches of the agency); the cost of travel to the meetings of the

6. Feinstein, O., and K. Ingram, *Lessons Learned from World Bank Experiences in Joint Evaluation*, Room Document 3, DAC Working Party on Aid Evaluation, Paris 27-28 March 2003.

evaluation governing bodies and also to workshops, seminars, field visits; and the employment of consultants and/or the hiring of services necessary to cope with the additional work requirements posed by a joint evaluation.

59. Most of the concerns expressed with regard to indirect costs relate to staff time. This is especially true for those donors who assume the lead role in a joint evaluation. Clearly, the size of an evaluation unit is an important element in assessing the strains on staff time. As a rule, a large evaluation unit will find it less difficult to absorb the time spent by staff on a joint evaluation than a smaller one. But even large evaluation units can reach the limits of their absorptive capacity, if they have agreed to act as lead country or as a member of the management group.

60. The Evaluation Department IOB of the Netherlands Ministry of Foreign Affairs estimates that the time spent on the Basic Education evaluation by the key staff member amounted to an average of one third of the working time of this person for close to three years. In other words, one staff member devoted a full year's work to this evaluation. About a third of one staff member's total work time was also needed at the evaluation unit of Germany's BMZ to lead the WFP evaluation and that of UNFPA and IPPF. DFID has allocated two and a half staff to work almost exclusively on the GBS evaluation. Even this allocation does not always seem to meet the requirements. It is, therefore, not surprising, that the DFID Head of Evaluation recently wrote to all members of the GBS evaluation Steering Committee that there was a "clear indication that there is something wrong with the way the bureaucracy is growing".

61. IOB and DFID are both examples of rather large evaluation departments. Danida, on the other hand, with a relatively small evaluation unit, estimates that about 50 per cent of its evaluation programme is dedicated to joint evaluation work. Danida also allocates approximately 50% of its evaluation staff time to joint evaluations. Any downward change in staffing would therefore imply serious repercussions either for the national evaluation programme or for the joint evaluations, or for both. This is not only true for Denmark, which is just one example of a small evaluation unit with a big stake in joint evaluation work. In two other cases, agencies had to reduce their staff inputs to the management of joint evaluation work drastically and at relatively short notice, because there were unexpected changes in the respective staffing situations of the evaluation units or strong internal pressures to devote more staff time of evaluators to national evaluation activities.

62. The issue of internal pressure on evaluation units with regard to the use of the time of their staff is a serious issue in a number of DAC countries. Especially with lengthy and complex evaluations, colleagues can get frustrated with the evaluation managers at the seemingly slow progress of work and the delays in getting results.

63. Similarly, delays in or postponement of other evaluation work has in some cases been attributed to the amount of time being spent on the joint evaluation work. It is therefore essential to devote sufficient attention to a realistic assessment of staff time implications before deciding if and how to join a multi-partner evaluation. It is also important to be as transparent as possible in explaining the allocation of staff time amongst the different evaluation activities.

64. It is highly commendable that, despite the considerable outlays in staff time and the criticism that this may provoke, many agencies continue to be ready to contribute to and support joint evaluation work. However, as pointed out in Section 1.2, the brunt of the burden of joint evaluation work is borne by only about half of the DAC membership. Therefore, a more broad based approach to joint evaluations – with active participation of more donors than at present - would be an important step toward a better burden sharing arrangement within the donor community. It would also help to allay the perception that joint evaluations are a "luxury" in evaluation work, caused by the fact that a significant number of donors stay away from it.

65. When an agency joins a multi-partner evaluation on a low-profile basis (ie with relatively few inputs) its investment of staff time is less heavy and more manageable and therefore less prone to criticism from other parts of the agency. With some contribution to the direct cost and only a limited staff involvement in the evaluation process, the agency may however have gained quite a deal from participating in the joint evaluation.

66. Meetings of management and steering groups, consultations, presentation workshops, and feedback events all require staff to spend time and money on travel. In larger and more complex evaluations, the number of meetings can easily be more than ten over the two or three years of the evaluation process. This may not cause problems for those who do not need to travel far to reach the meetings – but for those who have to travel long distances it can become a significant cost. However, little is known about serious difficulties encountered by evaluation managers with regard to this indirect cost.

67. The same can be said for the third component of indirect costs, namely the hiring of consultants to assist with the logistics and/or the substantive work of an evaluation. Most agencies have sufficient flexibility in their evaluation budgets to contract consultants for this purpose. But even if the intention is to hire consultants in order to offset some of the time that would otherwise need to be invested at the expense of the regular staff resources of an evaluation unit, there are limits to the role that a consultant can play in a management group or in a steering committee. Nevertheless, the hiring of additional capacity seems to be a useful and promising way of reallocating some of the workload which otherwise falls on the regular staff. There is some evidence that the potential of this approach has not yet been fully exploited.

68. Turning to the benefits that would attract donors to join a multi-partner evaluation, it is not easy to arrive at many general conclusions. We have already discussed the wide range of interests and motives that could persuade a donor to become part of a joint evaluation. Burning issues of development cooperation, such as general budget support, which are of concern to many donors provide a powerful incentive to join an evaluation. Similarly, a strong commitment to harmonisation and alignment can be another driving force. The intention to be able to influence the recommendations on a controversial subject is yet another incentive.

69. Some agencies make it clear that they are prepared to join multi-partner evaluations because their own evaluation capacity is too limited to meet all the evaluation needs of the agency. It also happens that agencies join multi-partner evaluations in order to allow for a quicker disbursement of large evaluation budgets that may not be matched by an adequate number of staff to organise and manage evaluations.

70. The evaluation of multilateral organisations is a case where the motives to join the evaluation are relatively easy to identify. A joint assessment of the successes and failures of an international organisation is more likely to be given credibility than an individual assessment. It also helps to lay the foundation for informed decision making on funding or other issues. But there are also other good reasons to join. A country may be hosting the headquarters of an organisation and want to follow the evaluation process and outcomes. Another may wish to join because it is the home country of the chief executive officer of the organisation being evaluated. So, there is a whole range of potential motives for joining any given multi-partner evaluation, and also a range of different expectations as to the benefits that could be drawn from the evaluation.

71. Learning from working with others is a benefit that many hope to realise when they join a multi-partner evaluation. There are even some joint evaluations that are specifically organised for learning – for example some of the joint country strategy or programme evaluations of the World Bank with regional development banks. But it is very difficult to capture and quantify these learning impacts in a way that would meaningfully feed into a cost–benefit analysis of joint evaluations.

72. There is another benefit of joint evaluations that is equally difficult to quantify or substantiate, This is the argument that the findings and recommendations of joint evaluations carry more weight and legitimacy than those of individual evaluations. While joint evaluations can be complex and provide many challenges, they are a useful tool for evaluating something that one donor would find hard to review on its own. In these cases, a joint effort means more muscle, backbone and legitimacy.

1.5 The political economy: Do big fish eat small?

73. In recent years, the international debate on development cooperation has focused on questions such as ownership of the development process by developing countries, genuine partnership between donors and recipients, alignment, participation of as many stakeholders as possible - including beneficiaries, community based organisations, NGOs, and the private sector -, harmonisation of donor procedures, mutual accountability, etc. These debates are fundamentally changing the paradigms of international development cooperation. The OECD Development Assistance Committee (DAC) has been central to this change process – coordinating the agreement of the 2003 Rome Declaration on Harmonisation and Alignment and the 2005 Paris Declaration on Aid Effectiveness, agreed by more than 100 DAC members, partner countries and development institutions.

74. To date, evaluation has not played a central role in this debate - although it is strongly linked to the monitoring, reporting and accountability dimensions of the new aid modalities such as general budget support and SWAps. Paragraph 45 of the Paris Declaration refers to these links and to the continuing, but transitional role of donor monitoring and evaluation until donors can rely more extensively on partner countries' statistical, monitoring and evaluation systems.

75. There are a range of ways for working towards more ownership of monitoring and evaluation by partner countries – but the cooperation of donors and partner countries on joint evaluations is one of the key possible ways forward. Accordingly, the terms of reference for this study stipulate that representatives of partner countries be involved as resource persons, both through individual consultations as well as through a workshop to discuss and validate the main findings and conclusions of the study. The workshop took place in Nairobi, Kenya on 20 and 21 April, 2005. It was attended by 21 participants, 17 of them representing partner countries.⁷ In addition, the National Centre for Science and Technology Evaluation of China (NCSTE) provided a written input by its Executive Director, Professor Chen Zhaoying.⁸

76. The workshop and the core presentations succeeded in capturing prevailing partner country perceptions of evaluation as well as of joint evaluations specifically. The picture that emerged in Nairobi has a number of distinct features:

- A general feeling of frustration with the present state of affairs, especially a dissatisfaction with the level of partner country involvement in evaluation work;
- A growing awareness among partner country representatives of the need for them to play a more proactive role in setting and implementing the evaluation agenda for development cooperation in their countries;
- A clear understanding of the opportunities and benefits, as well as the problems and challenges, of evaluation work and of carrying it out jointly with donors;

7. Bangladesh, Cameroon, Egypt, Ghana, India, Kenya, Mauritania, Nicaragua, South Africa, Tanzania, Uganda, and Vietnam. A full workshop report is included as Annex 3.

8. Chen Zhaoying, *Lessons Learned from a Joint Evaluation between Donor and Recipient: the Netherlands' Mixed Credit Programme in China*, Beijing, undated.

- A strong interest in learning from evaluation models and success stories in other developing countries; and
- A number of concrete proposals and steps that should be taken to strengthen ownership in the area of evaluation.

77. This picture may not be fully representative of the thinking in all partner countries, but it is based on solid evidence emerging during the Nairobi workshop and should be taken seriously. Let us therefore take a look at the picture in more detail.

78. Professor Chen Zhaoying underlines in her paper that development assistance has been moving towards a policy-oriented, recipient-driven approach, with recipient ownership constantly improving over recent years. She then continues: “Unfortunately, however, evaluation practice has been slow in responding to external changes. It can be discovered that the activity that always comes out as most donor-driven is evaluation. Although over the years, the evaluation units of the World Bank, UNDP, DAC, as well as some donor governments have been developing approaches to increase partner country ownership of evaluations, the current situation remains to be that most evaluations of development aid have been led by donors and were done to satisfy donors’ requirements. It is difficult to find a successful case which is generally acknowledged as a recipient-driven evaluation in the development evaluation community.”

79. Joint evaluations, too, do not fare much better in Professor Chen Zhaoying’s judgement, because “at present, ideas on joint evaluation are not lacking but practical ways are missing.”

80. Similarly, many participants at the Nairobi workshop argued that, “the demand for evaluation is donor-driven [and] local partners, including government agencies, undertake evaluations mostly to comply with donor requirements”; that donors are not sufficiently transparent in communicating how, when and why they are planning for certain evaluations; that strong donors tend to dominate the evaluation process, even if they aim for partnership; that partner country representatives are often brought into the evaluation process too late and only after key decisions (ToR, selection of consultants, etc) have been taken; that donors prefer external consultants to nationally available expertise and emphasize the value of their own approaches and methodologies; that their evaluations are often not driven by a set of clear and transparent evaluation questions, but rather by general objectives which may be difficult to turn into a manageable set of evaluation questions; that local stakeholders and beneficiaries are not sufficiently involved in the evaluation process; that “nationals are seldom associated to the critical part of the evaluation, i. e. design, data analysis and interpretation, and the formulation of recommendations”; and that the role of national consultants is often “limited to facilitating contacts with respondents and orientating the team during fieldwork.”

81. The list of concerns is long. One participant summarized the feeling of uneasiness and, indeed, of frustration in the following words: “Evaluations need to be demystified, democratized and simplified”. However, there also was a strong consensus among participants at the Nairobi workshop that all the blame should not fall on the donors.

82. Workshop participants agreed that an impression of donor-driven partner country ownership of evaluation should be avoided. Partner countries themselves were called on to take the lead and the driver’s seat, work towards more ownership of evaluation, and develop the modalities to achieve this - for instance by putting evaluations on the agenda of government consultations and negotiations. Equally, partner countries need to strive towards participation in the evaluation process right from the outset and take an active role in all stages of the process: agreeing the terms of reference, the inception report, the evaluation report and the recommendations. Participants also acknowledged that better interdepartmental coordination is needed within aid recipient countries.

Box 4. Joint Evaluation of the Netherlands's Mixed Credit Programme in China

In 1999, the Policy and Operations Evaluation Department (IOB) of the Netherlands' Ministry of Foreign Affairs organised a review of the Mixed Credit Programme ORET/MILIEV, which is to support sustainable development in China through the generation of employment, the boosting of trade and industry and through improvement of the environment.

When the Netherlands decided to evaluate the programme for a second time, IOB approached the National Centre for Science and Technology Evaluation (NCSTE) as a potential partner in the evaluation. NCSTE already had some experience from the evaluation of the mixed credit programmes of Norway and Denmark, though not working in full partnership on these evaluations (the ToR were set before NCSTE joined the team and the evaluation report was written by the European consultants with participation of NCSTE only in providing background papers and case stories).

In September 2003, IOB and NCSTE agreed on a joint evaluation of the ORET/MILIEV programme and presented a proposal to the Chinese Ministry of Foreign Affairs which expressed positive support and welcomed the independent character of the proposed evaluation. IOB and NCSTE worked together on the evaluation design, including the ToR. Agreement on the basic principles and methods for the evaluation, based on the DAC Principles for Aid Evaluation, was reached relatively quickly. It took the two partners five months, however, to arrive at a common understanding of the key evaluation questions and the evaluation matrix.

NCSTE and IOB acted as co-chairs of the Steering Committee and assumed overall responsibility for the evaluation. A reference group, comprising experts and officials from China, the Netherlands and UNDP, was also established. Five team leaders were appointed to take charge of field work, questionnaires, desk studies and other tasks. Four of these team leaders were from NCSTE. The fifth, nominated by IOB, was responsible for desk studies and the questionnaires issued to Dutch suppliers. Presently, both parties are working on drafting the final report which is scheduled for completion by the end of 2005.

According to Professor Chen Zhaoying, there have probably been no significant cost savings in undertaking the evaluation jointly. In terms of duration, undertaking the work jointly has most likely meant that the evaluation process has taken a longer period of time to reach completion. However, the length of time taken is offset by a number of positive gains. These include:

- Compared with donor-driven evaluations, this evaluation is more relevant to the policies, needs and priorities of both parties;
- Stakeholder participation was strengthened considerably, for instance through 17 roundtables with stakeholders in all major regions of China. This helped to gather a wide range of views and perspectives, and also encouraged the use of the evaluation results by these stakeholders at a later stage;
- The involvement of NCSTE will help the wide dissemination of the evaluation findings in China, including in the National People's Congress which is represented on the Reference Group;
- NCSTE benefited from the evaluation in terms of capacity development.

As to the issue of asymmetry in the donor and recipient relationship in evaluation, Professor Chen Zhaoying concludes: "We think that asymmetry is inevitable. Donor-recipient partnerships are often based on differences in power, and in access to information. In the process of this evaluation we have found ways to reduce these differences. For example, this evaluation established a totally equal governance structure that has helped to build trust and to reduce the power asymmetry between the two sides."

83. Participants from South Africa, Tanzania and Vietnam informed the workshop about the approaches developed in their countries to strengthen ownership of monitoring, review and evaluation. In the case of Vietnam, the government has introduced an evaluation component in its Decree on ODA Management, which provides for project evaluations at three major stages of the project cycle, i.e. an initial ex ante evaluation, a mid-term evaluation, and a completion evaluation. However, it was pointed out

that the implementation of the Decree is hampered by the lack of one single national system for evaluation, by inadequate local evaluation capacities and by insufficient funds.

84. Vietnam's experience with joint evaluations is fairly limited. While there has been some joint evaluation work undertaken with Japan (on infrastructure), with Australia (on water supply) and with the World Bank and the Asian Development Bank (on energy), the government wants to do more work on joint evaluations, including developing appropriate methodologies and capacity building. The Decree on ODA Management will be revised in the near future, and it is envisaged to include joint evaluations as an important component within the revision. Furthermore, it is planned to make joint evaluations compulsory for sector wide approaches, to create a focal agency in charge of initiating joint evaluations, to include joint evaluations as a standing item on the agenda of government–donor negotiations, and to provide a budget that allows for modest funding contributions to joint evaluations. Many donors see this as a promising way forward which deserves full support by them.

85. Tanzania and South Africa have chosen to follow slightly different paths towards more ownership of monitoring and evaluation. South Africa, for example, has initiated a series of evaluations of individual donor programmes, carried out jointly between the South African National Treasury and the respective donor (see box below).

86. Tanzania has turned a number of performance assessment and review processes, such as the Tanzania Assistance Strategy (TAS) Process, the Poverty Reduction Strategy (PRS) Process and the Public Expenditure Review (PER) Process, into joint activities between the Government, the donors and, on occasion, the Independent Monitoring Group. It is hoped that the government, development partners and other stakeholders will use these mechanisms as the primary modalities for monitoring and evaluating development in Tanzania. At the same time, it is understood that this joint work also is already contributing to a significant reduction of transaction costs for the Tanzanian government, especially by reducing the number of uncoordinated review missions from abroad, facilitating mutual learning, and strengthening the commitment of all actors to implement corrective actions, increase transparency and accountability, harmonize approaches and modalities and strengthen government ownership.

86 bis. There are also concerns. The Report of the Independent Monitoring Group for the year 2005 stresses that “monitoring and evaluation is beginning to be institutionalized. However, GOT [Government of Tanzania] needs to define more clearly what is meant by evaluations and for what purpose they are made. In some cases, annual reviews are so frequent or are so delayed that learning from those evaluations and reviews is not encouraged. In addition, various systems of monitoring and evaluation have yet to be harmonized. Currently, the Ministry of Finance and the President's Office Public Service Management are working towards harmonization of their M&E systems.

87. It should also be noted, however, that the Tanzanian example is presently focusing on monitoring, review and performance assessment, with only limited evaluation components. The Tanzania experience has also raised interesting questions with regard to differing perceptions of these in-country arrangements by local donor representatives and by donor headquarters. For example, some donor headquarters have called for a full external evaluation of the Health SWAp, but local donor representatives in Tanzania believe that they already have sufficient information from the in-country monitoring and review processes. The debate has been deadlocked for some time.

Box 5. The South African Joint Evaluation Model

ODA plays an important role in supporting the reduction of poverty, inequality and vulnerability and the consolidation of democracy in South Africa - although it represents less than 2 per cent of the national budget. To allow South Africa to optimise its ODA, the country and its donors agreed that more efficient and effective evaluations are required. The International Development Cooperation (IDC) in the National Treasury therefore decided to establish a system of joint evaluations which would allow for a recipient–donor assessment of the validity, relevance and success of different programmes of support. The aims are accountability, learning and developing the basis for new programmes of support.

As a first step towards implementing the joint evaluations, a Development Cooperation Report was published in 2000 - which reviewed all development cooperation and assistance from 1994 to 1999. Based on this experience, new joint evaluation modalities were developed and, since 2002, joint evaluations have been undertaken by the South African Treasury and a range of donor agencies - namely Norway, Switzerland, Ireland, the European Union, Sweden, the Netherlands and Belgium. The intention is to make this kind of joint evaluation standard and best practice for all development partners in South Africa.

The process which was agreed for joint evaluations is as follows:

- IDC of the National Treasury takes the initiative and requests a joint evaluation;
- The terms of reference are jointly developed between IDC and the respective donor;
- A decision on the kind of joint management for the evaluation is taken;
- A decision on the kind of procurement for the evaluation team is taken. This could be international tender, in-house recruitment or local recruitment. The need for a South African component in the team is non-negotiable;
- The evaluation leads to a draft report and recommendations for discussion;
- The Treasury initiates a consultative process, on the draft report, among all relevant South African agencies;
- IDC initiates discussions with the donor on the way forward;
- The evaluation report becomes an important resource document in developing a new cooperation strategy between South Africa and the donor country.

An interesting aspect of this approach is the inclusion, within the terms of reference, of the broader view of cooperation. Thus, the evaluation is placed against the background of the comprehensive framework of bilateral relations between South Africa and the donor country, including ODA, but also economic relations, trade, investment, cultural cooperation, research, and so on. This is to ensure that the review is forward-looking and addresses opportunities for improved integration and coherence between the two countries' various cooperative relationships. The evaluation teams are also encouraged to identify new areas where there is potential for longer-term institutional relationships beyond aid. This approach to the ToR and evaluation scope is holistic and progressive.

There are, of course, issues of conflict that come up during joint evaluations and need to be addressed and resolved by partners. These can include: disagreement on the ToR and procurement modalities; attempted donor dominance over the evaluation process; critical findings - that both sides are hesitant to share upwards; preconceived ideas on both sides as to why certain difficulties are encountered in projects; lack of institutional memories; cumbersome donor management; poor performance indicators; limited time and capacity to evaluate.

However, the overall assessment of the South African joint evaluation model is very positive. Important assets include:

- Enhancement of South African ownership and management of ODA;
- Reducing the management burden on recipient agencies;
- Increasing donor understanding of South African government strategies, priorities and procedures;
- Lesson learning on both sides, leading to common best practice and innovations;
- Improved quality of programming;
- Encouraging greater donor harmonisation and alignment.

88. The workshop spent significant effort and time identifying and discussing the opportunities and benefits of joint evaluations. It also addressed the problems and challenges of joint evaluation work. The following list summarizes the discussions, with opportunities and benefits in the left column and problems and challenges in the right:

Opportunities and benefits of joint evaluations

Enhances government ownership of ODA and its management, empowers recipient countries, and strengthens government structures and systems

Encourages evaluation capacity building in governments and other organisations

Increases the transparency and accountability of all development partners

Problems and Challenges of joint evaluations

Donors still tend to strive for evaluation leadership and to control findings

There often is a divergence of interests among development partners on the setting of terms of reference and the agreement on objectives, indicators, benchmarks, etc.

The procurement modalities and the composition of evaluation teams (national vs. international) are often controversial

National M&E systems and focal points, also independent of governments, need to be set up and institutionalised to allow for the incorporation of joint evaluation work

Recipient countries must become more pro-active in taking the lead on joint evaluations, including in heavily aid-dependent countries

All partners must agree funding mechanisms that enable at least some financial contribution to joint evaluations from aid recipient countries

Recipient country policies on evaluation need to be clarified and made transparent, and guidelines on how to do joint evaluations should be developed

Horizontal coordination within recipient country governments needs to be strengthened

More training and evaluation capacity development in recipient country governments is needed

Governments should make every effort to broaden their evaluation capacity base and include universities, NGOs, the private sector, and so on

Care needs to be taken to emphasise institutional, rather than individual, evaluation capacity building

Upstream and timely communication of evaluation plans, work programmes and schedules must be improved (e.g. with annual or bi-annual evaluation planning matrix)

Hidden agendas may also come into play and politicise and frustrate evaluation efforts (actual vs. official

	motivation)
Allows for firsthand inputs and contributions by the recipient country and thus a stronger focus of the evaluation on its own interests and requirements	Donors need to consider this a chance to make evaluation more relevant, rather than a limitation
Reduces transaction costs for recipient country governments	Care needs to be taken to avoid a duplication of evaluation efforts through joint evaluations and individual donor evaluations
Allows for cost sharing and thus a reduction in costs for individual agencies	
Allows for enhanced use of local consultants and their expertise, which broadens the evidence-base, adds to cost-effectiveness, contributes to capacity building and streamlines methods and procedures	There is often a lack of clarity with the key evaluation questions. Instead, there is a set of rather general evaluation objectives which consultants are expected to turn operational National consultants may find it difficult to work independently of social and political pressures in their countries. This may be perceived externally as a lack of independence and impartiality of the evaluation process as a whole In future, local consultants should be given a stronger role and more responsibility in the substantive aspects of joint evaluation work
Enhances harmonisation of donor processes and requirements and leads to better alignment	Meeting the different reporting requirements of each of the partners can burden the evaluation process. More delegation of responsibilities would provide some relief
Increases donor understanding of national strategies, priorities and procedures	
Strengthens partnership through better mutual understanding	Needs early agreement of all partners on the methodologies, management and time framework
Contributes to the identification of opportunities and areas for fruitful cooperation and harmonisation	
Facilitates mutual learning and improvement through the identification of lessons learned and best practice	Lack of time and evaluation capacity can impair the learning process and lead to a degree of frustration
Leads to stronger commitment to conclusions and recommendations and corrective action	Commitments to results are sometimes kept deliberately vague, and the translation of recommendations into action may differ from partner to partner Critical evaluation findings are not always easily shared upwards within the partner organisations
One report and one message reduce the number of confusing and conflicting messages emanating from a variety of evaluations, and help to get agreement on common goals and priorities	It is important to clarify whether the goal of the process is to review or evaluate
Can help to break new ground more easily than individual evaluations	In this context, it is important to agree on the type of joint evaluation, for instance whether it should be a single or multiple-donor- recipient evaluation
Different actors, based on their comparative advantages, enrich the evaluation process, increase the potential for objective and independent review and make it more participatory	There can be a risk that joint evaluation partners have preconceived ideas on why certain problems exist There are often obstacles related to a lack of institutional memory within the partner organisations The involvement of as many stakeholders as possible in the design phase of an evaluation needs to be prioritised The participatory involvement of stakeholders and

beneficiaries adds to the cost of an evaluation

Allows for broadening the scope of the evaluation to cover a wider range of interests and issues

Gives greater legitimacy, credibility and weight to evaluations

It is essential to strengthen the interest of national institutions - such as parliaments, the media, the private sector and NGOs - in accountability through evaluation work,

89. This overview leaves no doubt that to date there are more challenges than solutions. However, the Nairobi workshop indicates that there is new dynamism towards more national leadership and ownership of evaluations. Partner countries are accepting the challenge of changing and re-balancing the political economy of joint evaluations in their favour. It will need perseverance to find mutually acceptable solutions to the challenges detailed above. The Nairobi workshop identified a number of recommendations, which taken together form an action programme for more partner country participation and ownership in evaluation:

- 1) The term joint evaluation should denote only those evaluations which are carried out jointly by donor(s) and recipient countries. Other forms of evaluation undertaken by more than one actor should be called multi-partner evaluations.
- 2) A joint evaluation should be of relevance to all partners joining it and should involve all partners from the outset in all key decisions (ToR, procurement of consultancy services etc.) and must not bring recipient countries on board only at a later stage.
- 3) A greater proportion of evaluations should be undertaken as joint evaluations with the full and active participants of recipient countries.
- 4) Recipient countries must take a stronger lead in initiating joint evaluations.
- 5) Better partner coordination is required within recipient countries - among government departments and other stakeholders.
- 6) In the case of joint evaluations with the participation of several developing country partners, those countries should be enabled to meet together and coordinate their inputs. Steering Committee meetings should be held in partner as well as in donor countries
- 7) Developing countries should review and build on the South African experience of initiating and leading joint evaluations with different donors.
- 8) Developing countries should also review and build on the experience of Vietnam, where the government has made monitoring and evaluation a legal requirement in the Decree on ODA Management.

CHAPTER 2: KEY STEPS IN PLANNING AND CONDUCTING JOINT EVALUATIONS: AN UPDATE

90. This chapter aims to distil lessons learnt from joint evaluations in order to help improve the planning, organisation, management, implementation and follow-up of future multi-partner evaluations. The chapter looks at four key areas: (1) Upstream planning requirements for joint evaluations; (2) Setting up and organising joint evaluations; (3) Implementing joint evaluations; and (4) Following up on joint evaluations.

91. This report has no intention of starting from scratch and reinventing the wheel, but instead aims to build on existing material; especially on the DAC publication *Effective Practices in Conducting a Joint Multi-Donor Evaluation (2000)*. Each of the four sections of this report therefore begins with a short summary of the main relevant recommendations from that earlier publication. This introductory summary is then followed by a description of recent experience and new evidence from joint evaluation work. At the end of each of the four sections, the conclusions and recommendations are summarised for quick and easy reference.

2.1 Upstream planning requirements for joint evaluations

92. The increasing number of joint evaluations should be encouragement for development agencies to pay more attention to the upstream conceptualisation and planning of joint evaluations. Two aspects are of special importance. The first is the need for more information on principles and policies governing the approaches of the different agencies to joint evaluation work. The other is the insufficient attention paid by joint evaluation partners to agreeing clear ground rules for their collaboration.

93. *Effective Practices in Conducting a Joint Multi-Donor Evaluation* does not contain a great deal of detailed advice on upstream planning requirements for joint evaluations. However, it does draw out the following relevant points:

- The focus of any given evaluation should be the key factor in deciding whether to undertake the evaluation jointly or singly. Evaluations focusing on jointly funded programmes, on broad and/or sectoral development goals, or on the evaluation of a multilateral or regional development agency can often best be undertaken jointly.
- If the purpose of an evaluation is primarily lesson learning (and not accountability), a joint evaluation may be appropriate. 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.
- While joint evaluations tend to increase the management burden, they do provide small donors with the option of participating in comprehensive evaluations that might otherwise be beyond their capacity.
- Apart from focus and purpose, the other factors which help bring donors together include: similar development philosophies, organisational cultures, evaluation procedures and techniques, regional affiliations and proximity.

- With regard to ground rules, some of the problems that need to be addressed are also mentioned. Prior agreements on processes for resolving potential conflicts are considered useful, as are special precautions to preserve the independence and objectivity of an evaluation. Also, careful attention should be given to the evaluation process - which means sharing decisions concerning the evaluation's planning and management, its scope of work, the team composition, the methodology, the reporting format and findings, etc.
- Finally, the crucial question of partner country involvement needs to be taken up early in the process.

Making principles and policies transparent

94. At present, little is known of the orientations, policies and guidelines on and towards joint evaluations of the development co-operation agencies. Only a few evaluation manuals (or other publications) contain references to joint evaluation work, and as a rule, these are not very detailed references. The donor community (and other partners) have no standard way of researching the level of interest in joint evaluations of another individual agency; and thereby of knowing which donor agency should be a potential candidate to contact when looking for partners for a joint evaluation.

95. Evidence collected for this study demonstrates that many actors would welcome more information on the orientations, principles and guidelines of the development agencies. These should set out the general policy with regard to joint evaluations (both at headquarters and at country level), the degree of the agency's commitment to this form of evaluation work, the opportunities that the agency can offer in this context as well as the limitations that it has to keep in mind. Furthermore, they should explain: the preferred areas of focus and scope of joint evaluations which the agency would like to get involved in; the minimum criteria and standards required by the agency if it is to consider joining a multi-partner evaluation; the process of decision making;⁹ and the availability of resources for this kind of evaluation work, including possible financial contributions.

96. Information of this kind would help to bring more transparency to the arena of joint evaluations. Such transparency is urgently required, both for internal agency purposes and for the outside world. It would make an important contribution to reducing the risk of creating (or perpetuating) the impression that agencies deal with joint evaluations in an ad hoc manner, and that personal preferences or influences might even come into play.

97. Once this required information is available, it should be made public in a systematic manner. One way of doing this is to put it on the evaluation website of each agency, together with other information on evaluation policies and practices. Another possibility would be the establishment of a kind of clearing house for all matters related to joint evaluations and to allow a one-stop approach for those who are interested in information and/or communication on the subject.¹⁰

9. The decision to join a multi-partner evaluation is not always left to the discretion of the head of the evaluation unit. In some cases, joint evaluations which were not initially included in the approved evaluation programme of the agency need clearance or approval, on a case-by case basis, by the senior management. In at least one DAC member country, approval for joint multi-partner evaluations has to be sought from the minister. This is partly because of the risk that the multi-partner evaluation could impinge on national sovereignty - because a joint process is not under the full control of any one country.

10. It has been suggested, for example, that such a clearing house could be established as a project within the Development Co-operation Directorate of the OECD, financed via project funding from outside sources. This idea, however, would need further elaboration and discussion in the DAC Evaluation Network

Clarifying purpose, objectives, focus and scope

98. Purpose, objectives, focus and scope are crucial building blocks for the foundation of a joint evaluation. A framework of common understanding must be established as the very first step for a planned evaluation.

99. One of the near-universal lessons learned from recent experience with joint evaluations is that it is imperative to allow sufficient time at the beginning of the process to develop and agree this framework of common understanding about the proposed evaluation's purpose, objectives, focus and scope. When this is not done with the necessary time and patience, as is the case in a number of the evaluations analysed, there is a strong chance that the evaluation process will run into difficulties later on. Evaluation partners who have avoided discussing these critical issues at the beginning of the process (perhaps in the expectation of saving time and escaping conflict) will regret this decision later on. Any time and effort saved in bypassing these discussions - and moving more rapidly towards the implementation phase - will very likely need to be invested at a later stage and under more stressful conditions.

100. Despite the fact that evaluators have learnt this lesson over and over again, many joint evaluations continue to be rushed through at the outset. There are many reasons for this, including:

- The main sponsor of the proposed evaluation is already running behind schedule when initiating the evaluation and starting to look for partners; budgetary requirements such as committing or disbursing funds can also add further time pressure;
- Adequate time and staff resources have not been put aside to prepare the first proposal, although this proposal is usually of pivotal importance in launching a joint evaluation. This document should very thoroughly anticipate objections and problems, and make every effort to understand the motives and stakes of the potential partners. Ideally, the first proposal should be informally discussed and agreed; so that the draft can reflect an emerging consensus before it is officially presented;
- The exact opposite of the above problem can also occur; a proposal for a joint evaluation can appear so well thought-out and designed that partners do not feel encouraged to present their own points for discussion and consideration. Not surprisingly, this lack of real consensus can backfire later on when the evaluation process and implementation becomes complicated and contentious;
- The initial discussions on purpose, objectives, focus and scope of a joint evaluation will often disclose open or hidden conflicts and/or conflicting agendas among the group of partners. This situation is not always understood and/or used as an opportunity to address some of these underlying tensions; and to resolve them early on. If these tensions are not addressed at this stage, it is most likely that they will reappear later; and impact negatively on the evaluation process and implementation;¹¹
- Finally, the purpose, objectives, focus and scope of a joint evaluation may be only superficially defined; and the evaluation partners decide to leave the further clarification to the consultants. This decision may cause the evaluation to get trapped in a vicious circle: the basis for the selection of consultants is weak because the profile of expertise that is needed has not been fully

11. It can be observed occasionally that in the case of unresolved conflicts among partners in a joint evaluation, it is tacitly agreed to "leave it to the professional competence of the consultants" to deal with the problems and find a solution. This usually does not work, and the consultants often complain that they are blamed for something completely outside of their control.

explored; this entails the risk of selecting a suboptimal or wrong consultant(s) who, in turn, will not adequately or appropriately define the purpose, objectives, focus and scope. There is a strong probability that long and difficult discussions on these unresolved issues will emerge later in the evaluation process.

Establishing a set of ground rules

101. All joint evaluations require ground rules. These have to be agreed early on and openly among all the partners involved in the evaluation. Among other things, ground rules should determine the standards governing the evaluation, for instance with regard to the degree of application of the DAC Principles for evaluation; they should fix the rights and obligations of partners, including the conditions under which one partner may opt out during the process; they should also help to balance the different roles that partners will assume, so that a more level playing field can be maintained throughout the evaluation process; they should define the degree of commitment that partners will show to the findings and recommendations of the joint evaluation; they should provide agreement on a formula for burden sharing; they should decide on the basic composition of the governing bodies of the evaluation and whether they should be comprised solely of evaluators, or of a mix of people, including evaluators, policy, sector and technical staff; and, finally, they should take a decision on whether, at what point, and in what role partner country representatives should be included.

102. The need for ground rules is not only one of the lessons learned from recent experience, but is also one of the issues raised over and over again during the research for this report. This is because there has been a tendency among partners in joint evaluations not to address difficult issues for as long as they remain dormant and do not demand action. This can be a pragmatic way of working together; as long as it is understood by all partners that this kind of approach can only work if one single, important ground rule is accepted and observed: no one in the group should try to gain advantages over and at the expense of others. This, of course, can be a difficult rule to deliver on in reality. Therefore, in nearly all cases, it is pragmatic to spend more time agreeing the ground rules during the initial phase of a joint evaluation.

103. Ground rules can help to define the basic relationships between the group of partners sponsoring the evaluation and between them and the team of consultants. If the ground rules stipulate the full acceptance of all or of some of the DAC Principles for aid evaluation as the standards to which the consultants must adhere, then it is unlikely this will be challenged later on. On the other hand, if there are no clear standards laid out in the ground rules, the evaluation runs the risk of prolonged debates at a later stage on the issue of standards. Such debates can become acrimonious; especially when they deal with the independence of the consultants and their judgement, with the way critical findings should be handled in the report, and whether particular recommendations should be included in the report.

104. There is some demand from donors for a ground rule specifying the conditions under which one partner can opt out of the evaluation. It should be emphasised, however, that only a few cases of withdrawal have occurred so far. One such example did occur in the Rwanda evaluation, *The International Response to Conflict and Genocide: Lessons from the Rwanda Experience*, when one partner withdrew due to a high-level political decision. However, this is a rather special and probably unique case of a politically motivated withdrawal.

105. A more recent case, of a different nature has a lot to do with the fact that no full set of ground rules for the joint evaluation had been established; neither for the approach to methodological issues nor for opting out. The withdrawal of one of the partners was due to conflicting views on methodological issues. As a consequence of the withdrawal, the evaluation process drifted into rough waters and it took extra efforts by the other partners to overcome the uneasiness left behind. In all likelihood, this could have been managed a great deal more smoothly and efficiently had the appropriate ground rules been in place.

106. Ground rules for burden sharing are among those that should be agreed and established early in the evaluation process. It is essential to agree on the overall cost and finances for a particular evaluation; on the procedures for securing the required financing (cash or in-kind contributions? Equal sharing of the costs among all partners, or individual pledging?); on the kind of financial administration to be set up (pooling of resources or individual payments of expenditure? One agency to administer the funds on behalf of all? According to what rules? With what kind of control and audit in place?); and on a formula for sharing unforeseen expenditures.¹²

107. The issue of establishing a level playing field among all partners is not an easy one to accommodate within the set of ground rules. It is usually a relevant issue in evaluations where a large number of different actors work together. This is especially important when the size of the agencies, their backgrounds, evaluation philosophies and their length and depth of experience in evaluation differ significantly, and where the sharing of burdens is uneven, and, perhaps most importantly, where partner country involvement is sought, implemented and financed from the evaluation budget or from other third party sources.

108. A level playing field does not mean establishing total equality. However, it is essential to provide safeguards to ensure that voices of the weaker partners will be heard and respected by the stronger players. This objective can be achieved in many ways: through procedural precautions; through the assignment of specific key roles (e.g. chairpersons, rapporteurs, etc.); through a division of labour which gives additional weight to the not-so-strong; or through a clear and unequivocal de-link of influence from the size of financial contributions. The most important element of all, however, is to articulate the issue, put it squarely on the table early on in the process, and discuss it, in a frank and open manner, with a view to sensitizing everyone to the concerns that may exist.¹³

109. With regard to the composition of governing bodies, such as steering committees and management groups, recent evidence does not provide a clear indication as to the preference of the majority of aid agencies, consultants and others on how such bodies should be constituted. Evaluators tend to believe that they are more likely than others to be able to guarantee the full independence and impartiality of an evaluation process and to make sure that the evaluation is carried out according to state of the art knowledge. On the other hand, participants in evaluations with a mixed steering committee have emphasised the useful role of sector experts, operational staff and policy people - and their merit in bringing a strong sense of reality to the evaluation. There may also be drawbacks. These are discussed further down in paragraphs 118 – 120.

110. Last, but not least, there are two final issues which are absolutely critical in connection with ground rules for joint evaluations. The first one is the question of how committed the partners will be to the findings, conclusions and recommendations. Recent evidence in this regard is quite mixed. A good example is the evaluation of basic education, *Local Solutions to Global Challenges: towards Effective Partnership in Basic Education*. Some of the agencies that supported this evaluation with considerable efforts in time, staff resources and money, have gone to great lengths to disseminate the findings in their own institutions as well as externally, and to make sure that the recommendations are taken seriously by

12. Incidentally, the case of opting out halfway through which is mentioned above, also has a strong connotation of burden sharing, because the partner who withdrew decided not to honour the obligations under the burden sharing agreement worked out earlier for that particular evaluation. So, the remaining partners had to come up with additional financing to fill the gap.

13. The CDF evaluation is a good example of how the problem of a level playing field can be addressed, once the feeling of uneasiness among some of the partners is articulated openly and dealt with in a constructive manner.

their policy teams.¹⁴ Others, with equally strong contributions to the evaluation, appear to have largely ignored the outcome - or even decided not to present the final report to their senior management. This illustrates that joint evaluations without serious and early commitment to the results and recommendations can easily be perceived as useless and a waste of scarce resources.

111. Equally critical is the issue of partner country involvement. This issue should rank very high now on the agenda of any emerging joint evaluation, and there is some evidence that it is being given more proactive attention in recent years. However, the timing for addressing this issue is, in many cases, far from optimal. The issue needs to be addressed at the very beginning of the process, but at present the dimension of partner country involvement usually enters the discussion only later when other fundamental decisions have already been taken.

112. There are, of course, practical obstacles to finding the right partners from developing countries early enough to involve them fully in the development and design of an evaluation. Nevertheless, donor agencies must make stronger efforts to find ways and means to resolve this problem in order to remain credible in their drive for partner country ownership and leadership. It will be in that moment, that donor agencies will finally demonstrate that they have given up their over-dominant role in evaluations and in joint evaluations.

Main conclusions and recommendations

- At present, little is known of the orientations, policies and guidelines on and towards joint evaluations of the development co-operation agencies. Evidence demonstrates that many stakeholders would welcome more information of this kind. It should set out the general policy of the agency with regard to joint evaluations, clarify the degree of commitment to this form of evaluation, explain the preferred areas of focus and scope of interest to the agency, the minimum criteria and standards required, the process of decision making, and the availability of resources.
- The above information should be made public on the evaluation website of each agency. Another possibility is that a one-stop address for all matters related to joint evaluations should be established.
- It is absolutely imperative to allow sufficient time at the beginning of the process in order to agree a common framework of understanding for the evaluation's purpose, objectives, scope and focus. Partners who avoid these critical discussions at the beginning of the process - in the hope of saving time and escaping conflict - will find later that this was a false economy.
- The initial proposal for a joint evaluation should be fully thought through to anticipate objections and making every effort to understand the motives and stakes of potential partners. Ideally, the first proposal should be informally discussed and agreed; so that the draft can reflect an emerging consensus before it is officially presented.
- Ground rules should determine the standards governing the evaluation, for instance with regard to the degree of applying the DAC principles for evaluation; they should: fix the rights and obligations of partners, including the right and the conditions under which one partner may decide to opt out during the process; help to balance the different roles that partners will assume,

14. CIDA, for instance, submitted the final report of this evaluation, together with a management response, to its Audit and Evaluation Committee that meets under the chairmanship of a Senior Vice-President of the organisation. The committee decided that a plan of action for the implementation of the recommendations emanating from this evaluation was to be developed for CIDA.

so that a level playing field can be maintained throughout the evaluation process; define the degree of commitment that partners have to the findings and recommendations; provide agreement on a formula for burden sharing; decide on the basic composition of the governing bodies for the evaluation; and take a decision on partner country representation.

- A level playing field should not be confused with total equality. However, it is essential to provide a sufficient number of safeguards for the weaker in the group to ensure their voices are heard and respected by the stronger players. This can be achieved in many ways. The most important element of all, however, is to articulate the issue, with a view to sensitizing everyone to the concerns that exist.
- The issue of partner country involvement is critical, and should rank high on the agenda of any emerging joint evaluation. It needs to be addressed at the very beginning of the process. In too many cases, partner country involvement enters the discussion only when several fundamental decisions have been taken, such as on the terms of reference; the appointment of a consultant; and even the selection of partner countries where case studies should be carried out.

2.2 Setting up and organising joint evaluation work

113. *Effective Practices in Conducting a Joint Multi-Donor Evaluation* contains a large number of practices and recommendations in this category. Some of the most pertinent are the following:

- In cases where only a few donors are involved, the management structure is usually quite simple. Participants may decide to share equally in all management decisions. Alternatively, they may decide to delegate management to just one agency.
- In cases of a relatively large group, day-to-day management is typically delegated to a few members. Often this small management group is self-appointed, comprised of representatives of those donors with a special interest in the evaluation and willing to volunteer substantial efforts.
- A key lesson from experience is this: giving adequate time and attention to the planning of joint evaluations can be critical to their success.
- 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.
- A major challenge in preparing a scope of work 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.
- Prepare a timetable for the key phases and establish due dates for submitting reports. 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.
- State when various draft interim and final reports are due and who is responsible. Provide adequate details on their format, content and length and on language requirements.
- Joint evaluation teams are often comprised of consultants who are nationals from the participating donor countries. Experience indicates that it is often useful to include donor agency

officials on the team as well as consultants. Teams may also have members who are representatives from the implementing agencies or recipient country organisations.

- Selecting a team leader requires special attention. The team leader must not only have evaluation and subject-matter expertise, but also strong management, communication and writing skills.
- When selecting developing countries for fieldwork case studies, it is a useful practice to devise selection criteria to help ensure that key aspects are covered.

114. Development cooperation is undertaken in the world of governments, aid agencies, and international organisations. This is a milieu dominated by complex bureaucracies, and this fact has some bearing on all the work of development cooperation. Bureaucratic culture is characterised by specific forms of action, conduct and behaviour, such as the inclination to establish rules and regulations, to create power hierarchies, to structure work along organisational considerations, and to attempt to avoid surprises by imposing tight control mechanisms. Bureaucratic culture is a determining factor in the way in which joint evaluations are set up, organised and run.¹⁵ This can be a complicating factor, because several partners in one evaluation means different bureaucratic cultures and regulations have to be accommodated and merged into one mutually acceptable management structure.

Creating a governance and management structure

115. The governance and management structure for a joint evaluation depends to a large extent on three factors: (I) the size of the group of partners; (II) the character of the evaluation (desk work; meta evaluation; field work; case studies; etc.); and (III) the fundamental approach to the exercise (centralised or decentralised; delegation of work; silent partnerships; parallel work; and so on). If there are only two or three partners working together, there is not much need for complex governance and management arrangements and a light, probably flat, structure will most likely be sufficient.

116. However, the more complex and multi-partner joint evaluations need a more elaborate governance and management structure if a fair degree of efficiency is to be achieved. In practice, the trend is to adopt hierarchical elements into the structure. In concrete terms, this “hierarchy” normally entails (a) a steering committee in which all partners are represented on equal terms. The steering committee would oversee the whole evaluation process. In addition, it would be responsible for certain key decisions, such as the approval of the terms of reference, the selection of the consultants, and the approval and release of the evaluation products (inception report, case studies, synthesis report); and (b) a smaller management or core group that runs the day-to-day business of the evaluation and is the primary point of contact for the consultants.

117. In this kind of arrangement, the lion’s share of the work rests with the management group. Therefore, agencies need to be sure that they can afford and maintain, over the whole evaluation period, the required outlay in staff time, backstopping and travel necessary to fulfil their role in the management group. Likewise, the other partners in a joint evaluation should not hesitate to assure themselves that candidates for the management group really can afford the commitment and can guarantee to maintain it over the full period of time.

15. In this context, it is interesting to take a look at the way evaluation work is approached, organised and implemented among NGOs, for instance. As a rule, NGO evaluations are much more action oriented, participatory, open-ended, etc. This, of course, can create problems with regard to widely accepted evaluation standards, such as impartiality, independence and methodological rigour.

118. Although a steering committee and a management group represent a straightforward overall structure, there are variations additional aspects that complicate the picture. One is the composition of the groups; should steering committees and management groups be made up of evaluation specialists only? Or is there a case to be made for the inclusion of sector, policy or operational staff as well? The answer is relatively easy for evaluations which are initiated and implemented at the partner country level: they will most likely be run by sector and operational people from the local representations of aid agencies and embassies, as it is quite rare to find evaluation specialists posted to the field. However, this can lead to conflicts of interest between local representations and evaluation services in headquarters.

119. In other categories of joint evaluations, decisions on the composition of the groups have to be taken with prudence. Inclusion of policy and sector staff can work well in the steering committees of evaluations that deal with global policy and impact and effectiveness analysis as well as with thematic and sector issues. Non-evaluation staff bring useful expertise and a strong notion of realism to the table, and often contribute to making sure the evaluation results are relevant for practitioners. On the other hand, the evaluation specialists in the group will be responsible to ensure that this realism and pragmatism are not pushed too far and are not used as an excuse for compromising the rigour of approach and methodology, or for the restricted presentation of those findings that are less palatable.

120. However, the involvement of policy and operational staff in the management structure of joint evaluations dealing with the analysis and assessment of specific multilateral institutions, of country strategies and of individual programmes or projects is less advisable; as it can be perceived as an attempt to influence the outcome of the evaluation. Even in evaluator-only steering groups, it can be wise to pay close attention to holding the risks of such perceptions at bay.

121. As far as the composition of the management groups is concerned, evidence shows that these groups are often made up of evaluation managers or specialists only. That makes sense, because it is the management group which stays in close liaison with the consultants on all matters pertaining to the evaluation process and methodology, and which has to assess the professional performance of the evaluation team and the quality of the work presented.

122. In parenthesis, it should also be noted that language can become an issue in forming a steering committee and management group. Although English is the dominant language in joint evaluation work so far, that could change in future - depending on the geographical focus of joint evaluations and on the degree of genuine involvement of the partner countries.

123. Several recent large multi-partner evaluations demonstrate a serious and intensified effort by the donor community to involve partner countries at the earliest stage in the work of steering committees and management groups. However, a key structural deficit in this regard has not yet been resolved: in most cases, the invitation to partner countries is extended only after some initial decisions have been taken. These decisions usually include: the decision to do a specific evaluation; the definition of the ground rules and of the governance and management structure; the terms of reference or, at least, a fairly advanced draft of them; the selection of consultants or their pre-selection, for instance in the case of the expression of interest procedure; the evaluation design - including the methodology; and in many instances, the selection of country case studies.

124. No patent formula has yet been found to allow partner countries to have their voices heard more upstream in the process of developing a joint evaluation, and to become fully and unrestrictedly involved in all the early decisions. More innovative thinking is clearly required from the donor community in this regard, including on the potential usefulness in this context of national or regional evaluation associations or similar.

125. Two more issues in connection with the governance and management structure of joint evaluations remain to be addressed. The first one is quality assurance. In a number of recent joint evaluations, the question of how to deal with quality assurance in the context of the overall governance structure was resolved in interesting innovative ways. In the basic education evaluation, a senior and experienced consultant (with a solid background both in the subject matter and in evaluation work) was engaged to accompany the evaluation process throughout, starting with sketching out the first ideas for the evaluation until the very end. The consultant acted as the senior advisor to both the management group and the steering committee, helping them to bring the process to a satisfactory conclusion. In his own words, his role was that of “catalyst, facilitator and communicator, also of honest broker at times, helping the members of the steering and management groups to overcome their inclination to avoid controversial issues and hard decisions”. He emphasises the importance of his not providing any ground on which his neutrality could have been questioned.

126. Another innovative way of dealing with quality assurance issues was chosen by the steering committee of the evaluation of the Enabling Development Policy of WFP. When the quality assurance work of the consortium of consultants began to show flaws, the steering group engaged two experienced senior quality advisers. Their main function was to carefully examine and check all early drafts of evaluation products and to feedback their reactions and comments to the consultants before any of the drafts were submitted officially to the steering committee. This helped to streamline the process in the steering group of reviewing drafts. It also gave the consultants some psychological incentive to accept the relatively discreet criticism of the senior quality advisers. And, finally, it helped to take out some of the potential animosity that can easily emerge in the debates of steering committees faced with drafts that they consider to be inadequate.

127. The second issue is the need for full transparency in the governance structure, especially in cases where the steering committee has delegated authority and decision-making power to a management group. The simplest and most efficient way of establishing transparency is to make provision for the quick and timely circulation of full records of all meetings. Evidence shows that it is a wise investment of effort and time – and, perhaps, even of a person specifically hired for the purpose- to produce records of meetings which reflect the individual contributions and opinions of members, summarize definitions agreed upon, and list all the decisions taken. Preferably, these records should be written and circulated immediately after the event so that members can comment on them while the memory of the meeting is still fresh.

The terms of reference

128. Once the governance and management structure has been set up, the terms of reference have to be developed to define the scope and subject as well as the objectives of the evaluation; the methodology to be followed; the overall purpose; and how it will be achieved. This describes, in a nutshell, one of the most difficult, often contentious and occasionally highly controversial processes in any joint evaluation. Why? Joint evaluations bring together a range of actors who, although they have met to pursue a common goal, have different political and cultural backgrounds, different interests, different aspirations and also different evaluation cultures. It could even be said that at the beginning of a joint evaluation process there are often as many, or even more, differences than common ground.

129. There is no universal or ideal way to solve this problem. Each joint evaluation is different, but there are some issues that do always emerge. First, early agreement should be sought on what the participants hope to see as a terms of reference (ToR), and how detailed these should be.¹⁶ There are (at

16. It should be recalled here that the definition of terms of reference presented in the *DAC Glossary of Key Terms in Evaluation and Results Based Management* is relatively general. It reads: “Written document presenting the purpose and scope of the evaluation, the methods to be used, the standard, against which performance is to be assessed or analyses to be conducted, the resources and time allocated, and reporting

least) two different schools of thought on ToR. One school requires the ToR to cover the full evaluation process and to describe in great detail the purpose, objectives, subject matter, scope, methodologies, etc; and to include a full evaluation matrix with objectives, indicators, benchmarks etc. The other philosophy is the opposite: counselling that the initial ToR should be kept short and simple; describing the purpose, the setting, the subject and the framework; and tasking the consultants to develop this further during the inception phase.

130. These two views are hard to reconcile, although efforts to this effect should be made in the steering committee. If this does not work, it is strongly advised not to opt for a poor compromise between these two methodologies in the hope that the consultants will be able to function with this later on. The problem will surface again and again in meetings of the management and steering groups. Therefore, if you can't win the debate, it is better to give in and to agree for the other option than to negotiate an unworkable compromise. Otherwise, the whole evaluation process may be badly hampered.

131. Another issue that comes up in almost every joint evaluation is the question of methodology; particularly that of quantitative versus qualitative methods. Again, there are no easy solutions - although the budgetary constraints that exist in most cases may well work against a predominantly quantitative approach with expensive data collection. The issue of recommendations is another crucial and problematic area in shaping the terms of reference. Some donors are opposed to allowing consultants to form the recommendations. Others want exactly the opposite, and insist that consultants should submit recommendations. When country case studies are part of a joint evaluation, experience shows that the aid practitioners on the spot (i.e. those managing the project or programme being evaluated) are often very keen to have recommendations - to help them solve the problems of their work. Indeed, country case studies conceived as one element within a global evaluation, will only find the necessary logistical and other in-country support if the terms of reference clarify also the usefulness of these studies for the work at the country level. Recent evaluations that were ambiguous on this point have begun to meet with growing uneasiness and signs of reluctance to provide more than a minimum of logistical support.

132. A final point about ToR (that has been given more attention recently) is the inclusion of possible approaches for follow-up and the dissemination of results. To date, the prevailing attitude has been to wait and see what the final product is like, and only then to start discussions about dissemination and follow-up. This approach entails the risk that the consultants may no longer be available to assist when it comes to dissemination. There is also a danger that the occasions which would be particularly appropriate to present the evaluation results are missed, because they were not taken into account from the beginning of the evaluation process.

133. Follow-up is another issue which merits more attention in the terms of reference. Occasions for follow-up could (again) be events that will take place anyway. These events need to be worked into the evaluation timetable early on in the process. But there are also possibilities to organise special follow-up activities, including the option of presenting the evaluation results to a joint workshop of policy staff from the agencies involved, information sessions for the staff of individual agencies, or, very importantly to learn from experience, an evaluation workshop on the evaluation process (or "post mortem").

Costing and budgeting for joint evaluations

134. Agreeing the budget for a joint evaluation is another crucial step in preparing the ground for a successful exercise. Although funds are often scarce and limited, the costing of a joint evaluation must be realistic and should encompass all elements of the evaluation cycle that require financing as well as

requirements." It then continues to point out: "Two other expressions sometimes used with the same meaning are 'scope of work' and 'evaluation mandate'."

allowing for unforeseen developments and expenditures. A shoestring budget is unlikely to yield all the desired results.

135. Costing and budgeting for a joint evaluation is a difficult balancing act. First, there are the agencies that have agreed in principle to finance the multi-partner evaluation and who, most of the time, will try to keep it cheap. They either have financial limitations to observe due to tight aid budgets, or will have decided to put a ceiling on their share of the cost or to contribute a fixed sum only. Sometimes, they may not want to appear too generous with the funding of joint evaluations, because this may lead to criticism from within their own agency. If early cost estimates become too high, it is quite common to cut out certain items - for instance on dissemination and follow-up - and to agree (often euphemistically) to return to them later on in the process.

136. There is also the role of the consultants. Many consultants state that they tend to minimise expenditure items and to play down cost in the bidding process; in order to be able to submit competitive budget estimates. Should they win the contract, there will be opportunities later to make the budget more realistic and to adjust it upwards. This tactic has as many flaws as the opposite approach, followed less frequently, which inflates the initial cost estimates with the intention of being able to lower the costs during the negotiation process.

137. The practices described above are in striking contrast with what all parties say they would like to see: realistic costing, full budgeting of all expenditures within the evaluation cycle, and budgetary provisions for meeting additional costs that may be incurred during the evaluation due to unexpected circumstances. Experience with recent joint evaluations shows, however, that there is much room for progress towards more expedient handling of financial questions:

- It is preferable to start the budget process with the question of cost, and not with the financing aspect. There will be opportunities thereafter to try and match expenditure and income in a balanced budget.
- The preliminary costing of a joint evaluation should be based on experiences with comparable exercises, plus an additional safety margin. However, to the extent possible, no final financial commitments should be made on the basis of these preliminary figures, as they may still change substantially in the budget process which follows. Early commitments by some donors entail the risk for the others of becoming stuck with all of the additional unforeseen costs.
- A budget should cover the full cycle of the evaluation and should include items such as quality assurance; translation, editing, printing and dissemination of the reports; follow up through workshops, seminars and other events; and contingencies. Although contingency items in a budget are never very popular, they help to avoid cumbersome procedures at a later stage, if the original amount budgeted for an evaluation needs to be augmented.¹⁷
- More time and effort should be spent on the scrutiny and critical assessment of the financial aspects of the bids submitted by consultants who are competing for a contract. It requires a great deal of expertise and experience to compare different bids and to discover the flaws in a financial proposal. Therefore, steering committees or management groups should look into the possibility of soliciting professional advice for the evaluation of the bids. Ted Freeman of the consultant firm Goss Gilroy Inc. tells his consultant colleagues that if the financial envelope offered for an

17. In some administrative systems, contingencies can be budgeted for without too much ado, because they will only be released once a steering or management committee, for instance, has been convinced of the necessity for the additional expenditure and approved it.

evaluation is not adequate, they should not accept to undertake the evaluation. This sound advice also applies in the inverse situation; if a consultant offers to do a job and it appears to be below cost, the steering committee should not accept the bid. Consulting is a business which needs to make a profit, and if there is no profit either the quality of the product will be sub-optimal, or the consultant will go bankrupt or, most likely, the consultant will demand more money halfway through the exercise, when his negotiating position is strong.

- The true size of the final budget will only be known after the contract has been awarded and the contract negotiations with the bidder have come to a successful end. That should be the moment at which all partners in an evaluation firmly pledge their financing contributions.¹⁸

Bidding and contracting for joint evaluations

138. The bidding and contract procedures for evaluations are varied, complex and often tedious. Joint evaluations are certainly no exception to the rule. On the contrary, the fact that different agencies work together and that each of them brings its individual rules and regulations to the table can make matters even more complicated. Again, there are no patent solutions to the problem, but there are a number of models, developed and tested in recent years, that could be of help in working toward pragmatic and manageable solutions.

139. Evidence from recent experience demonstrates that there is a strong, almost universal, trend to avoid direct contracting for joint evaluations and to opt instead for full competitive bidding procedures. There is an argument that direct contracting has some advantages over competitive bidding; such as no delays in commissioning the work; less resources needed to sift the piles of bidding documents; the carefully targeted recruitment of expertise and therefore a reduced risk of inadequate performance; and greater speed in completing the contracting processes. Nevertheless, competitive bidding prevails, because it is the *modus operandi* best suited for the highest possible degree of transparency, for value for money, and for competition on substance.

140. Forms of competitive bidding differ widely. Different evaluations have used the World Bank, UN, and European Union rules. There are also examples of using a pre-qualification exercise to identify qualified consultants who would then be invited to submit a full bid. In other cases, a list of qualified consultants is put together by the agencies working together on a joint evaluation and is then used to invite bids.

141. Whatever kind of bidding procedure is selected, in almost all joint evaluations the time needed for the bidding, selection and negotiation process has been grossly under-estimated. This has resulted in a number of significant delays in the evaluation schedule. It is therefore important to make absolutely sure that enough time has been scheduled for the selection of the consultants. As an approximate rule of thumb, a minimum of three to four months should be allowed for the selection process (after the date of publication of the invitation to bid).

142. There are several reasons why consultants show a relatively strong inclination to form consortia of firms to bid for joint evaluations. One of the main reasons is that the complexity and size of many joint evaluations can easily overstretch the capacity of one consultant alone. Consequently, an interested consultant will look for partners that could contribute the experience, knowledge and expertise that he himself does not possess.

18. This approach may not work in some cases, due to the requirements of some donors to secure the full funding of an evaluation before the bidding process can be initiated.

143. But, quite often, other considerations also come into play for tactical reasons. The country of origin or residence of a consulting firm can be an important reason to seek the inclusion of this company in the consortium. It is often felt, though this is not really substantiated, that the chances of a consortium winning a contract are greater if it is composed with consultants from many of the countries that are represented in the group of agencies supporting the evaluation.

144. As a result, consortia can become complex and relatively heavy in structure. The bigger a consortium is, the greater the need for it to dedicate significant resources and a lot of energy to organising itself and developing smooth working relationships amongst the group. In quite a few cases of consortia, the potential for synergies was overshadowed by quarrels and arguments about the shares of the cake for each consultant and fighting about the pecking order in the initial phase of working together.

145. Therefore, experienced consultants advise members of steering committees and management groups to allocate all the time possible to the selection process and to carefully analyse and scrutinize the various consortia. Such a prudent approach can go a long way in establishing the credentials of a consortium. A thorough scrutiny of a consortium should comprise (at least) the following questions:

- Is the composition of the consortium primarily based on complementary experience, expertise and knowledge? Or is there a tactical notion involved that partners are also (or even predominantly) selected because of their country of origin?
- Does the number of members in the consortium appear reasonable (not more than three as a rule) or is it too many?
- Have members of the consortium already worked together before, in a comparable setting and on similarly complex tasks? What were the experiences with this cooperation?
- Have members of the consortium discussed and agreed on the approach to, the methodology, and the division of labour for the evaluation, before submitting their bid? Has that agreement been set down in some binding form?
- Are tasks, duties, responsibilities and income divided in a transparent and fair manner among the members of the consortium?¹⁹ Has this been set down in an agreement prior to submitting the bid?
- Has the consortium agreed on a system of quality assurance to guarantee overall adherence to the quality standards stipulated in the bidding documents?
- Are the administration and the financial management of the evaluation scattered across the consortium or are they concentrated in the hands of one partner, as should be the case?

146. Management groups and steering committees are strongly advised to check and make sure that, prior to submitting a bid, the consortium has discussed the range of key issues involved in working together, and has established a firm understanding of the roles of each partner and on the rules governing their collaboration in the exercise. This advice is necessary because almost all examples of joint evaluations implemented by consortia in recent times show that these questions are regularly not

19. One consultant interviewed stressed, that as a consequence of her previous experiences in a consortium the members of which were fighting bitterly for their shares of the cake, she would only consider joining a consortium again in future if it was clear from the beginning that both the work and the income would be shared among members of the consortium in equal parts.

sufficiently addressed early on in the process, neither among the members of the consortium, nor in the selection process. As a result, unresolved conflicts will emerge during the work of the consortium. In most cases, they will spill over into the work of management and steering groups and create unnecessary discussions and problems.

Coping with the legal issues involved in joint evaluations

147. The contract agreed with the consultants is a vital document. There is no standard or model contract that would contribute to harmonisation in this area. One experienced consultant pointed out that he had worked on comparable assignments under contracts varying in length from 5 to 125 pages. The contracts normally reflect the legal system, the requirements and the established practice of the agency which is taking the lead on behalf of the group.

148. Another contractual issue that needs to be given full attention is the question of lump sum agreements versus negotiated contracts (with or without a strong element of reimbursable expenses). There is also the difficult question of cancellation clauses. Should a clause be included in the contract that allows the early termination of a contract in the case of poor performance? One method that has been used to deal with this difficult issue is the inclusion of an option clause in the contract. An option clause requires the customer to explicitly request the continuation of the work by the contractor at certain stages of the process. Should this request not be made, the contract will expire. While this is a relatively easy form for the customer to terminate the contract, it clearly puts the pressure and onus on the contractor.²⁰

149. One important aspect of the legal implications of joint evaluations is the contractual needs that emerge amongst the partners. In the case of pooled funding, most notably, different kinds of contracts and agreements are required to establish legally binding relationships between the different partners. Agencies providing pooled funding will often need a legal instrument before they are able to transfer their contributions to the country managing the pool of funds. A specific issue that comes into play in this context is the legal ban, in some countries, for the government to support another government financially, even if this support is earmarked for activities jointly agreed and implemented.

150. Another legal difficulty can be the stipulation in the co-funding agreement that the pool country has to submit progress reports to show that the funds are being put to proper use. The hitch is that the one partner is thereby accepting full responsibility for the disbursement of funds over which it does not have full control because the key decisions are taken by a steering committee or management group. Difficulties can also arise when one or more of the co-funding countries want to reserve the right to audit the accounts kept by the pool country. Normally, the audit of government accounts is the exclusive domain of the national boards of auditors. Again, the pool country may be expected to accept responsibility for what is not fully under its control, if, for instance, an external audit leads to a request for reimbursement of expenditure considered inappropriate.

151. In practice, these and other legal difficulties can mostly be overcome, although they may well create delays in the evaluation process before they are sorted out. So far, a strong spirit of cooperation prevails among the partners of joint evaluations. This, and the general readiness to search for pragmatic, workable solutions if problems of principles endanger the joint work, has helped to avoid the collapse of the financing arrangements. In the search for solutions to blockages, some agencies have even

20. As a follow up to this study, the DAC Evaluation Network may want to consider commissioning a legal expert with a more thorough analysis of existing contractual arrangements in joint evaluation work with a view to come up with practical guidance and recommendations in this field (“Key elements of contract design for joint evaluations”).

demonstrated significant creativity and ingenuity (not always contributing to the edification of legal advisers within their home ministries).

Using modern communications technology

152. One consultant, working as part of a consortium for a recent joint evaluation, received over 850 e-mails, within a few weeks, concerning the organisation of the work of the consortium. He had to deal with almost all of these before he could spend any time on the substance of the evaluation. This example highlights the problem with the electronic mail system that is always at hand; anywhere in the world, 24 hours a day. While e-mail can go a long way in helping to alleviate workloads, it can also contribute to more work through indiscriminate information traffic. Joint evaluations are prone to this danger, and all partners must exercise discrimination in their use of e-mail and other communications. All information does not need to be known by all partners, and the copy and blind copy functions of the e-mail system should be used with particular restraint and moderation.

153. This is not to deny the importance of communications and information sharing, but recent joint evaluations have explored a new way of handling information and making it available to those who need it, without necessarily adding to the problem of information overload. The method is to set up a special website for the joint evaluation, which becomes the platform where information emanating from and related to the evaluation process, such as inception and progress reports, records of meetings, draft reports, and so on, is posted. If necessary, part of the website can be turned into a restricted area, only accessible for members of the steering committee, the management group and the consultants. A good example of a joint evaluation website is the GBS site, which is housed within the DAC Evaluation website.

Main conclusions and recommendations

- The governance and management structure for a joint evaluation depends to a large extent on three factors: (I) the size of the group of partners; (II) the character of the evaluation; and (III) the basic approach to the exercise (centralised or decentralised; delegation of work; silent partnerships; parallel work; and so on).
- The inclusion of policy and operational staff can work well within steering committees of those evaluations dealing with global policy and impact and effectiveness analyses as well as with thematic and sector issues. Non-evaluation staff bring useful expertise and a strong notion of realism, and often contribute to ensuring the evaluation results are relevant for practitioners.
- In contrast, the involvement of policy and operational staff in the management of joint evaluations dealing with the specific multilateral institutions, with country strategies and individual programmes or projects is less advisable; it can be misinterpreted and seen as an attempt to influence the outcome of the evaluation.
- There is a need for full transparency in the governance structure, especially in cases where the steering committee has delegated significant decision-making authority to a management group. The simplest way of establishing transparency is to make provision for the quick and timely circulation of full records of all meetings.
- It is a wise investment to hire a person specifically to keep records of the meetings, which reflect the individual contributions and opinions of members, summarize definitions agreed upon, and list all the decisions.

- There are, at least, two different schools of thought on ToR. One requires the ToR to describe in great detail the purpose, objectives, subject matter, scope of work, methodologies, and so on, and to include a full evaluation matrix with objectives, indicators, benchmarks etc. The other philosophy is the opposite: to keep the initial ToR short and simple; describing the purpose, the setting, the subject and the framework, and then leave it to the consultants to develop this into a full proposal. Evaluators should agree on one or the other option, not a compromise between the two.
- If country case studies are part of a joint evaluation, aid practitioners on the spot are likely to be keen on recommendations to help them with the problems of their daily work.
- Follow-up merits more attention in the initial phase of drawing up the ToR. Events which can be linked to need to be worked into the evaluation timetable early on. There are also possibilities to organise specific follow-up activities, including the presentation of the evaluation results to a joint workshop of policy staff from the agencies involved, information sessions for the staff of individual agencies, or, very important to learn from experience, an evaluation workshop for the evaluation process (“post mortem”).
- Developing the budget for a joint evaluation is a crucial step in preparing the ground for a successful exercise. It is preferable to start the budget process with the question of cost, and not with the financing side. The costing of a joint evaluation must be realistic, and should encompass all elements of the evaluation cycle, such as quality assurance; translation, editing, printing and dissemination of the reports; follow up through workshops, seminars and other events; and contingencies.
- The size of the final budget will only be known after the contract negotiations with the bidder have come to a successful end and the contract is ready to be awarded. That should be the moment for all partners to pledge their contributions.
- It is important to make absolutely sure that enough time has been scheduled for the selection of the consultants. A minimum of three to four months should be allowed for the selection process (after the date of publication of the invitation to bid).
- Consultant consortia can be complex and relatively heavy in structure. The members of steering committees and management groups must reserve full time to carefully analyse and scrutinize the bids.
- Different kinds of contracts and agreements are required to establish legally binding relationships between the different partners in an evaluation. This is particularly relevant for co-funding agencies in pooling arrangements. This may lead to delays in the process and may require looking for pragmatic and administratively creative solutions.
- A particularly helpful way of handling information and making it available to those who need it, without necessarily adding to the problem of information overload, is to set up a dedicated website. This website acts as a platform on which to post the information emanating from and relayed to the evaluation process. If necessary, part of the website can be turned into a restricted area.

2.3 Implementing joint evaluations

154. Effective Practices in Conducting a Joint Multi-Donor Evaluation lists several recommendations for implementation:

- Holding a planning workshop at the beginning of a joint evaluation will help the team to get off to a good start. Evaluation team members as well as representatives from the donor management group should, if possible, participate. The purpose of the workshop should be to build an effective team that shares common understandings of the evaluation's purpose and plans.
- Evaluation teams often benefit from considerable diversity, such as multi-national 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, as may differences in perspectives and opinions.
- While adding to cost and effort, a fieldwork phase will also add to the timeliness, quality, and credibility of the evaluation's findings.
- All too often, teams do not gather information and views directly from the intended programme beneficiaries. While this does require extra cost and time, it is often very enlightening.
- While it makes sense for team members in the field to split up 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. With continuous interaction, reaching team consensus on evaluation findings will be easier.
- It is good practice to share the preliminary draft report with those organisations and key individuals in the field that have been especially involved in the evaluation.
- The team may also consider a stakeholder workshop, at which the report's findings and conclusions may be presented, discussed, and commented on by participants.

155. The way a joint evaluation is planned, prepared, organised and set up largely, but not exclusively, determines and shapes the implementation phase. However, other factors will also play a role once the evaluation is underway. These factors include the cooperation among the sponsoring agencies, the collaboration among the consultants, the relationship between the management and steering groups and the consultants, and the quality of the intermediate and final products. The implementation phase is the litmus test for the quality and adequacy of the initial preparations, but it is also normally a period of unforeseen challenges that require all actors to demonstrate flexibility, mutual understanding and patience.

156. An update on lessons learned in implementing joint evaluations are grouped and presented under the following categories: (I) Steering committees; (II) Management groups; (III) Consultants; (IV) Quality assurance; (V) Field work; (VI) Crises.

Steering committees

157. The steering committee is the central forum in any joint evaluation. It represents the sponsors and financiers of the evaluation, who are normally also important stakeholders in the exercise. The steering committee has responsibility for the smooth running and the quality of the results of the evaluation, and is held accountable for both. The steering committee is also the employer of the contractor(s) (i.e. the

consultant(s)), and the central contact for the evaluated (be it an institution, a concept, a country, a sector, a programme or a project).

158. It may well be the awareness of these strong responsibilities that often tempts steering committees to opt for tight control mechanisms *vis-à-vis* the consultants. The result of such an approach is rarely satisfactory, and the steering committee (with or without the support of a management group) can get bogged down in micro-management and become unable to 'see the wood for the trees'. A broad and comprehensive perspective is, however, essential if the steering committee is to steer the evaluation through to a successful conclusion.

159. Recent examples of joint evaluations demonstrate that the actors are becoming much more aware of the risk of getting lost in the details of micro-management. Steering Committees have started to experiment with innovative modalities that allow them to concentrate on playing the oversight role, without neglecting to assure the necessary degree of quality control over the consultants. Three modalities are of particular significance, and are explored below: management groups; quality assurance systems; and external advisors.

160. It is always important for a steering committee to find its own balance between oversight and control, depending on the specific circumstances of the joint evaluation. However, this balance should be reviewed regularly and in the light of external reactions and feedback. When an evaluation is perceived as running into difficulties, the bureaucratic instinct is to tighten the grip and control. This, however, can easily damage good intentions and good relations with the consultants. It is recommended to set aside time, within each of the steering committee meetings, to reflect carefully on its own role and how this is evolving.

161. One way of tackling some of the challenges of a steering committee, used in the evaluation of the Enabling Development Policy of WFP, is to institutionalise the rotation of the Chair. In the context of the WFP evaluation, this proved to be a 'win-win' solution. It contributed to broadening the sense of ownership of steering committee members, to demonstrating the joint responsibility for the success of the exercise, and to avoiding the resource-burden imposed on one agency by a one-chair arrangement. Most importantly, it also helped to allow the different temperaments and characters, represented in the group, to come fully into play. Although this is not a universally accepted approach, it can go a long way in balancing the overall approach of the steering committee and in setting the atmosphere surrounding it.

162. A steering committee must also agree, early on, as to the degree of its direct involvement in the evaluation process. Especially with country studies being undertaken as part of a larger evaluation, members of steering committees – and management groups – sometimes wish to join the evaluation mission as observers, either for a few days or for the full mission. The circumstances of the specific case have to be looked at carefully, and a cautious approach is advised. There are various problems that can occur: one is that the 'observers' can become labelled as watchdogs, putting them in an uncomfortable position. Another problem one can call the 'Animal Farm' syndrome, meaning that some members of the steering committee become 'more equal' (and exert greater influence) than the other members because they travel with the consultants. This raises questions of preferential treatment of or discriminating against specific evaluation missions (depending on your perspective). There is clearly also the issue of the independence and impartiality of the consultants when steering committee members join an evaluation mission. Judging from past experience, steering committees are well advised to refrain from getting too involved in the evaluation process itself, and to limit such involvement to exceptional and well-justified cases.

163. There is always a temptation, for steering committees, to add to the existing tasks of the consultants as new and interesting questions emerge during the evaluation process, or as new people join

the steering group. A steering committee, however, and especially the Chair of the group, should resist this temptation as strongly as possible and avoid allowing its members to raise many additional questions or to instil a notion of special interests into the evaluation. Consultants will normally try to accommodate additional wishes coming from their employers, but more often than not, this will complicate the process, reduce the manageability of the evaluation, and blur the focus of the work. It can also, of course, add to the cost of the exercise or spread the available resources more thinly.

164. Lack of continuity in the membership of steering committees is a key problem in their work and, more recently, also in that of management groups. The longer the evaluation lasts, the more likely it is that changes in membership in the group will occur. This is rarely due to deliberate withdrawal, but more normally to outside factors such as the systems of staff rotation, transfers and promotion within the member country governments. However, as a high degree of continuity among the members of the steering committee is a vital precondition for a successful evaluation, agencies contemplating joining a joint evaluation should be encouraged to make sure, to the extent possible, that their representatives in the steering and management groups will be able to remain as representatives throughout the process.

Management groups

165. It is now almost routine procedure to establish a management group for larger and more complex joint evaluations. A management group is a small and flexible body that can communicate quickly among itself and move and meet at relatively short notice. It is normally made up of members from the agencies that are taking a particularly strong interest in the evaluation. The members of a management group have needed to set aside a significant share of their work capacity for the joint evaluation. Management groups usually have significant latitude for running the day-to-day business of a joint evaluation. Only major decisions are subject to ex ante approval or ex post confirmation by the steering committee.

166. In order to fulfil its crucial role, the management group has to be assured that it can function without too many obstacles. First and foremost, it is vital that the management group does not suffer from changes in personnel (as has happened recently in some major joint evaluations). Agencies willing to join a management group should be prepared not only to invest the required staff-time, but also to maintain the continuity of their personnel. If that cannot be guaranteed with reasonable assurance, agencies should seriously reconsider their decision to join the group, or at least make provision for an overlap of predecessor and successor to allow for a smooth takeover of duties and the continued functioning of the management group.

167. One of the most important roles of the management group is efficient liaison with the consultants on a regular (at times of peak activity sometimes even daily) basis. A good and open relationship between the two is of key importance for the smooth and successful functioning of the evaluation process. Instructions to the consultants emanating from a steering group are often unclear or even contradictory. Therefore, the management group must be as precise and specific as possible in its guidance. There is even a role for the management group in translating complex steering committee ideas and inputs into reasonable and manageable instructions to the consultants. Conversely, the management group can also play a valuable role in filtering the requests and requirements formulated by the consultants and turning them into meaningful items for discussion in the steering committee. In other words, the management group has discretionary power of its own and is also an important interface in the communication between consultants and steering committee. That role requires strong sensitivity and communication skills.²¹

21. An interesting example of communicating key information to a relatively large number of evaluators who were about to go out to do a series of country studies, was developed and tested in the evaluation of the Enabling Development Policy of WFP. A short, two and a half pages paper called *Pointers for evaluators* summarized all the important points discussed between steering and management groups and the

168. Part of the role of the management group is also how to handle the question of access to confidential material. Can that be organised for an international team of consultants, and if so, how and under what conditions?

169. Another issue is the careful preparation of field visits. The green light of partner country governments for such visits must be obtained beforehand. Someone from among the sponsors of the evaluation has to accept responsibility for this and for the team of consultants that undertakes the field studies. This involves additional work for embassies or aid missions, who may want to know why they should assist the work of an evaluation team that may have only very little bearing on their day-to-day problems. The management group therefore has to undertake a great deal of explanatory, diplomatic and mediating work; which requires sensitivity, patience and perseverance.

170. Preferably at the very start of an evaluation, the management group should discuss the possibility of commissioning external expertise to support its work. This support could encompass help with the keeping of full records of meetings (and their quick turnaround and distribution); a catalytic role in bringing new and emerging issues to the attention of the management group; the role of moderator or facilitator, if needed; and advice on substance and evaluation matters, to the extent possible.

171. Another important function of such an advisor could be to keep track of the chronology of the evaluation and to register all the lessons learned during the evaluation process. This information could be collected in an "evaluation diary". This diary would be submitted to the steering and management groups as well as to the team of consultants as the basis for a "post mortem" or critical *ex-post* assessment of the evaluation process. This would facilitate systematic and objective reflection on the experiences gained, the strengths and weaknesses of the process, and the lessons learned for future work.

172. A critical phase in a joint evaluation is when draft reports become available and are circulated for comments. The comments received are usually wide ranging, often only very general in nature and frequently omit concrete proposals for changes or new formulations. On the other hand, many comments are very detailed (and sometimes overly so). Comments usually cover the whole range of issues; dealing with substance, methodology, findings, conclusions and judgements, and often they miss important points made by the consultants, or create misunderstandings. Most importantly, however, they can be quite contradictory in nature and substance. The consultants expect, or at least hope, that the management group will make the effort to consolidate the various comments into one set of comments, ironing out all contradictions, before passing them on to the evaluation team.

173. However, this expectation expects the impossible, mainly for two reasons: (I) a management group cannot substitute for the professional expertise and judgement of the evaluators to assess the relevance of comments made; and (II) it would be diplomatically insensitive for a management group to decide that the comment of agency X on a certain aspect carried more weight than that of agency Y. Some possibilities to help the consultants come to grips with this problem will exist if there is a proper quality assurance system in place (see below). Nevertheless, the need for consultants to continue to deal with unconsolidated comments will not disappear.

evaluation team during the initial phase, thus providing everyone with a clear picture of the purpose, subject matter, scope, and substantial and methodological thrust of the evaluation. It also encouraged the evaluators to contact one of the two senior quality advisors in cases where, in their work they "would come across particularly sensitive information that they would not like to share too widely." The strictly confidential treatment of such information by the senior quality advisors was guaranteed. Feed back later on indicated that evaluators had found this way of drawing their attention to the crucial points of the evaluation as helpful as the indication of where to resort to in difficult situations.

174. Finally, a word should be said on the specific role of the lead agency or lead country in the management structure of a multi-partner evaluation. A lead role is normally ascribed to an agency or a country that has taken the initiative for the evaluation, or accepted certain duties and responsibilities such as administering the pool of funds, acting as employer of the contractor, shouldering a particularly large share of the total cost, or playing a more prominent role - for example as chair of the steering committee or management group. Although many of these characteristics tend to coincide in many cases, and therefore make it easy to identify a lead agency or country, this is by no means a natural law. Some thought should therefore be given in future, as appropriate, to the possibilities of dividing some of these functions among different actors according to their comparative advantages. The pool of funds, for instance, may be best established in an agency that has flexible rules for financial management. Similarly, the contractual arrangements with the consultants could be the responsibility of the agency which allows the most flexibility and least bureaucratic expense.

Consultants

175. Consultants and partners in a joint evaluation are inseparable entities - one would lose its *raison d'être* without the other. Their relationship is dialectical; and therefore far from easy. Nevertheless, it is of the utmost importance to establish a professional and friendly working relationship.

176. As a first step to building this relationship, it is essential to understand the interest of the consultants in a specific evaluation, the stake they have in it, the risks they are willing to accept, and the limits which they are not prepared to exceed. A number of important points which should be taken into consideration in this context have already been made in section 2.2 (above). In addition, consultants should be encouraged to share with the steering and management groups their motivation (aside from the financial motive) for bidding for a specific evaluation. Such motivations could include the potential reputation building; an express interest and experience in the subject matter; the hope for follow up contracts after the successful completion of the joint evaluation; or, in the case of research institutions, their specific research interests. It is important to have transparency in this regard so that it will be easier at a later stage to interpret and understand the reactions of the consultants.

177. Other important elements of establishing an early understanding between the consultants and the steering committee and management group are the clarification of expectations, the agreement on certain rules, and the definition of key terms used in the bidding documents, the terms of reference and other material. A lot of the complications and frustrations that commonly occur, especially at a later stage, are due to the lack of early efforts to work towards a common understanding. One recent joint evaluation experienced serious delays because the consultants produced an inception report which did not meet the needs of the steering committee. A large part of the problem was that no-one had thought to agree in advance what was meant by an inception report. As a result, the consultants produced a long report basically summarising the work up to that point. However, it did not contain key elements that the steering committee wanted to see, for example a SWOT analysis. There was frustration on all sides, because no one had thought of establishing a common understanding up front.

178. There are many different ways of establishing the needed common understanding. Partners can prepare and agree a paper containing the essential definitions and interpretations. Differences of views that can exist at the outset will become apparent through this process and can then be addressed. If more time is needed, a teambuilding workshop might be held to undertake this activity.

179. Another issue that needs to be addressed early on (this issue is usually referred to in the bidding documents) is the question of incompatibility. This term needs clarification, so that it can be turned into an operational category for the consultants. The group of agencies sponsoring the evaluation will have to decide how they wish to define incompatibility (some agencies are more rigid in this regard than others).

The line of compromise between the different positions is in most cases the need to avoid any danger of being accused by stakeholders or third parties of jeopardising the independence and impartiality of the evaluation.

180. As in steering committees and management groups, the continuity of key personnel in the consultant group is a prerequisite for the smooth execution of a joint evaluation. Therefore, the continuing availability of the key consultants, throughout the evaluation process, should be a firm part of the contractual arrangements. The early agreement of a yardstick with which to measure incompatibility helps with the continuity of evaluation personnel; evaluation teams are often put together before such yardsticks are agreed and have to be reshuffled later because an individual cannot be considered impartial in light of the incompatibility rules.

181. If consultants work for extended periods of time without the opportunity to present intermediate results and raise issues for discussion, they will not get the feedback which they may require and which would help them to understand better if they are on the right track. Opportunities for feedback loops should be jointly explored and identified.

182. Many recent joint evaluations have included extensive field work, mostly in the form of a series of country case studies to collect empirical evidence. This is an area in need of more attention; it is essential to make sure that the results of field studies are comparable when the synthesis work begins. There are many ways of working towards comparability. Clearly, the terms of reference and the evaluation matrix are key tools. The same is true for designing questionnaires and interview guides that are to be used in all case studies.

Box 6. 16 Golden Rules for Consultants

The International Programme for Development Evaluation Training (IPDET) which is sponsored by the World Bank, Carlton University of Ottawa, and several bilateral donors takes place each summer in Ottawa, Canada. In 2004, a new training module on joint multi-partner evaluations was added to the course programme. One of the speakers was Ted Freeman, Partner of the Canadian consulting firm Goss Gilroy Inc. His 16 rules for organising and managing the external evaluation team, presented at the IPDET course, are extremely useful:

1. While encompassing important expertise on the sector, sub-sector and geographical areas under evaluation ensure that the core expertise in complex, large scale evaluations of development cooperation is as solid as possible.
2. In most cases a consortium of firms and/or research institutions will be needed – keep the organization as simple as possible and, whenever possible, work with organizations you have worked with before.
3. The lead organization in the consortium should be one with a strong commitment and track record in the evaluation of international development cooperation. The project should be a natural fit with its core business and markets.
4. Commitment to the evaluation should be made clear at the board level of the main external evaluation organization.
5. National consultants should be integrated into the process of the international competitive bid and should take part in methodology selection and design.
6. In multi-country studies, each field team should combine resources from different organizations in the consortia rather than having each organization specialize geographically

or institutionally.

7. Evaluation team workshops to develop a common approach to measurement and reporting are invaluable.
8. The evaluation team is ultimately responsible to the overall Evaluation Steering Committee; whenever possible this should be done with and through a smaller management group.
9. Meetings with the Steering Committee will be less frequent given the expenses of assembling the committee but sufficient time will be needed to allow for full discussions and working out of a common position.
10. It is always useful to present preliminary evaluation findings to the Steering Committee (along with basic evidence) in advance of the presentation of the Draft Report itself.
11. In joint evaluations it is essential that the external evaluators operate openly and transparently and are able in presentations and reports to directly link methods to evidence gathered to findings, conclusions and recommendations.
12. In negotiations for additional resources, when they are clearly needed, the evaluation team and the management group will need to begin by agreeing on the split between work which should be undertaken as a result of the original contract (and using the original resource envelope) and work which is the result of new issues and interests or arises from unforeseeable circumstances. This will require the team to prepare detailed, costed and time-bound plans for any new work required.
13. Large, lengthy, complex and high-stakes joint evaluations require all stakeholders to maintain a strong positive orientation throughout the exercise.
14. It is essential that the external evaluation team is responsive to all members of the Steering Committee as having essentially equal weight (...). It is equally essential that the evaluation team can demonstrate an absence of institutional bias.
15. Draft reports are never perfect. Both the evaluation team and the Steering Committee should enter discussions on drafts with an open attitude to improvements which can be made. At the same time the evaluators must be able and willing to maintain their objective responsibility for evaluation findings and conclusions.
16. The cost, complexity and duration of joint evaluations argue strongly for investing a substantial proportion of the budget in dissemination and follow up activities.

183. Moreover, there will be different evaluation teams going to different countries, but all of them are expected to produce comparable results. The teams will often be made up of a mix of international and national consultants, with the latter only joining the team upon its arrival in country. Therefore, a lot of effort has to be put into the harmonisation of approaches and methods. As a minimum, there should be a team leader workshop early enough before the country studies to allow for the full incorporation of the workshop results in the preparation of the field visits. Depending on the complexity of the field work envisaged, and on the role the local consultants will be assigned in the evaluation, preparatory trips of the team leaders to “their” countries may actually be a good investment to acquaint them with the local situation, with the opportunities and risks of the planned evaluation, and, last but not least, with the national consultants and their qualifications and potentials.

184. In addition, there are cases where the first country-level empirical study is used to test the approach and methods before the other country studies are undertaken. It is advisable, in this circumstance, to include more than one of the team leaders in the test mission.

185. Finally, a debriefing workshop for the team leaders, after they have returned from the field, is a useful way of comparing notes and of working towards synergies that will feed into the synthesis report. Obviously, these activities raise the cost of a joint evaluation and therefore need to be factored into the budget (but compared to the overall cost they are relatively small expenditure items with significant promise for high returns).

Quality assurance

186. It is an absolute must that consultants bidding for an evaluation contract present a convincing internal quality management and assurance system that covers the whole range of work to be performed; such as substance, methods, language, editing, and so on. Management groups and steering committees should insist on hard proof that this kind of quality assurance is in place from the very first day of work. Shortcomings in this regard, which are not redressed quickly are among the main reasons for difficulties, debates and conflict affecting joint evaluations.

187. The emerging trend is that it is no longer considered sufficient to rely on the quality assurance and control systems of the consultants. Quality assurance by the agencies which support a joint evaluation has gained in importance over recent years. Most of this quality assurance is undertaken by the members of the management and steering groups (with or without support and inputs from colleagues in the substantive divisions of their agencies). In a few particularly complex evaluations, such as the CDF and IFAD evaluations, advisory panels were established; with a view to assessing the quality of the work performed and to securing the acceptability of the products of the evaluation.

188. However, awareness is also growing that steering and management groups find it increasingly difficult to cope with all the quality management work themselves. The complexity of many evaluations, and the work pressure on members of governing bodies, make it more and more imperative to add a quality assurance component to the governance structure. Although this quality assurance component is primarily created for the purposes of the management and steering groups, it also provides a mechanism for early and efficient liaison with the quality management systems of the consultant. In this way, significant synergies in quality assurance efforts can be achieved.

189. An example of this new form of quality management which turned out to be successful in helping to improve the quality of the products upstream in the process and in relieving the governance groups of some of their work load, was the appointment of two senior quality advisors for the evaluation of the Enabling Development Policy of the WFP. One of them was an expert in the subject matter of food aid for development and the other one an evaluation specialist.

190. It was agreed that the consultants would submit early drafts of their reports to the senior quality advisors. The advisors would then provide the consultants with initial comments, express encouragement and criticism, and make suggestions for improvements, clarifications and for additions and deletions. All this was kept relatively informal, so that the potential for confrontational situations was largely avoided. In one or two cases, the criticisms of the senior quality advisors also contributed to strengthening the hand of the internal quality management of the consortium. Once the early drafts had been revised in view of the reactions of the senior quality advisors, revised drafts were then officially submitted to the steering group. Thus, the informal round of comments reduced the need for official comments and for long discussions and significant savings of time were achieved in the meetings of the steering committee.

191. The fact that a well designed quality management and assurance system can increase the time-efficiency of a joint evaluation, and also strengthen its effectiveness should be taken as encouragement for further experiments in this area. These should also begin to address the issue of a possible new weighting of the relative importance of steering committees, management groups and quality assurance systems in the setting up of a joint evaluation.

Field work

192. The first step toward the field work (in those joint evaluations that contain a field work component) is the selection of countries or case studies. Although there have been some serious efforts to rationalise the selection process, and to base it either on a set of clearly defined criteria for selection or else on random sampling, a lot of ad hoc decisions still prevail. If outsiders perceive an arbitrary and/or biased selection of case studies this will be seen to adversely impact on the impartiality and credibility of the evaluation. Partners should therefore make every effort to rationalise their selection process and to make it as transparent as possible.

193. Transparency is also required when it comes to the preparation and implementation of the field studies. Partner governments have to be officially informed about the proposed visit of a team of consultants. This is important to avoid any suspicions, and should be done early enough to allow for sufficient lead time to find another country if something goes wrong with the first choice. Partner country authorities should also be given copies of the ToR as well as the CVs of the team members; and they should be consulted on the programme of visits. Finally, partner governments should be encouraged to nominate a participant in the evaluation, either as an observer or as a resource person. In most cases, this invitation will be politely turned down, for lack of staff resources or qualified evaluation specialists.

194. Partner country governments usually show a strong interest in a briefing workshop at the beginning of the evaluation, which can be attended by all or by many of the stakeholders, and a debriefing event at the end of the field work. The initial briefing workshops should allow for an open discussion. At the debriefing event, the consultants should present their main findings and conclusions, and also indicate the kind of recommendations they are intending to make. It is very damaging for the in-country perception of the character of an evaluation if there is a strong discrepancy between the initial presentation of findings and recommendations at the end of the field work, and in the final report. The best way to ensure that such discrepancies do not occur is by preparing a written *aide-mémoire* of the debriefing session.

195. Field work can become necessary if the purpose of a joint evaluation is the assessment of an institution, or of a global or sector theme. When an institution is evaluated, the proposed programme for the field visits is often drawn up by the local office of that institution. The consultant team's input into the programme is then very limited, and this can entail the risk of a perceived lack of independence and impartiality in these decisions.

196. This perception will be further enhanced if the evaluation team makes extensive use of the logistical support offered by the local office; such as transportation in official vehicles, translation, and so on. The readiness of key informants to speak to the consultants in a frank and open manner can be impaired if the evaluation team emerges from a cavalcade of official cars. A good way to deal with this problem during the preparation of field visits is to enable the team leader to travel to the country on a scouting mission. This allows him to exert more influence on the programme of visits.

197. Finally, we come to the crucial role of local consultants in the evaluation process. There are a growing number of dedicated and well qualified consultants in partner countries that are very capable of delivering thorough evaluation work. Even though it is no longer difficult to identify these consultants, for

instance through networks such as evaluation associations, there are a number of obstacles to contracting them.

198. The likelihood that a local consultant has been in prior contact with the subject of the evaluation, and that he or she is therefore disqualified due to the incompatibility rule, is much greater for national consultants than for their international colleagues (because the consultancy scene in a partner country is usually small). Moreover, local consultants find a great deal of their employment among the bilateral and multilateral aid agencies operating in a country, and that may also affect their neutrality and impartiality. In addition, there may be political and social pressures in a country that can reduce the freedom of action for a local consultant or else put him/her in awkward, perhaps even dangerous, situations. A critical report signed off by a local consultant together with his/her international colleagues, often entails the risk of impinging on his/her prospects for new assignments in future; because this kind of criticism may be culturally and socially unacceptable.

199. However, international consultants occasionally use the arguments cited above to design a low-key role for their local colleagues and thus restrict their chances for a full contribution to the evaluation. This is short-sighted and can cause resentment among team members. It is also unnecessary. Local consultants themselves know best what risks they can and cannot afford to take. Similarly, local consultants are often not assigned full responsibility for the drafting of a whole chapter of the report. Technical communication problems are cited in justifying that decision.

200. It seems that international consultants in a team vary significantly in the amount of information they share, with their local colleagues, acquired from files and documents during the inception phase and prior to the field work. This information may be privileged and not readily available to the local consultants - unless shared freely by their international colleagues.

201. And finally, insufficient evidence exists of serious efforts for team building among local and international consultants at the beginning of the field work.

202. A lot remains to be done to develop a level playing field for local and international consultants. Both have a lot to contribute to a joint evaluation, both have their own comparative advantages, and both have certain disadvantages. But even in shrinking markets for consultancy work (with increasing feelings of trade jealousy and competition) there must be room for all when working together as a team; and there is certainly no need to try and keep the other group away at arm's length.

Crises

203. Very few joint evaluations have not gone through some form of crisis.

204. The symptoms of such crises may include; a rapidly deteriorating relationship between employer and contractor; the slow or sudden withdrawal of support by individual agencies; or the resignation of key personnel (without offering an acceptable justification).

205. The recognition that the quality of work of a consultant is unacceptable, that it cannot be sufficiently improved, and the evaluation process therefore returns to square one can easily lead to a crisis with evaluation partners almost instinctively looking for a scapegoat. The same happens if there are serious delays in the schedule of an evaluation, or a decreasing interest in its results because of new and unforeseen developments or due to too long a duration of the process.

206. There is very little that can be learned for a crisis situation from past experiences; because every crisis is unique. What can be learned, however, is one key generic lesson; if there is a crisis appearing on the horizon, do not wait and see if it will go away - it won't. On the contrary, the longer a crisis is allowed

to develop, the less manageable it will become. The way to handle the situation is to do deal with it in a determined manner as soon as it is recognized. If a consultant needs to be fired, do it. If a partner in an evaluation wants to withdraw from it, let him go (making sure that he honours his financial commitments). If there is a decline in interest in the results of an evaluation, think of ways to inspire renewed interest in them. In summary: if there is a crisis looming, don't drag your feet, demonstrate leadership, and act, quickly.

Main conclusions and recommendations

- The steering committee has full responsibility for the smooth running and quality of the results of the evaluation and will be held accountable for both. Not surprisingly, the awareness of this responsibility often tempts steering committee members to opt for relatively tight control mechanisms vis-à-vis the consultants.
- A steering committee must find its own balance between oversight and control - depending on the specific circumstances of the evaluation. When an evaluation is becoming problematic, the bureaucratic instinct to tighten the grip can easily gain the upper hand and damage good intentions. It is important to set aside time to reflect on the steering committee's role and how it evolves with the progress of the evaluation.
- Lack of continuity in the membership of steering and management committees is one of the key problems for joint evaluations. This is mainly due to outside factors that are hard to control. However, agencies considering joining a multi-partner evaluation should make sure that their representatives in the steering and management groups will be able to remain there throughout the evaluation process.
- One of the most important roles of the management group is close, efficient and regular liaison with the consultant team. A good and open relationship between the two is therefore of key importance.
- Preferably at the very beginning of an evaluation, the management group should discuss the possibility of commissioning an external expert to support its work. This support can encompass help with the keeping of full records of meetings and their quick turnaround and distribution; a catalytic role in bringing new and emerging issues to the attention of the management group; the role of moderator or facilitator, if needed; and providing advice on substance and evaluation matters.
- One important function of an advisor of the kind mentioned above could be to keep track of the chronology of the evaluation and to register all the lessons learned in an "evaluation diary". At the end of the evaluation, this diary should be the basis for a "post mortem" or critical assessment of the evaluation process.
- It is essential to understand the interest of the consultants in a specific evaluation. Consultants should be encouraged to share with the steering group their motivation for bidding for a specific evaluation (aside from the financial incentive). Such motivation can include the potential reputation to be gained; an express interest and experience in the subject matter; the hope for follow up contracts; or, in the case of research institutions, their specific research interests.
- Other important elements of establishing an early understanding between the consultants and the steering committee and management group are the clarification of expectations, the agreement on

certain rules, and the definition of key terms used in the bidding documents, the terms of reference and other material.

- The issue of incompatibility needs to be addressed early and openly. The term needs clarification quickly so that it can be turned into an operational category for the consultants.
- The continuity of key personnel in the consultant-teams is also a prerequisite for the smooth execution of a joint evaluation. The continued availability of the key evaluation staff should be a firm part of the contractual arrangements.
- Depending on the complexity of the field work and on the role of the local consultants, preparatory trips of the team leaders to “their” countries can be a good investment to acquaint them with the local situation, with the opportunities and risks of the planned evaluation, and with the national consultants,
- Consultants bidding for an evaluation contract must present a convincing internal quality management and assurance system, that covers the whole range of work to be performed, such as substance, methods, language, editing, and so on. Management groups and steering committees are advised to insist on hard proof for this kind of quality assurance.
- The complexity of many evaluations and the work pressure on members of governing bodies make it more and more imperative to add a quality assurance component to the governance structure of a joint evaluation. This also provides a mechanism for early and efficient liaison with the quality management of the consultant. Thus, significant synergies in quality assurance efforts can be achieved.
- Partner country governments usually show a strong interest in a briefing workshop at the beginning of the evaluation and a debriefing event at the end of the field work.
- When the purpose of a joint evaluation is the assessment of an institution, the proposed programme for the field visits is often drawn up by the local office of that institution. This entails the risk of a perceived lack of independence and impartiality of the evaluators. This perception will be further enhanced if the evaluation team makes wide use of logistical support offered by the local office, such as transportation in official vehicles, translation, and so on. A good way to deal with these problems is for the team leader to travel to the country on a scouting mission. This allows him/her to exert more influence on the programme of visits.
- A lot remains to be done to work toward a level playing field for local and international consultants. Both have a lot to contribute to a joint evaluation; both have comparative advantages and both have disadvantages. But even in shrinking markets for consultancy work, there must be room for all to work together in a team.
- Each evaluation crisis is unique. The key generic lesson is that if a crisis starts to appear on the horizon, do not wait and see if it will go away. It won't. On the contrary, the longer a crisis is allowed to develop, the less manageable it will become. The way to handle the situation is to deal with it determinedly as soon as it is recognised.

2.4 Following up on joint evaluations

207. *Effective Practices in Conducting a Joint Multi-Donor Evaluation* contains very little guidance on follow-up practices. There are some references to the need to communicate the results of joint

evaluations, through publications, conferences, workshops, and so on. There is also a reference to the potential of well-structured monitoring systems as a way to encourage agencies to account for their responses to the findings and recommendations of a joint evaluation. The Joint Evaluation Follow-up Monitoring and Facilitation Network (JEFF) set up in the wake of the Rwanda evaluation, is cited as a good example for this approach. The question of compliance with the recommendations of joint evaluations is played down, because “joint multi-donor evaluations usually raise broad system-wide issues and recommendations that relate to a diverse range of organisations.” Compliance with recommendations can therefore not be compelled.

208. This new study indicates that follow-up on joint evaluations remains a weak link in the chain. There is very little systematic knowledge on follow-up that can be shared. This suggests that there is no common agreement yet on the kind of follow-up suitable and appropriate for joint evaluations.

209. All stakeholders point to the fact that an evaluation is a learning process that will have an impact on the way business is done in future. Some also stress that there are examples of joint evaluations which, although still ongoing, have already had an impact on the policies and practices under review. It is true that some of the institutions evaluated in recent years have shown a propensity to implement emerging recommendations before the end of the evaluation process.²² In addition to the need to implement helpful recommendations as early as possible, one additional reason can be the prospect of discussions in the governing bodies - and the perceived need to be ready to show a good degree of responsiveness to the evaluation results in those high-level discussions.

210. Every bilateral agency has its own way of dealing with the outcomes of joint evaluations. Some have decided to give joint evaluations the same treatment as their own evaluations - ie they request a management response, submit it together with the report to their audit and evaluation boards for review and discussion, or send to parliament for information. Some even go as far as establishing action plans for the implementation of the recommendations.

211. However, many agencies proceed more cautiously and on a case-by-case basis. Depending on their interest and stakes in the joint evaluation, they prepare the follow-up process in a tailor-made *ex-post* fashion. Unfortunately, this approach can potentially lead to the decision to file the final report away rather than to support its recommendations.

212. To date, one of the major shortcomings in following up on joint evaluations is the fact that in most cases follow-up questions are not addressed at the start - but only at a later stage in the process once the first contours of the emerging results become apparent. That is often too late to develop a well-designed, common strategy for follow-up which, for instance, could include the preparation of joint events to present, discuss and disseminate the findings and conclusions of the evaluation, or joint presentations at international meetings and events that relate to the subject of the evaluation and would allow the wider dissemination of the work accomplished. As a consequence, evaluations that started out as joint activities, can evolve into fragmented follow-up processes and so lose a lot of the coordinated drive and impact.

213. A key issue with regard to follow-up is the commitment of the agencies to the joint evaluations they join, including to their results and recommendations. The level of this commitment should be articulated early on in the evaluation process so that everyone knows how far the partners are prepared to go in their commitment. The degree of commitment also depends on the motives of an agency in participating in a joint evaluation. Some of these motives may actually work against firm commitment;

22. A notable example of this pro-active approach to intermediate evaluation results is the UNFPA-IPPF evaluation. Both organisations informed their senior management and governing bodies regularly about the progress of the evaluation and any findings considered relevant to be discussed in greater detail.

especially in cases where an evaluation is seen as threatening or risky. Finally, the question of commitment is certainly linked to the rewards provided in the incentive system of an agency for the individuals who actively pursue and engage in joint evaluation work.

Main conclusions and recommendations

- There is little systematic knowledge on follow-up. There is no common agreement yet on the kind of follow up that is suitable and appropriate for joint evaluations.
- All stakeholders stress that the evaluation is a learning process that will have an impact on the way business is done in future. Some also stress that there are examples of joint evaluations which, although still ongoing, have already had a noticeable impact on the policies and practices under review. Institutional evaluations are most cited in this context.
- Each bilateral agency has its own way of dealing with the outcomes of joint evaluations and the follow-up to it. Some give joint evaluations the same treatment as their own evaluations, and even establish action plans for the implementation of the recommendations. However, many agencies proceed more cautiously, on a case-by-case basis. This approach can potentially lead to the decision to file the final report away rather than to support its recommendations.
- A key issue with regard to follow-up is the level of commitment of the agencies to the joint evaluations they join, including to their results and recommendations. The level of this commitment should be articulated early on in the evaluation process so that everyone knows how far the partners are prepared to go in their commitment.

CHAPTER 3: OPTIONS FOR THE FUTURE – A CRUCIAL CHALLENGE FOR THE DAC

214. “In four short years,” writes the Secretary-General of the United Nations, Kofi Annan, in his Report on the Implementation of the United Nations Millennium Declaration,²³ “the eight Millennium Development Goals derived from the Millennium Declaration have transformed the face of global development co-operation. The broad global consensus around a set of clear, measurable and time-bound goals has generated unprecedented, coordinated action, not only within the UN system, including the Bretton Woods institutions, but also within the wider donor community, and, most importantly, within developing countries themselves.”

215. It is against this background that the DAC and its members have embarked upon a system-wide reform of aid principles, strategies and modalities to support the Millennium Declaration and the MDGs. Priorities have been to remove obstacles to country ownership of the development process and to the broad participation of all stakeholders, to harmonize donor approaches and procedures and to align development programmes with country-led planning. The Rome and Paris Declarations are the key donor commitments to this ‘aid effectiveness’ agenda.

216. The international community has also agreed the importance of monitoring the various indicators that enable us to measure progress towards the Millennium Development Goals and the Paris Declaration on Aid Effectiveness. Moreover, new modes of assistance such as sector-wide approaches, general budget support and other collaborative multi-donor programmes are creating a growing need for joint work in measuring and monitoring of implementation. The same logic – that joined-up development assistance is best reviewed jointly – should also apply to undertaking effective evaluations of the new aid modalities.

217. Thus, one might expect a mushrooming of new multi-partner evaluations - focused on the new aid modalities – brokered and organised within groupings such as the DAC Evaluation Network or the UN Evaluation Group or the Evaluation Coordination Group of the IFIs. However, apart from the ongoing GBS evaluation led by the UK, and a few other examples including the UNDAF evaluation of UNEG members and evaluation work on the PRS process, there is only limited progress towards taking up the challenge of evaluating the new development paradigms.

218. When it comes to the role of evaluation in assessing and promoting the DACs efforts in harmonisation, alignment and aid effectiveness, there are also few signs of concrete progress by DAC member evaluation units. There are no clear indications that DAC donors are planning to combine their evaluation capacities to identify and redress bottlenecks and impediments to the implementation of the Rome and Paris Declarations. Likewise, there are no known plans to critically evaluate the implementation of the Rome Declaration, which so far has only been subject to a self-assessment of the activities and achievements in the group of pilot countries. This assessment was carried out by policy and operational staff of donors.

219. Members of the DAC Evaluation Network are strongly urged to take up these challenges and to ensure that the evaluation community prioritises workstreams relevant to the core developments in the DAC. This report identifies a range of challenges for joint evaluations that will need to be addressed in

23. UN Document A/59/282 of 27 August 2004.

order to maintain evaluation's long-standing and traditionally important position and focus within the overall work programme of the DAC.

220. In the following sections, challenges and options for the future are laid out in three areas of multi-partner evaluation work that are of key importance, namely:

- Improving the existing practice of multi-partner evaluations;
- Enhancing developing country involvement and ownership; and
- Focusing multi-partner evaluation work in the DAC.

3.1 Improving the existing practice of multi-partner evaluations

221. Multi-partner evaluations are a continuous process of experimenting with different configurations of actors, changing thematic and methodological challenges, and varying modalities of organising the common work. This variety generates relatively large numbers of lessons learned. However, it is sometimes difficult to apply these lessons to other evaluations because they emerge from specific activities and circumstances. Nevertheless, DAC members and other interlocutors have suggested a significant number of ideas, practical proposals and lessons learned on improving multi-partner evaluations, as detailed in Chapter 2.

222. The following section outlines some underlying trends and notions, which will help the DAC community to chart the way forward:

- The planning process for multi-partner evaluations is crucial for the success of the exercise. This needs to be improved and allocated the time necessary to agree among the partners a full understanding of the purpose, objectives, ground rules, key evaluation questions, terms of reference, procurement modalities for consultancy services, governance structure and so on.
- The degree of commitment to the findings and recommendations of a multi-partner evaluation will determine its weight and relevance. Therefore, questions related to commitment must be discussed and agreed at the earliest stage.
- Dissemination, follow-up and feedback are essential if an evaluation is to be practicably useful. These aspects need to be given full attention from the beginning of the evaluation process.
- Multi-partner evaluations should be built on mutual trust and confidence – not on tight control, process and systems. The present trend towards the bureaucratisation of multi-partner evaluations through slow and heavy procedures for decision making, drafting and approving of products, managing the consultants, and so on needs to be urgently stopped and reversed. This trend is already beginning to discourage DAC members from joining multi-partner evaluations.
- The decentralisation of decision-making authority to agency Country Offices suggests that increasingly the idea to undertake a joint evaluation will be taken in the field. This can lead to friction, including sometimes a perceived lack of sufficient independence, if not properly coordinated with central evaluation units. Communication between headquarters and the field on planned and ongoing evaluation work has to be strengthened in order to avoid overlap and clashes.

3.2 Enhancing developing country involvement and ownership

223. The Nairobi workshop underlined the urgent need for donors to enable a stronger degree of participation and ownership by developing countries of multi-partner evaluations. In return, the developing country participants pledged their willingness to take a more proactive stance in conceptualising, requesting, initiating and leading joint evaluations. Participants also agreed the need to prepare their own governments for joint evaluation work, for example through better and more efficient interdepartmental communication, coordination and collaboration, through developing common evaluation policies and guidelines, through increased efforts to contribute to the funding of joint evaluations – even though this may initially be on a small scale only - and through early information sharing and planning. We hope that these lessons will be taken up by other developing country governments.

224. DAC members have many opportunities to encourage and facilitate more joint evaluation work. They include the following steps which would help developing countries get ready for more joint evaluations, both in the short and longer term perspectives:

- Donors should support and contribute to a mutual exchange of information with developing country partners, on a regular and systematic basis, about which evaluations are being planned by recipient governments, aid agency headquarters and/or field offices. Information collated should also encompass the work of multilateral, NGO and other development actors.
- Donor consultations and negotiations with developing countries should make space for a regular agenda item on joint evaluations. These discussions should review work accomplished, ongoing activities and plans for future work.
- Donors should assist developing countries to hire local consultants to participate in joint evaluations. Funds could be made available by adapting donor financed study funds or through other mechanisms.
- Recipient-donor coordination groups need to ensure better differentiation between monitoring, assessment, review and evaluation - with a view to strengthening the evaluation component.
- Support should be increased for evaluation capacity development in partner countries, for governments and for the non-governmental sector (parliaments, consultancies, NGOs, academic and research institutions, private sector). This would consist of light forms of capacity development, such as the provision of stipends for training courses existing in the country or region or overseas, places for internships (for instance to observe the process of a joint evaluation in another country), study visits, attendance at seminars, and so on.
- However, evaluation capacity development should also include longer-term activities, more focused and targeted on building institutional capacities. These could include an evaluation capacity development module in technical assistance programmes in the area of democracy, good governance and decentralisation, assistance to establishing evaluation units in government departments or elsewhere, help with the drafting of evaluation policy and legislation and/or guidelines, and similar.

225. Not all of these activities will fall directly under the mandate of the evaluation units of DAC members. In these cases, evaluation units are urged to proactively work with and through their colleagues in policy, programme, operational or sector departments to help make available the necessary support to evaluation capacity development and to creating the conditions in partner countries conducive for more joint evaluations.

226. Evaluation units do have full control over the way a multi-partner evaluation is organised and managed. If the decision is taken to invite developing countries to join a multi-partner evaluation, their participation must be from the very beginning of the process - and not after initial key decisions (terms of reference, selection of consultants and similar) have been taken. It is also important that developing country partners are represented not only on the steering committee, but also on the management group - and that both groups meet in donor as well as in developing countries to perform their work. This is an important psychological aspect and has a strong link to evaluation capacity development.²⁴

227. If these recommendations are implemented for multi-partner evaluations, significant steps will have been made to meet the demands and aspirations of developing countries, summarized by one of the workshop participants in Nairobi as: "Evaluations have to be demystified, democratised and simplified."

3.3 Focusing multi-partner evaluation work in the DAC

228. Without the work and efforts of the DAC, and especially of its Evaluation Network, joint evaluations would not have become as established as they are today. The DAC remains the obvious choice for the forum in which to carry forward the debate on joint evaluations, particularly with a view to adapting to the new development paradigms and the emerging challenges.

229. During the consultations for this study, all interlocutors agreed that there are important issues and challenges in joint evaluations that must be taken up and addressed by the DAC. These issues are outlined below:

- The new development paradigms – MDGs, PRSPs, harmonisation, alignment, aid effectiveness, and so on – provide a strong case for joint evaluation work. However, some donor evaluation units have been hesitant to take this agenda forward. Therefore, the Evaluation Network as well as the DAC itself must take up and move forward this agenda. A number of questions must be urgently addressed: Will there be more joint evaluations in future – in response to the needs arising from the new aid paradigms? What would this imply for traditional evaluation work? Or will the number of joint evaluations remain stable, but with a new focus on different subjects and with enhanced emphasis and thrust?
- Is there a need, as many think, for a DAC role in identifying priority areas and subjects and coordinating joint evaluations? If so, at what level of the DAC should this be taken forward? Who would make the proposals - and who would approve them? Would such a role for the DAC entail the risk of politicising decisions on joint evaluations? How could any risk of impinging on the independence and impartiality of evaluation units be avoided?
- How can a more broad-based constituency in support of joint evaluations be built among a wider range of DAC members - to ensure that the burden of joint work is shared more equitably among the DAC members?
- Should the DAC continue to deal with the whole range of questions connected to joint evaluation work? Or should it concentrate on fewer subjects, for instance on those linked to the new development paradigms, and perhaps a few other subjects of crucial importance (such as quality standards or the evaluation of development effectiveness), and leave the rest to individual members or other donor groupings?

²⁴ In a room documented submitted by Denmark to the 34th meeting of the DAC Working Party on Aid Evaluation in May 2003 on the lessons learned from Managing the Joint Evaluation of Ghana's Road Sub-Sector Programme, it is even stated that Steering Committee meetings should preferably be held in the partner country.

- How can the risk of duplication of evaluation work, including in joint evaluations, between different donor groupings (DAC, ECG, UNEG, EU, Nordics, and Utstein etc.) be reduced? What role should the DAC play in better networking between these groups - with a view to attaining synergies and value added?
- Should the DAC play a role in encouraging members to use multi-partner evaluations to experiment with new forms of evaluation work, such as impact and ex post evaluations, longitudinal studies, and others?
- Should the Development Cooperation Directorate of OECD, on behalf of the DAC, play a focal role in collecting data on joint evaluations, maintain an inventory of them, provide information on lessons learned and good practice, and become a clearing-house and institutional memory for the donor community for joint evaluations? This would need funding by DAC members.
- Should the DAC agree to the compilation and publication of another manual on how to organise and run a multi-partner evaluation, based on the lessons learned and good practices contained in this report, including the aspirations of partner countries?
- Should the DAC commission short technical papers on specific issues related to joint evaluation work - for example on the legal questions involved in establishing financing pools; on minimum requirements for consultant contracts; on different options for bidding procedures for consultancy services; on the assessment of bidding proposals in an effective and transparent fashion; etc?
- Finally, there is the question of the DAC Development Evaluation Network assuming a more proactive role in the planning and implementation of Meta evaluations. Under DAC guidance and supervision, these evaluations would bring together dispersed evaluation knowledge, validate it and feed it into planned or ongoing international processes. Meta evaluations would also help to identify areas of sub-optimal evaluation coverage - perhaps, for example, in the proliferation of individual country strategy and programme evaluations; which are largely disconnected from each other and risk losing sight of the common aid effort.

230. During the extensive consultations for the preparation of this report, one question surfaced constantly: Is there a future for joint evaluation work - and where will this future lie? The answer is yes - there is an important future for multi-partner and joint evaluations. However for this future to be realised, DAC donors (and other evaluation stakeholders) must streamline evaluation to ensure its central place in development cooperation. To meet today's challenges, the evaluation community must become more proactive, be responsive to the new modalities in development cooperation, more participatory and open to developing country ownership, and more accountable in its role and purpose as a crucial element in the global effort to fight poverty and realise the MDGs.

ANNEX 1 JOINT EVALUATIONS BY FOCUS/SCOPE SINCE 1990

Evaluation	Participants ¹	Initiative/Reason	Remarks
1.0 GLOBAL POLICY EVALUATIONS			
1.1 Assessment of DAC Members' Women in Development Policies and Programmes, 1994	All DAC members	WID Working Party together with Working Party on Aid Evaluation	Meta evaluation
1.2 The International Response to Conflict and Genocide: Lessons from the Rwanda Experience, 1996²	Australia, Austria, Belgium, Canada, Denmark, Finland, Germany, Ireland, Italy, Japan, Luxembourg, Netherlands, New Zealand, Norway, Spain, Sweden, Switzerland, United Kingdom, United States of America, Commission of the EU, OECD/DAC Secretariat, IOM, UN/DHA, UNDP, UNHCR, UNHCR, UNICEF, WFP, WHO, IBRD, ICRC, IFRC, ICVA, Doctors of the World, INTERACTION, Steering Committee for Humanitarian Response, VOICE. ³	Denmark; in view of the magnitude and complexity of the Rwanda emergency, it was felt that a thorough analysis of experiences through an evaluation process was urgently required, not least to identify the lessons learned and to develop good practices to be applied in future	Until recently, this was the largest joint evaluation ever undertaken. The five volume report of this evaluation had a major impact on the political debate of the Rwanda case and far-reaching repercussions for humanitarian assistance. These included the Sphere Project to develop minimum technical standards for a range of key intervention areas in an emergency and the formation of the Active Learning Network for Accountability and Performance in Humanitarian Action (ALNAP)
1.3 Evaluation of Programs Promoting Participatory Development and Good Governance, 1997	All DAC members	Ad hoc Working Party on Participatory Development and Good Governance	Meta evaluation
1.4 Joint Multi-Donor Evaluation of European Union Aid Managed by the Commission, 1999	All 15 EU members with <i>France</i> as the chair of the Evaluation Working Group (EWG = Steering Committee), the European Commission and EIB (partly)	Council of Development Ministers; assessment of all aspects of the EU development co-operation programme and its major components (Lomé; ALA; MEDA; ECHO)	A very large and occasionally cumbersome joint evaluation of EU members and the European Commission that went on for about five years (1995 – 1999) and created high transaction costs for everyone involved, with 29 meetings of the EWG alone to discuss altogether 50 different drafts by consultants

<p>1.5 Toward Country-led Development: A Multi-Partner Evaluation of the Comprehensive Development Framework, 2003</p>	<p><i>World Bank (OED and Development Research Group), UNDP, AfDB, UNECA, EU, OECD, Bolivia, Ghana, Pakistan, Romania, Uganda, Canada, Denmark, JBIC, Netherlands, Norway, Sweden, Switzerland, United Kingdom, United States of America, Oxfam</i></p>	<p>Executive Board; OED and DRG merged their individual initiatives to assess and evaluate the CDF into this joint initiative of the Bank. This reflected the high profile that CDF had gained in the World Bank as a key policy initiative taken by President J. Wolfensohn</p>	<p>A relatively heavy governance and management structure due to the need to accommodate a large number of participants in this evaluation. Relatively high transaction costs. A great part of these costs were funded through a Multi-Donor Trust Fund, established by Canada, Denmark, Sweden and the United Kingdom for this evaluation, and from existing individual donor trust funds (Netherlands, Norway and Switzerland). A serious effort to include the partners in governance, management and substance of this evaluation. The report could probably not develop its full impact everywhere because at the time of its publication CDF had already been overshadowed by the PRSP approach to partner-owned development</p>
<p>1.6 Joint Utstein Study on Peace Building, 2004</p>	<p>Utstein countries Germany, Netherlands, Norway, United Kingdom</p>	<p>Norway; the purpose of the evaluation was the analysis of the four countries' peace building initiatives and activities in 13 countries with particular reference to the underlying strategic concepts and the degree to which specific activities were embedded in them</p>	<p>This evaluation had an interesting management structure in that the Oslo Peace Research Institute (PRIO) as the lead consultant did not do the fact-finding and analytical work in the four countries itself. In each of the four countries, decentralised consultants were commissioned to collect and analyse the relevant material and to produce a national input for PRIO that could be used for the meta analysis. While this approach created some problems with regard to the timely delivery, contents, and comparability of the four inputs, it did produce country reports with a strong individual flavour and with interesting insights that might not have been</p>

<p>1.7 Evaluating and Learning about Coherence, Coordination and Complementarity in the European Union's development co-operation policies and operations (so-called Triple C Evaluation), ongoing</p>	<p>All members of the EU with <i>Netherlands, Sweden, temporarily the United Kingdom and the European Commission</i> taking the lead and forming the management group. The full Steering Committee was made up by the members of the management group plus France, Germany and Portugal</p>	<p>EU Heads of Evaluation Services; the triple C concept (co-ordination, complementarity, coherence) is a key element of the Maastricht Treaty to guide the Union's development co-operation activities. At the same time, the concept is politically sensitive and operationally difficult to apply and implement. On a number of occasions, EU development ministers demonstrated great interest in understanding better and learning more about this subject matter</p>	<p>obtained had the work been organised in a centralised manner. Also, the idea of this evaluation right from the beginning was to feed into a high level policy seminar with a view to improving the coordination of peace building approaches among Utstein countries</p>
<p>1.8 Evaluation of General Budget Support, ongoing</p>	<p>Australia, <i>Belgium</i>, Canada, Denmark, France, Germany, Ireland, Japan, <i>Netherlands</i>, New Zealand, Norway, Portugal, Spain, <i>Sweden</i>, Switzerland, <i>United Kingdom</i>, United States of America, <i>European Commission</i>, Inter-American Development Bank, International Monetary Fund, OECD, World Bank, Burkina Faso, Malawi, Mozambique, Nicaragua, Rwanda, Uganda, Vietnam</p>	<p>Probably the largest joint evaluation ever undertaken so far and an ambitious effort to come to grips with the crucial questions of substance, methodology and approach in evaluating GBS. The framework developed for this evaluation is also used in several GBS evaluations in other countries (e.g. Tanzania) to work through new GBS programmes</p>	<p>The evaluation had a slow and cumbersome start in 2000/2001. Several efforts of consultants to come to grips with the subject of this evaluation failed, also due to difficulties in understanding and scoping the triple C concept and in designing an evaluation approach satisfactory to the EUHES as well as avoiding the traps of a politically difficult field. Meanwhile on track, as a result of the decision to aim at a lighter and relatively decentralised structure of and an iterative approach to this evaluation</p>

<p>2.0 GLOBAL IMPACT AND EFFECTIVENESS EVALUATIONS</p>			
<p>2.1 EU Programme Food Aid Evaluation, 1997</p>	<p>Pre-1996 EU members (12), European Commission</p>	<p>Council of Development Ministers and EU Heads of Evaluation Services; the evaluation addressed, in a selective rather than comprehensive manner, key issues of programme food aid as a developmental tool.</p>	<p>In 1989, the EU development ministers agreed to undertake joint evaluations in future. Subsequently, EUHES identified food aid as a suitable first subject for such an endeavour. However, it was only in 1992 that they invited proposals for stage one of this evaluation, which then was followed by three more stages (12 rapid evaluations of case study countries; four extended studies of major recipients; synthesis study) through 1996. The evaluation report can be seen as contributing to the shift in international food aid policy, i. e. de-emphasising this form of aid in favour of financial support for food security and targeted emergency and project food assistance</p>
<p>2.2 Independent Evaluation of the Special Programme of Assistance for Africa, 1997</p>	<p><i>World Bank (OED)</i></p>	<p>World Bank; the purpose of this evaluation was to analyse the Special Programme of Assistance for Africa and to assess the extent of its contribution to promoting the adjustment and development in Sub-Saharan Africa</p>	
<p>2.3 UNAIDS Evaluation, 2002</p>	<p>Members of the Joint UN Programme on HIV/AIDS (UNICEF, UNDP, UNFPA, UNESCO, WHO, IBRD, UNODC, ILO) with some bilateral inputs</p>	<p>The objective of this evaluation was a comprehensive external assessment of the effectiveness and impact of UNAIDS after the first five years of its existence. This evaluation was already foreseen at the time of establishing the joint programme. Canada, together with Australia, Norway, Sweden and UK, provided</p>	<p>This evaluation is an interesting case in point for the usefulness of joint efforts among multilateral and bilateral actors, as it was possible for a group of members of the DAC Working Party on Aid Evaluation to help UNAIDS to qualify, in a professional sense, its approach to this evaluation and to agree on some</p>

<p>2.4 PRSP Evaluations of the IMF and the World Bank, 2004</p>	<p>Independent Evaluation Office of the IMF and Operations Evaluation Department of the World Bank</p>	<p>financial support, including to the 3-person Management Support Team.</p> <p>High profile topic for Executive Boards and Management. Therefore, demonstrating results early on, identifying lessons learned and reacting to criticism raised are important components of the motivation for these evaluations</p>	<p>additional elements to be worked into the evaluation framework with a view to reducing some of the weaknesses of the original approach.</p> <p>The two evaluations were carried out and published independently of each other.⁴ However, in the prefaces of the reports, both institutions refer to the work of the other as something conducted “in parallel”.</p>
<p>2.5 UNDAF evaluation, ongoing</p>	<p>UN organisations system-wide</p>	<p>UN Country Teams at the level of heads of agencies; the key objectives of the evaluation include an assessment of the role and relevance of UNDAF in relation to the underlying causes and challenges identified in the CCA and as a reflection of international norms and standards guiding the work of the UN system and adopted by UN member states. Emphasis is placed on outcomes, effectiveness and efficiency as well as sustainability.</p>	<p>This evaluation represents an ambitious attempt to move beyond process analysis to results and impact evaluation. It is foreseen that the UN Evaluation Group plays an important role in this work once heads of agencies have approved the approach and the TOR</p>
<p>3.0 THEMATIC AND SECTORAL EVALUATIONS</p> <p>3.1 Inter-Agency evaluation of the UN humanitarian response in Afghanistan 1998 – 2000, 2001</p>	<p>UNHCR, UNICEF, UNOCHA, WFP</p>	<p>The purpose of this joint work was an evaluation of identifying vulnerable populations and assessing needs in a complex emergency. These were issues of interest to policy and operational staff in all four agencies. It was also decided to concentrate on one country only, Afghanistan, so as to allow more focus and in-depth analysis</p>	<p>This evaluation suffered from a number of problems and obstacles. These included: the lack of a lead agency and strong leadership; insufficient upstream consultation with the field and, as a consequence, a lack of interest in the field offices to see the evaluation implemented; insufficient clarification of concepts, definitions, methodologies etc. early in the process; poor contractor performance and haggling between</p>

<p>3.2 Local Solutions to Global Challenges: Towards Effective Partnership in Basic Education, 2003</p>	<p><i>Canada, Denmark, Germany, Ireland, JICA, Netherlands, Norway, Sweden, United Kingdom, Burkina Faso, Bolivia, Uganda, Zambia, UNESCO, UNICEF, European Commission, World Bank</i></p>	<p>Netherlands; following the World Education Forum held in Dakar in April 2000, the purpose of this evaluation was to develop a strategy for assessing the combined contributions of external support to basic education in elected partner countries in order to draw lessons for policy and programme improvement</p>	<p>the contractor and the inter-agency evaluation management team; time pressure; and so on. Eventually, the evaluation was more or less abandoned and the report neither finalised, nor fully cleared and published.</p>
<p>3.3 Addressing the Reproductive Health Needs and Rights of Young People since ICPD – The Contribution of UNFPA and IPPF, 2004</p>	<p>Denmark, <i>Germany</i>, Netherlands, Norway, United Kingdom, IPPF, UNFPA (Office of Oversight and Evaluation)</p>	<p>Germany; starting from a need to prepare funding decisions, the purpose of this evaluation was to look more deeply into the two organisations' role in implementing one of the less prominent elements of the Cairo Programme of Action, i. e. the reproductive health needs and rights of young people, and to compare the respective performance and comparative advantages of the two organisations. In addition, the evaluation was to assess the degree of complementarity between UNFPA and IPPF in their activities in partner countries</p>	<p>This is one of the major joint evaluations in more recent times that has produced a number of challenging results. The way the participating agencies will take up and implement the recommendations of this evaluation should throw light on the issue of how deep a commitment agencies are willing to make to the results of a joint evaluation. This evaluation also produced new and interesting insights into the involvement of partners in the evaluation process.</p> <p>This evaluation presents a useful example of how the evaluation process as such can impact on the organisations under review. Moreover, a fairly comprehensive follow-up programme in order to maximise the impact secured the broad dissemination of the results and recommendations of this evaluation within UNFPA and IPPF as well as among member countries and the general public. The co-operation between the Steering Group and the consortium of consultants was occasionally rocky, but did provide useful lessons learned which should be observed in future joint evaluations.</p>

<p>3.4 Lessons Learned on Donor Support to Decentralization and Local Governance, 2004</p>	<p>Members of the DAC Working Party on Aid Evaluation with <i>Germany and UNDP</i> in the lead</p>	<p>Germany, UNDP; the purpose of this meta evaluation under the auspices of the DAC Working Party on Aid Evaluation was to broaden the knowledge base on the role of development co-operation in supporting decentralization and strengthening local governance, beyond the findings and conclusions of the preceding UNDP/Germany evaluation of <i>The Role of UNDP in Decentralization and Local Governance published in 2000</i>. To this end, all relevant evaluation material produced by members of the DAC evaluation group was analysed with a view to distil the lessons learned so far and to identify good practices</p>	<p>This evaluation is a good example to demonstrate of how a relatively limited joint evaluation initiative taken by a small group of partners can prepare the ground for larger and more encompassing follow-up activities with a view to producing guidance for donors. Practical experience with this particular evaluation showed that the language skills of the consultant chosen to do the work are particularly important in the case of meta evaluations which are likely to be dependent on reports written in a number of different languages.</p>
<p>3.5 Evaluation of the Effectiveness and Impact of the Enabling Development Policy of WFP, 2005</p>	<p><i>Canada, Denmark, Finland, France, Germany, Italy, United States of America, WFP</i> (Evaluation Office)</p>	<p>Germany; the purpose of this evaluation was assessment of the impact and effectiveness of the Enabling Development Policy which was strongly influenced by the results and recommendations of the tripartite evaluation of WFP (cf. 4.1 under Institutional Evaluations below) and approved by the Executive Board in 1999. An important part of the background to this evaluation is the declining trend in funding toward the development activities of WFP and the continuing controversial debate on food aid for development</p>	<p>This evaluation was not conceived as the wide-spread top down, institutional analysis approach to evaluating the work of this international organisation, but rather as a study which would be primarily based on empirical evidence collected in the course of seven country case studies. Despite some shortcomings and setbacks, this approach worked quite well. The evaluation also raises a number of interesting issues with regard to the work of a Steering Group that had quite diverse interests in this evaluation</p>

<p>3.6 Evaluation of Assistance to Internally Displaced Persons, ongoing</p>	<p><i>Denmark, Ireland, Netherlands, Sweden, United States of America, ECHO, OCHA, UNHCR</i></p>	<p>EU Humanitarian Assistance Committee, Denmark; assessing the outcomes and impact and identifying the lessons learned and good practices in this area of development assistance</p>	<p>This evaluation is an interesting example of how several agencies can work jointly in a largely decentralised fashion through joint, parallel or single evaluations under a common frame of reference</p>
<p>4.0 INSTITUTIONAL EVALUATIONS</p>			
<p>4.1 Evaluation of UNICEF, 1992</p>	<p>Australia, Canada, Denmark, Switzerland</p>	<p>Canada; full-fledged institutional evaluation to assess the role of the organisation and the results and impact of its work with a view to identify strategic choices for UNICEF.</p>	<p>One of the first evaluations of UN agencies</p>
<p>4.2 Evaluation of the United Nations Population Fund, 1993</p>	<p>Canada, Finland, Germany</p>	<p>Canada; first comprehensive institutional evaluation of UNFPA and its work</p>	<p>see 4.1</p>
<p>4.3 Evaluation of the World Food Programme, 1994</p>	<p>Canada, Netherlands, Norway</p>	<p>Canada, Netherlands, Norway; the purpose of this evaluation was to analyse and assess the efficiency, effectiveness and impact of the WFP and to examine the relevance of key operational objectives considering emerging trends of the 1990s. The final report, it was hoped, could also serve as a valuable input into the formulation of the future policies of the participating agencies as well as into determining their funding commitments.</p>	<p>This evaluation, known as the Tripartite WFP Evaluation, turned out to become one of the relatively influential joint evaluations. The results and many of the recommendations expressed in the report fed into the discussion process in WFP leading to the eventual adoption of the Enabling Development Policy in 1999.</p>
<p>4.4 Global Environment Fund Evaluation, 1994</p>	<p>UNDP, UNEP, World Bank</p>	<p>GEF Council; the objective of this evaluation was an institutional analysis of the pilot phase of GEF, with particular emphasis on how the innovative structure and governance of GEF had worked out in practice</p>	<p>This evaluation was part of the process of preparing the discussions of the replenishment of the GEF</p>

<p>4.5 Evaluation of the UN Institute for Social Development, 1997</p>	<p><i>Nordic Group</i></p>	<p>Denmark, Finland, Norway, Sweden; the purpose of this evaluation was to analyse the work of UNRISD with a view to assessing its efficiency, effectiveness and impact.</p>	<p>This evaluation was initiated and carried out by the group of Nordic countries. It is not known if the Nordics invited any of the other donor agencies supporting UNRISD to participate in this evaluation.</p>
<p>4.6 Evaluation of the United Nations Capital Development Fund, 1999</p>	<p>Belgium, Denmark, France, Japan, <i>Netherlands</i>, Norway, Sweden, Switzerland</p>	<p>Netherlands; evaluation of the work of UNCDF, particularly in the light of and as a follow-up to the 1996 Capacity Assessment of UNCDF.</p>	<p>Classic case of an institutional analysis.</p>
<p>4.7 Independent External Evaluation of IFAD</p>	<p>IFAD member countries, IFAD Executive Board and IFAD Office of Evaluation</p>	<p>The assessment mission of member states helping to prepare the replenishment negotiations, IFAD Executive Board; first full-scale evaluation of IFAD's contribution to rural poverty reduction, the results and impact it has achieved in this area and the relevance of the organization's mission and objectives in relation to international development goals and the national development strategies of IFAD borrowing countries. Canada chaired the Steering Group.</p>	<p>This evaluation had to go through a difficult gestation process, which included a substantive reform of the role of the evaluation function in IFAD, resulting in an independent evaluation office reporting directly to the Executive Board. The governance structure of this evaluation was quite unique and innovative. It placed the overall supervision of the evaluation in the hands of the Director of Evaluation who acted on behalf of the Executive Board. The Steering Committee was composed of representatives of member countries from the different regional groups; it served in an advisory capacity to the Director of the Office of Evaluation and reviewed and endorsed the TOR and the selection of the evaluation team as well as the reports emanating from the Independent External Evaluation.</p>

<p>4.8 Evaluation of ITC, ongoing</p>	<p>Canada, <i>Denmark</i>, Finland, Germany, Norway, Switzerland, United Kingdom, ITC, Algeria, Benin, Egypt, El Salvador, Iran, Senegal</p>	<p>Denmark; first ever evaluation of the work, the results and the impact of the International Trade Centre and of its comparative advantage within the international trade and development community</p>	<p>This evaluation has a strong component of partner country representation in the Core Group (= Steering Committee) which plays an advisory role to the smaller management group made up of sponsoring donors. Altogether, this is a relatively heavy evaluation of a comparatively small international institution. Methodologically, the identification of outcomes and impact in ITC's field of work, mainly training, is particularly difficult</p>
<p>5.0 COUNTRY STRATEGY AND COUNTRY PROGRAMME EVALUATIONS</p> <p>5.1 Country programme evaluation Kazakhstan, 2001</p>	<p><i>World Bank (OED)</i>, MIGA, IFC, EBRD</p>	<p>World Bank; the main thrust of this evaluation was the analysis of how the three institutions of the World Bank Group helped to facilitate the development of the private sector.</p>	<p>This evaluation was a joint exercise within the World Bank family, bringing together the key players in the field of private sector development. The three institutions fielded a joint mission to Kazakhstan. Despite some difficulties in reaching agreement on a common methodology and how to report findings and recommendations, this collaboration led to in-depth learning about the roles of each institution in developing the private sector and created a significant potential for synergies. The participation of EBRD in this exercise was limited to specific inputs because a full involvement of the bank was not possible due to differences in disclosure requirements and the impossibility for EBRD to add new evaluation work midstream in their work programme approved by the EBRD Board</p>

<p>5.2 Country Strategy Review Papua New Guinea, 2002</p>	<p><i>New Zealand, Australia</i></p>	<p>New Zealand; the purpose of this evaluation was to review the relevance, effectiveness and impact of the NZAid country strategy and programme developed for and implemented in Papua New Guinea. AusAID as one of the key players in the development co-operation field in Papua New Guinea was invited to join</p>	<p>This joint evaluation is a good example of a rational decision-making process of why and when to join with others in evaluation work. In this specific case, there was a strong commonality of interest in the partner country as well as in the region</p>
<p>5.3 Country programme evaluation Lesotho, 2002</p>	<p><i>World Bank (OED), African Development Bank (Evaluation Office)</i></p>	<p>African Development Bank; the purpose of this evaluation was not only a full assessment of the outcomes and impact of the country programmes of the two organisations, but also to provide a practical training experience on country programme evaluations for the Evaluation Office of the AfDB</p>	<p>This was the first ever country evaluation of the AfDB. As it was impossible for the two teams to spend the full time of their field work together in Lesotho, arrangements were made for them to overlap for a few days which. This time was used for coordinated meetings with the government, key donors, NGOs and civil society. Again, the chapters on Lesotho's economic and social development and the developmental challenges facing the country were prepared jointly. The rest of the work was done separately. On the cost side, the main costs for AfDB and OED were devoted to coordination, to exchange of work plans, and to time needed to prepare comments.</p>
<p>5.4 Country programme evaluation Peru, 2003</p>	<p><i>World Bank (OED), Inter-American Development Bank (Office of Oversight and Evaluation)</i></p>	<p>World Bank, Inter-American Development Bank; the decision to evaluate the country assistance strategy (World Bank) and the country programme (IADB) was taken as part of the regular programme cycle</p>	<p>Both the World Bank and the IADB prepared their country programme evaluations in Peru in parallel and then agreed to work closely together. This decision was particularly important for the World Bank, as the IADB had made slightly larger commitments to Peru over the period 1992 to 2000. The two evaluations covered the respective bank's assistance programmes. Close</p>

<p>5.5 Country programme evaluation Rwanda, 2004</p>	<p><i>African Development Bank (Evaluation Office), World Bank (OED)</i></p>	<p>AfDB; the purpose of this evaluation was the full assessment of the relevance, outcomes and impact of the country programmes of the two organisations. For the AfDB which took the lead on this evaluation, it was also the application of their previous experience with the Lesotho evaluation</p>	<p>cooperation in implementing the evaluation included the attachment of one staff member of IADB's Office of Oversight and Evaluation to the World Bank mission to Peru, and the inclusion of the Executive Summary of the IADB evaluation in the World Bank evaluation report. One key obstacle to producing a truly joint World Bank-IADB evaluation is that OED rates outcomes explicitly while OVE does not</p>
<p>5.6 Country programme evaluation Jordan, 2004</p>	<p><i>World Bank (OED), Islamic Development Bank (Operations Evaluation Office)</i></p>	<p>Islamic Development Bank; the purpose of this evaluation was not only a full assessment of the outcomes and impact of the country programmes of the two organisations, but also to provide a practical training</p>	<p>Again, the evaluations were produced both in parallel (for the respective country programmes) and jointly (Rwanda's economic and social development and the key developmental challenges facing Rwanda). In addition, a descriptive chapter on the two banks' assistance to Rwanda and the nature of their cooperation was done jointly. This particular evaluation also raised some important issues to be kept in mind when planning for joint country evaluations: (i) how should political issues be handled? (ii) how should differences in operational policies be taken into account, e. g. on post-conflict assistance?, and (iii) which donors should be interviewed during the joint country missions if there are several of them?</p> <p>The reports of this evaluation, published in parallel under the covers of each of the participating institutions, are only partly the result of joint work. The first two chapters containing the general analysis of</p>

		<p>experience on country programme evaluations for the Evaluation Office of the IsDB</p>	<p>Jordan's socio-economic development and the country's development priorities and constraints are based on joint work. The evaluations of the country programmes of the two institutions as such, however, as well as the conclusions and recommendations for each of the institutions were carried out and presented separately. The learning effects for both partners are rated quite high.</p>
<p>5.7 Country programme evaluation Republic of Tunisia, 2004</p>	<p><i>World Bank (OED), Islamic Development Bank (Operations Evaluation Office)</i></p>	<p>Cf. above</p>	<p>The report of this evaluation reflects very much the procedure described above. Again, the learning effect is rated high, particularly with regard to the sharing of perspectives, lessons learned, methodologies and a better understanding of differing mandates of institutions. In addition, the World Bank points out that joint evaluations of this kind have reduced evaluation costs to recipient countries. On the other hand, partnership in evaluations such as between IBRD and IsDB can impose additional costs on the partnering institutions that need to be factored into the cost of the evaluation.</p>
<p>5.8 Joint Sida and Norad/MFA Review of the Development Cooperation with Malawi, 2004</p>	<p><i>Norway, Sweden</i></p>	<p>Norway, Sweden; Sida and Norad decided to carry out a joint review of their programmes in preparation for their negotiations with the Malawi Government on the future direction of their development cooperation. The focus of the exercise was to be on the review of ongoing programmes, particularly the Joint Programme, and on the Sida/Norad model of co-</p>	<p>This evaluation is a particularly interesting example of how evaluation work can support and strengthen joint activities in programme planning and implementation and, at the same time, contribute towards implementing the international agenda on harmonization and alignment.</p>

		operation. This model under which Norad manages both Swedish and Norwegian funds in Malawi is an innovative response to the Rome Declaration on Harmonization and Alignment.		
6.0 SPECIFIC PROJECT AND PROGRAMME EVALUATIONS				
6.1 Joint Nordic Evaluation of the Beira Port Transport System Programme Mozambique, 1995	<i>Denmark, Finland, Norway, Sweden</i>	Denmark; the purpose of this evaluation was to analyse the relevance of the overall BPTS programme objectives as formulated at the launch of the programme in 1986 and taking into account the new situation following the political changes in the Republic of South Africa; the relevance of the assistance provided by the Nordic countries; and the sustainability of the BPTS, also in a regional context		This programme is the largest commitment of the four Nordic countries in Mozambique and possibly the largest infrastructure investment in the country since independence. Altogether, there are 16 donors and four financing agencies supporting BPTS. However, the evaluation covered the 28 distinct projects in several sectors only which were supported by the Nordic countries.
6.2 Joint Nordic Evaluation of the Nordic Africa Institute, 1997	<i>Nordic Group</i>	Denmark, Finland, Norway, Sweden; institutional analysis and impact study		
6.3 The UNDP Role in Decentralization and Local Governance, 2000	UNDP, Germany	UNDP; the purpose of this joint effort was an assessment of the work of UNDP in the field of fostering decentralization and strengthening local governance		This evaluation represents a joint effort of UNDP as a multilateral organisation and Germany as a bilateral donor with broad experience in development co-operation for promoting decentralisation and local governance, to assess the quality, outcomes and impact of UNDP activities in this field and to draw lessons learned. As a matter of fact, the results of this evaluation , when they were presented at a joint dissemination workshop in Berlin, led to the proposal to the DAC Working Party on Aid Evaluation to sponsor a meta evaluation of experiences with support to decentralization and local

<p>6.4 Evaluation of Water Project Dakar III, Senegal, 2003</p>	<p>Kreditanstalt für Wiederaufbau (KfW), Agence Française de Développement (AFD)</p>	<p>KfW; the purpose of this joint evaluation of the contributions of Germany and France to this third water sector programme in Dakar, implemented on their behalf by KfW and AFD, was the analysis and assessment of the relevance, effectiveness, impact and sustainability of the external support</p>	<p>governance (cf. 3.4 above) This is one of the very rare examples of two development cooperation implementing agencies taking the initiative of jointly evaluating their respective contributions to a specific programme in the water sector which they support together with other donors such as the IBRD. The key question for this evaluation was to identify the opportunities and constraints of this sector programme in implementing a privatization strategy in the water supply sector of the country</p>
<p>6.5 Evaluation of Support to African Schools of Statistics, 2003</p>	<p>France, European Commission</p>	<p>France; the purpose of this evaluation was to assess the effectiveness, outcomes, impact and sustainability of the support by the two donors to African schools of statistics in Abidjan, Dakar and Yaoundé under the COMSTAT project of training African statisticians</p>	<p>Classical joint evaluation of the two key donors in the training of African statisticians programmes COMSTAT. As the different aid components supplied by the two donors complement each other, an individual evaluation by each of the donors would not make much sense</p>
<p>6.6 The Role of UNDP in Promoting Governance, ongoing</p> <p>7.0 JOINT EVALUATIONS WITH PARTNER COUNTRIES</p> <p>7.1 Country programme evaluation Burkina Faso, 2000</p>	<p>UNDP, Netherlands</p> <p>Government of Burkina Faso, World Bank</p>	<p>Country assistance evaluation as part of the regular programme cycle</p>	<p>Joint evaluation between the World Bank and the government of Burkina Faso with a view to involving the partner government more directly in assessing the outcomes and impact of the country programme and, at the same time, helping them to develop evaluation capacity of their own.</p>
<p>7.2 Country programme evaluation Tanzania, 2000</p>	<p>World Bank, Government of Tanzania</p>	<p>Country assistance evaluation as part of the regular programme cycle</p>	<p>Joint evaluation between the World Bank and the government of Tanzania with a view to involving the partner government more directly in assessing</p>

<p>7.3 Evaluation of the Road Sub-Sector Programme in Ghana, 2000</p>	<p><i>Denmark, Germany, Japan, Netherlands, United Kingdom, African Development Bank, European Commission, World Bank, Ghana</i></p>	<p>Denmark; the purpose of this evaluation was to analyse the results and impact of the Ghana road sub-sector programme and of the external support to it with a view to obtain lessons learned to feed into the preparation of the New Road Sector Development Programme</p>	<p>the outcomes and impact of the country programme and, at the same time, helping them to develop evaluation capacity of their own.</p> <p>This evaluation was characterised by the strong interest the Ghana government took in it and the lead it assumed throughout the evaluation process. The results of the evaluation strengthened the continuation of the existing donor cooperation and collaboration with the government in the road sector of the country. A follow-up study of the joint evaluation to assess progress made with the implementation of recommendations and to identify constraining factors preventing such implementation was carried out in 2002. It reconfirmed the positive feedback, both from the government of Ghana and especially from the donors, that the joint evaluation had received before.</p>
<p>7.4 PROAGRI evaluation Mozambique, 2003</p>	<p>Mozambique (Ministry of Agriculture and Rural Development),</p>	<p>Mozambique; this joint evaluation was initiated at partner country level by the Ministry of Agriculture and Rural Development (MADER) and agreed upon jointly between the government and the donors in the PROAGRI sector co-ordination group. Its purpose was to evaluate the degree of progress of PROAGRI towards creating an environment conducive to the development of the agricultural sector in the medium and long term, and to propose pertinent adjustments for the remaining period of the first phase of PROAGRI as well as recommendations for the formulation</p>	<p>PROAGRI has been established as a government – donor mechanism to pool support and funding for the agricultural sector and its medium and long term development. Its work is based on institutional capacity development in MADER, the setting up of appropriate financial management systems, decentralisation and ownership. Most donors have meanwhile joined PROAGRI, and very few financing agencies continue to earmark their funding for specific purposes. The initiative for this joint evaluation, taken at the country level, is very much in line with the logic of</p>

<p>7.5 Evaluation of French Aid to Chad, 2003</p>	<p>France, Government of Chad</p>	<p>of the second phase</p> <p>France; joint evaluation of overall country programme by relevance, effectiveness, outcomes, impact and sustainability.</p>	<p>the PROAGRI sector wide support mechanism</p> <p>A particularly interesting example of a broad policy-based evaluation, which did not only look at development cooperation but also at military cooperation, both over the period 1992 to 2005</p>
<p>7.6 Evaluation of the Canadian – South African Development Cooperation Programme, 2003</p>	<p><i>Government of South Africa, Canada</i></p>	<p>South Africa (National Treasury); the purpose of this evaluation was to analyse and assess the programme of Canadian – South African Cooperation for Development with a view to providing a basis for the continuation and further shaping of this cooperation</p>	<p>This is one of the series of evaluations of the cooperation between the South African Government and individual donors, that the National Treasury of South Africa initiated and led to obtain a thorough analysis and assessment of these cooperation programmes and also recommendations on how to develop them in future</p>
<p>7.7 Country programme evaluation Eritrea, 2004</p>	<p>World Bank, Government of Eritrea</p>	<p>Country assistance evaluation as part of the regular programme cycle</p>	<p>The government of Eritrea has assumed strong leadership in all areas of development assistance with donors. It was, therefore, an obvious candidate for a joint country programme evaluation. To this end, the National Statistics and Evaluation Office in the President's Office was designated as partner agency for OED. This office is responsible for all programme and project evaluations undertaken by the government. Most of these evaluations are subcontracted to the University of Asmara. Thus the government would like to use joint evaluations also as a way to develop evaluation capacity in the country, both within and outside of the government.</p>
<p>7.8 Evaluation of the Response to the 2002-03 Emergency in Ethiopia, 2004</p>	<p>Ethiopia (Disaster Prevention and Preparedness Commission), UN Secretary General's Special Envoy for the Horn of Africa, UN Strategic Disaster Management Team, Development Section of Canadian</p>	<p>Ethiopia, UN; the general purpose of the evaluation was, given the unprecedented scope and magnitude of the 2002-03 emergency, an evaluation of the overall response that would provide the government, at</p>	<p>The composition of the evaluation team as well as of the steering group reflected a strong presence of government and UN officials. This may well put a question mark behind the complete independence of this</p>

	Embassy	<p>all levels, the UN agencies, NGOs, donors and beneficiaries with an opportunity to understand the complexity and root causes of this crisis so as to mitigate further crises in Ethiopia and to improve further future humanitarian performance</p>	<p>evaluation, despite the fact that most of the analysis drew heavily on the findings and recommendations of previous evaluations carried out by bilateral and multilateral donor agencies, NGOs, etc.. On the other hand, the extensive involvement of parts of the Ethiopian government has brought about a strong sense of ownership of the responsible government departments of the conclusions and recommendations of this report and a remarkable commitment to their implementation.</p>
<p>7.9 Joint Government-Donor Evaluation of BPEP II Nepal, 2004</p>	<p><i>Denmark, Finland, Norway, European Commission, World Bank</i></p>	<p>Denmark; the purpose of this evaluation was the analysis and assessment of progress achieved in the second phase of the Nepalese Basic and Primary Education Programme II, supported in the form of basket funding, defined as the Core Investment Programme, by Denmark, Finland, Norway, the European Commission and the World Bank, and by JICA, UNICEF and the Asian Development Bank through direct funding of activities</p>	<p>The purpose of this evaluation, which followed shortly after the multi-donor evaluation of external support to basic education, was primarily to investigate macro-level critical issues related to the overall planning and implementation of the BPEP II and some key micro-level issues related to the involvement of de-central education authorities and communities in the planning and implementation process. In order to link this work to that carried out under the Evaluation of Basic Education, one of the objectives was to review and validate the main findings of that evaluation in relation to the Nepalese context</p>
<p>7.10 Evaluation of the Swiss-South African Development Cooperation Programme 2000 – 2003, 2004</p>	<p><i>Government of South Africa, Switzerland</i></p>	<p>Republic of South Africa (National Treasury); the purpose of this evaluation was to analyse and review the Swiss-South African cooperation programme, which is focused on the three domains of governance, education and land affairs, and to provide the South African Treasury</p>	<p>This evaluation looked only at the SDC-supported development cooperation activities in the Republic of South Africa, although the work of SCD is embedded into a larger cooperation mechanism between the two countries. The evaluation has to be seen in connection with the</p>

<p>7.11 Joint Evaluation of the ORET/MILIEV Programme in China, ongoing</p>	<p>Netherlands, Peoples Republic of China (National Centre for Science and Technology Evaluations)</p>	<p>and the Swiss Development Cooperation with a set of independent, flexible and forward-looking recommendations as to the future of this programme</p> <p>Netherlands; this joint evaluation between the Evaluation Department of the Dutch Ministry of Foreign Affairs and the National Centre for Science and Technology Evaluation in Beijing, China, of the Development and Environment Related Export Transactions Programme of the Netherlands Government is a deliberate effort to promote country-led evaluations that would increase partner country ownership of evaluations and enable developing countries to play a different role in the evaluation of their development policies</p>	<p>upcoming negotiations between the two countries on their future cooperation after 2004. It was also the first joint country evaluation ever for SDC and therefore an important learning exercise for the agency.</p> <p>After a good year of negotiations, the framework agreement for the evaluation, including the TOR, management structure and budget, were approved by both sides. According to this agreement, the main objective of this evaluation is to assess to what extent the ORET/MILIEV programme has fulfilled the policy objectives, needs and priorities of the Netherlands and China. The evaluation will also verify whether the funds have been appropriately and efficiently used. The evaluation is intended to provide information for both the Chinese and the Dutch that could be used to improve the programme, as well as for policy formulation.</p>
<p>7.12 Joint Country Programme Evaluation Tanzania, ongoing</p>	<p><i>Tanzania, Canada</i></p>	<p>Canada, Tanzania; replication of the evaluation format developed for this kind of country programme evaluations by the National Treasury in South Africa</p>	<p>Interesting approach towards broadening the range and numbers of country programme evaluations with strong ownership of the partner country. Worth further and comparative analysis once results from a number of countries are available.</p>

1. Countries/agencies written in italics are in the lead for a particular evaluation
2. Dates refer to the year of publication of the final evaluation report, unless otherwise indicated.
3. France suspended its participation in the Steering Committee in December 1995.
4. Evaluation of the IMFs Role in Poverty Reduction Strategy Papers and the Poverty Reduction and Growth Facility and The Poverty Reduction Strategy Initiative-An Independent Evaluation of the World Bank's Support Through 2003.

ANNEX 2

TERMS OF REFERENCE - JOINT EVALUATIONS: RECENT EXPERIENCES, LESSONS LEARNED AND OPTIONS FOR THE FUTURE

Background and Objectives

231. Since a number of years, the DAC Working-Party on Aid Evaluation (now: DAC Network on Development Evaluation) has been in the lead of promoting joint evaluations as a tool towards increased rationalisation of the process of evaluation, reduced transaction costs for partner countries, improved quality of the work undertaken, and increased weight and legitimacy of the evaluation (cf. Note on Joint Evaluations, prepared by Niels Dabelstein, Denmark, for the meeting of the Working Party on Aid Evaluation on 27-28 March 2003, and Lessons Learned from World Bank Experiences in Joint Evaluation, prepared by Osvaldo Feinstein and Gregory K. Ingram, OED, for the same meeting). Experiences with joint evaluations, involving different bilateral and/or multilateral aid agencies, and a first set of lessons learned were synthesised and presented in “Effective Practices in Conducting a Joint Multi-Donor Evaluation”, published in the Evaluation and Aid Effectiveness Series of the DAC Working Party on Aid Evaluation in 2000. This significantly added value to the efforts of members of the DAC Working Party on Aid Evaluation to promote the idea of joint evaluations. As a result, a number of major joint evaluations were initiated and have been concluded recently, or are under way and close to conclusion.

232. As the body of knowledge about joint evaluations grows rapidly, but is still scattered widely across the donor community, the need to review in a more comprehensive fashion and in a systematic way experience with joint evaluations, including emerging issues and new challenges, becomes more acute. Also, the changing environment for international co-operation for development and new paradigms for development co-operation strategies and modalities, such as the PRS-process or new and innovative forms of aid (SWAps, basket financing, budget support) result in additional challenges for evaluation, imply more, rather than less joint efforts, and therefore increase the urgency to take stock of the current knowledge and of the evidence with joint evaluations. So far, much of the evidence available tends to be anecdotal rather than systematic. Consequently, an in-depth analysis and a rigorous assessment of its findings could contribute to a better understanding of the benefits as well as of the costs of joint evaluations. Moreover, distilling a set of lessons learned could be useful in preparing and implementing joint evaluations in the future; in developing new procedures, processes and formulas for joint evaluations, not least in the area of evaluating the effectiveness and impact of the work of multilateral organisations (cf. Room Document No. 8b on Evaluation of Multilateral Organisations, submitted by Denmark to the 1st meeting of the DAC Network on Development Evaluation in Paris on 15-16 January, 2004); and in identifying new challenges that could help to develop orientations for joint evaluations as the collective effort of the donor community for development grows and requires new and more convincing approaches to demonstrating results and developmental impact (cf. OECD Development Co-operation 2003 Report – Overview by the DAC Chair).

233. Therefore, the DAC Network on Development Evaluation agreed at its meeting in Paris on 15-16 January, 2004, to collectively proceed with a new study on joint evaluations which would build on previous work already in existence, especially on “Effective Practices for Joint Multi-Donor Evaluations”, and which would update and broaden it to incorporate recent experiences and new issues.

Scope of the Study

234. The study on joint evaluations will need to be focused carefully on those issues that are of particular interest for the donor community to be addressed, in order to move the idea of joint evaluations forward. While stock-taking and the drawing of conclusions from the evidence collected would be a key focus of the study, it would be equally important to secure a broad enough emphasis on new and emerging issues, so as to provide early and experience-based guidance on how to deal with new challenges in joint evaluations, and to map out possible ways forward.

235. Key themes of the study would continue to be the benefits of joint evaluations on the one hand, and the costs of them on the other hand. More specifically, subjects that would need to be discussed in the study in some depth could include the following – and it should be noted that this list is illustrative rather than exhaustive:

- Rationalising the process of evaluation through joint work;
- Strengthening the quality and credibility of evaluations through joint efforts;
- Harmonising donor efforts and procedures in the field of evaluation through joint work (Rome agreement);
- Reacting to changing paradigms of international development co-operation and accounting for new and innovative forms of aid through joint evaluations;
- Reviewing the transaction costs for joint evaluations, for partner countries as well as for donors, including the lead country;
- Reviewing and categorising the different forms that joint evaluations may take (e.g. parallel evaluations on the same subject by different donors) and the mechanisms for their delivery, including management and governance structures, with a view to presenting the full range of choices available;
- Identifying early on opportunities for joint evaluations and potential partners in them;
- Strengthening accountability for results through joint evaluations.

236. There are also a number of new and additional questions that have surfaced more recently and might be addressed in a study of this kind. These include:

- Standards for, follow-up to, and dissemination of joint evaluations;
- The interest of partner countries in joint evaluations, including their fuller involvement and recipient leadership in them;
- The selection of and the guidance for consultants in joint evaluations, including the use of consortia of consultants, bidding procedures, and contractual and other legal issues;
- The differences as well as the potential linkages between joint evaluations and other joint activities, such as monitoring, data-collection, research, etc.;

- The link between joint evaluations and the results-based management systems in the individual donor agencies;
- The link between joint evaluations and evaluation capacity development, both in donor and in partner countries.

237. Finally, there are a number of emerging issues that members of the DAC Network on Development Evaluation may wish to see addressed in a study of this kind. Again, in an illustrative and not exhaustive sense, the following issues could be given consideration in this context:

- Broadening the range of partners in joint evaluations, including NGOs, political foundations, the private sector, and regional and sub-regional entities active in development co-operation, such as municipalities or regional authorities;
- Using joint evaluations as leverage to work towards harmonising national accountability requirements for aid money;
- Encouraging implementing agencies to do more joint evaluations together with other implementing agencies, at their respective levels of work;
- Creating a level playing field for donors and other partners of unequal weight in joint evaluations.

Approach and Methodology

238. The approach to the study would be characterised by using existing knowledge as a starting point; updating it; adding recent evidence and experience; and by complementing anecdotal evidence with more systematic analysis. The methodology to be applied for taking stock, analysing it and distilling findings and lessons learned/good practices, would be: desk work to absorb and analyse existing written material; interviews with the key actors in joint evaluations, both in Paris and in selected DAC member states' capitals, as well as in international organisations and from among the consultants' community with broad experience in joint evaluations; possibly, but not necessarily a questionnaire to solicit factual information; and focus group discussions with DAC evaluators and other stakeholders in the fringe of other meetings.

239. Representatives of partner countries are important resource persons for the study, as some of the key issues to be addressed (e. g. the question of transaction costs for partners, the issue of harmonising donor procedures and of reducing administrative burdens on recipient governments) cannot be answered satisfactorily without the involvement of partner country representatives who had some experience with the conduct of joint evaluations, as, for instance, with the Joint Evaluation of External Support to Basic Education in Developing Countries or the joint CDF evaluation. Therefore, it is foreseen to hold both individual consultations with partner country representatives as well as a workshop with a group of them to discuss and validate the main findings and conclusions of this work.

Timing

240. The actual work to complete the study would involve approximately 70 days of consultant time (lead consultant) spread over a longer period of time to allow for the necessary flexibility, particularly during the stock-taking exercise. In addition, provision has been made for up to 25 days of supplementary consultant time for specialised tasks in the context of this work. After the DAC Network on Development Evaluation has approved this work at its meeting in mid-January 2004, and as strong commitment to securing the funding of the study has been expressed by eight members of the Evaluation Network, it is

envisaged to present a first summary of tentative results and of remaining issues at the autumn 2004 meeting of the Network for discussion and validation, and a draft of the full report at the Network's first meeting in 2005.

Management structure

241. It has been suggested to establish a relatively light management structure for this study. There should be a Steering Committee to provide overall guidance for the work, consisting of the OECD-DAC Secretariat and members that have expressed their willingness to support the study actively (Austria, Belgium, Denmark, Germany, Italy, Netherlands, Norway and Sweden). The Steering Committee may wish to establish a small task team to supervise the work and to act as a sounding board for the consultants, as needed. Denmark, Germany and the Secretariat have already expressed their interest to be members of this task team. A few members of the Evaluation Network have indicated their interest in becoming "sleeping partners" in this exercise (Canada, Japan and the United Kingdom). They could be included for information purposes in the electronic consultation process, which should be the primary means of communication among the Steering Committee, the task team and the consultant.

Budget and Finance

242. The main components of the budget for the study consists of: consultant fees; travel, to OECD member capitals and partner countries, including some participation by the Secretariat in these missions; the partner workshop; support costs accruing to the Secretariat; and publication costs (editing and production.). The total budget for work in 2004-05 is 116.000 € and Denmark, the Netherlands, Germany and Austria have all made firm commitments to fund the project.

ANNEX 3

REPORT OF THE WORKSHOP ON JOINT EVALUATIONS

CHALLENGING THE CONVENTIONAL WISDOM - THE VIEW FROM DEVELOPING COUNTRY PARTNERS

Nairobi, 20-21 April 2005

Introduction

The DAC Evaluation Network workshop on ‘Joint Evaluations - Challenging the Conventional Wisdom; the View from Developing Country Partners’, was held in Nairobi from 20-21 April 2005. The Workshop was Chaired by Professor Samuel Wangwe of Tanzania on Day One and by Mr Kwesi Abbey Sam of Ghana on Day Two. Hans Lundgren, Head of OECD/DCD Evaluation Section, served as co-Chair.

Rationale

The DAC asked the Network on Development Evaluation to review and analyse past experiences and options for the future for joint evaluations. A literature review and consultations with over 100 representatives of donor agencies (bilateral and multilateral), civil society, and consultants were undertaken in 2004/05. The Nairobi Workshop constituted a vital stage in this consultation process; and solicited the view from developing country partners. The Workshop had two overall objectives: (1) To review past experience of joint evaluations and to analyse their benefits and challenges; and (2) To develop recommendations on how joint evaluations should be planned, implemented and followed-up for the maximum benefit of all partners. National consultants and representatives of developing country governments and civil society were invited to participate (Annex 2: Participant List).

Context and Background

Joint evaluations have been on the development agenda since the early 1990s. The 1991 *DAC Principles for Evaluation of Development Assistance* state, “joint donor evaluations should be promoted in order to improve understanding of each others’ procedures and approaches and to reduce the administrative burden on recipients”. The principles also underline the importance of involving the aid recipients.

Some, but not all, aid agencies have made significant efforts in delivering joint evaluations. In 1998, the *Review of the DAC Principles for Evaluation of Development Assistance* concluded that the 16 DAC members who had participated in joint evaluations, “found them highly – or, more often occasionally – satisfactory”. The report stressed that joint evaluations “have proven to be satisfactory as they allow first-hand learning from each other, give greater results, facilitate feedback, mobilise knowledge, improve follow-up and save resources”. However, respondents also voiced reasons for concern, namely “higher costs, since [joint evaluations] require more time and resources to assure co-ordination and foster mutual understanding. Hidden agendas, different approaches, too general and diplomatic conclusions as they have to combine different interests, increased complexity and delays and different political objectives, also work against effective joint evaluations”.

In 2000, the DAC Evaluation Network published a guidance booklet; *Effective Practices in Conducting a Joint Multi-Donor Evaluation*. The study currently being undertaken aims to build on and

update this earlier guidance; and to prioritise the perspective from developing country partners. The report, *Joint Evaluations, Recent Experiences, Lessons Learnt, and Options for the Future*, which will integrate the workshop outcomes, will be presented to the DAC Network on Development Evaluation in June 2005 and published thereafter. This work is expected to have significant influence on the way that future evaluations are undertaken.

Workshop Summary – Day One

- 1) Hans Lundgren welcomed all participants to the meeting on behalf of the DAC Evaluation Network and presented an outline of the workshop and its aims and objectives. All participants introduced themselves. A series of short presentations were then made on the benefits and challenges of some past joint evaluations: Juan Carlos Gutierrez of Nicaragua gave a presentation on the ongoing evaluation of General Budget Support; Joyce Mapunjo of Tanzania gave a presentation on the monitoring and evaluation systems in Tanzania; and Oumoul Khayri Ba Tall of Mauritania gave a presentation on the perspective of a national consultant. The meeting then divided into breakout groups, to discuss the benefits and challenges of using joint evaluation approaches, before reporting back and holding a plenary discussion. The key issues raised include:

- 2) Definition of Joint Evaluations

Participants felt that joint evaluations should be defined as any evaluation undertaken with the active participation of more than one agency. A typology was proposed with four categories of joint evaluation: (1) Donor + Donor; more than one donor agency working in partnership; (2) Donor + Partner Country; a donor and a partner country working in partnership; (3) Multi-Donor + Multi-Partner; more than one donor and more than one partner country working in partnership; and (4) Partner + Partner; more than one aid recipient country working in partnership on an evaluation.

Some participants argued that all evaluations should be undertaken with the active participation of the aid recipients while others felt that not all evaluations should be undertaken jointly. However, all agreed that a greater proportion of evaluations should be undertaken jointly than has been the case in the past.

- 3) Benefits

Key benefits of working in partnership on joint evaluations were identified as:

- Increased potential for objective and independent review; as the terms of reference and recommendations are not directed by one sole agency. This can increase the legitimacy of the evaluation. However, there will be less legitimacy where there is not real partnership and the evaluation remains donor driven.
- Joint evaluations provide the means for developing more systematic evaluation processes, and the evaluation process can be as important as the results.
- Cost savings for the developing country partner; as joint evaluations should reduce the overall number of evaluations and country reporting requirements.
- Joint evaluations facilitate mutual learning, sharing of best practice and capacity building. It was noted that capacity building must also take place at the level of institutions.

- Joint evaluations encourage more harmonised and aligned programming, and can enhance coherence and coordination between different development actors.

4) Challenges

The workshop also noted a range of challenges in implementing joint evaluations. It was stressed that joint evaluations must be carefully managed in order not to let the challenges outweigh the benefits. Key challenges of joint working were identified as:

- The larger number of participants increases the chances that competing or conflicting interests will frustrate the evaluation. For example, some partners could have political and/or other agendas that negatively influence the process.
- Development aid which is not implemented with a coordinated, harmonised and/or aligned approach may be difficult to evaluate with a joint approach.
- Risk of increased cost for the funder(s) of the evaluation as a result of large and complex evaluation teams and processes.
- Risk of lengthy evaluation process; as each step needs to be agreed by multiple partners.
- Joint evaluations may tend to become overly reliant on external consultants.
- A low level of commitment and participation, on the part of some stakeholders, may frustrate attempts at joint working.

5) Participants felt that joint evaluations have strong potential to empower developing countries. However, it was felt that when joint donor evaluations exclude developing country partners, they can increase the donor influence and disempower the aid recipients. It was also stressed that when a joint evaluation Steering Committee includes representation from several developing countries, those countries should be facilitated to meet together to coordinate their inputs. It was noted that in the case of the Evaluation of General Budget Support, the developing country representatives had not met without the donors. It was also recommended that Steering Committee meetings should be held in developing countries as well as in donor countries.

6) The workshop agreed that while joint evaluations have most commonly been donor-driven, the modality has the potential to lead to real partnership and country ownership. The experience of Tanzania was outlined as a strong model for national ownership of monitoring systems. The Independent Monitoring Group has played a strong role in coordinating M&E work and in putting the country partners in the driving seat. However, it was noted that full partnership and ownership will not be achieved when all the partner countries do not participate from the outset of the evaluation process and when they are not taking an active role in all stages: agreeing the initial terms of reference, the inception report and the recommendations. Participants stressed that developing countries need to themselves initiate and take the lead on joint evaluations. All agreed that ownership is vitally important and that even heavily aid-dependant countries should demand participation in and ownership of evaluation processes. It was also noted that a joint evaluation can be undertaken when a programme has not been implemented jointly; the evaluation team should be independent of the programme managers and a joint approach can help build both partnership and objectivity.

- 7) The participation of civil society organisations also needs further consideration. It was noted that society has a role to play in demanding government accountability. Evaluations were seen as one way of meeting accountability requirements, but it was also noted that lighter-touch and faster approaches such as PRA and small-scale reviews also have an important role to play.

Workshop Summary – Day Two

- 8) Kwesi Abbey Sam welcomed participants to the second day of the workshop and presented a summary of the first day of the workshop. The second day looked forward to the future, and asked how joint evaluations should be planned, delivered and followed-up for the maximum benefit of all partners.
- 9) A series of short presentations were made on future directions for joint evaluations: Sebastian Ling of the OECD gave a presentation on the context in which the DAC Evaluation Network is undertaking the ongoing study on joint evaluations, including the Paris Declaration commitment for donors to “Harmonise their monitoring and reporting requirements, and, until they can rely more extensively on partner countries’ statistical, monitoring and evaluation systems, with partner countries to the maximum extent possible on joint formats for periodic reporting”. Horst Breier, the report consultant, gave a summary of the findings and recommendations identified so far in the draft report, *‘Joint Evaluations: Recent Experiences, Lessons Learnt and Options for the Future’*; Vu Dai Thang gave a presentation on present and future directions in Vietnam; and Sharmala Naidoo on present and future directions in South Africa. The meeting then divided into breakout groups, to discuss (1) Upstream planning of joint evaluations; (2) Management and Governance of joint evaluations; and (3) Participation and Ownership.

The key issues raised included:

10) Upstream Planning

- The group recommended that all development interventions should have a joint evaluation embedded from the initial design phase. The decision to undertake an evaluation jointly should be made at the initial planning stage of every project or programme. This would increase ownership by developing countries and improve lesson learning and capacity building.
- It was felt that the key stakeholders in the evaluation process should be identified jointly by the donors and the aid recipients.
- It was noted that where donor programmes are harmonised within SWAps and/or are aligned with government planning, especially through GBS, it will be easier to plan and undertake joint evaluations.
- It was recommended that developing countries need to show greater initiative in planning and scheduling which evaluations will be undertaken – a possible tool could be an annual or bi-annual planning matrix coordinated by a central government ministry.
- The group recommended that the following ground rules should be agreed at the outset of every joint evaluation: (1) That the evaluation should be undertaken independently and objectively; (2) That the Steering Committee should have an agreed joint management and decision-making structure and that all partners should share accountability for the evaluation; (3) That the evaluation should have a clear and agreed purpose; and (4) That the ToR,

procurement arrangements, management structure, implementation, timeframe and dissemination policy should all be agreed jointly.

- Countries should review and build on the experience in Vietnam, where the Government has made internal M&E a legal requirement in the Decree on ODA Management.

11) Implementation: Governance and Management

- The group noted that multi-agency joint evaluations will normally need both a larger Steering Committee and a smaller Management Group. Both groups must, however, be of a functional and workable size and should include participation from developing countries. The roles and representation on both committees should be agreed between the key actors.
- In general, the Steering Committee should be responsible for the following areas: defining the scope of the work; agreeing the MoU and ToR; overseeing the evaluation process; approving the budget; selecting and appointing consultants; resolving conflicts; approving reports; and advising respective partners on recommendations and action plans.
- While the Management Group should be responsible for the following areas: managing, supporting and facilitating the evaluation process on a day-to-day basis; preparing draft ToR and other documents for the Steering Committee; providing technical and administrative support to the consultants; and reporting to the Steering Committee on progress and problems. The Management Group should normally be composed of evaluation professionals.
- The role of the consultants should also be agreed up-front. In general, they should be responsible for: implementing the ToR; developing the evaluation criteria; and writing the inception report and the final report and recommendations. It was recommended that local/national consultants should be contracted where possible and that innovative forms of funding should be made available to developing country governments to enable them to themselves contract national consultants.

12) Participation and Ownership

- The group stressed that participation is easier to realise than ownership.
- It was felt that the agency that has the idea for and initiates a joint evaluation is likely to take the initial lead and therefore have the greatest ownership. It was recommended that developing countries must themselves take the lead and initiate more joint evaluations.
- However, the group also noted that sufficient capacity is needed in order to take ownership. It was therefore recommended that the IPDET evaluation training should be expanded and rolled out in a broader range of countries. However, capacity building should not be limited to training of individuals, but must encompass institutional capacity building. Developing country partners may lack capacities in time and resources as well as in technical knowledge. M&E units should therefore be built and developed within partner country governments, possibly within a central ministry or at the Office of the Auditor General. All partners need to look at innovative ways of providing funding for aid recipients to build their own evaluation capacities.
- Strong participation of local consultants can also build national ownership.

- M&E networks and professional associations need to be built and developed within developing countries.
- Some developing countries may find it more practicable to take ownership of evaluations that have a stronger focus on lesson learning than on accountability.
- Procurement rules need to be harmonised within developing countries; e.g. all the donors should agree to a common set of Public Procurement Rules (PPR) and all evaluations should follow that common country guidance. Participants also commented that where aid remains tied, this can reduce the developing country capacity to make spending decisions and take ownership.
- Countries should review and build on the South African experience; where the National Treasury has initiated a series of seven joint evaluations in partnership with different donors and has also led a Development Cooperation Report; evaluating total country ODA from 1994-1999 (presentation attached at Annex 3).

13) Key workshop recommendations

- a) A greater proportion of evaluations should be undertaken jointly; with full and active participation of the aid recipients and other partners from the very outset. Further, developing country partners need to take ownership and must therefore take a more active role in initiating joint evaluations.
- b) Developing countries should show greater initiative in taking the lead in planning, coordinating and scheduling which evaluations will be undertaken – a possible tool could be an annual or bi-annual planning matrix coordinated by a central government ministry.
- c) Developing country governments should be supported to build their institutional capacity for initiating and leading joint evaluations. M&E units should be built and developed within developing country governments. All partners need to look at innovative ways of providing funding for aid recipients to build their evaluation capacity.
- d) Better coordination and knowledge sharing is needed amongst the various partners within aid recipient countries. National M&E networks and professional associations need to be built and expanded.
- e) When a large joint evaluation is undertaken with the participation of several developing countries, the developing countries should be facilitated to meet together to coordinate their views and inputs. Steering Committee meetings should also be held in developing countries as well as in donor countries.
- f) Developing countries should review and build on the Vietnamese experience; where internal M&E has been made a legal requirement in the Decree on ODA Management.
- g) Developing countries should review and build on the South African experience; where the National Treasury has initiated a series of seven joint evaluations in partnership with different donors and has also led a Development Cooperation Report, evaluating total country ODA from 1994-1999.

Workshop Programme

Workshop on Joint Evaluations
Challenging the Conventional Wisdom - the View from Developing Country Partners
 20-21 April 2005, Nairobi, Kenya

	20 April	21 April
AM	<p>DAY ONE: EXPERIENCES OF JOINT EVALUATIONS (LOOKING BACK AND THE PICTURE TODAY)</p> <p>09:00 – 09:30: Opening a. Introduction to the workshop b. Roundtable introductions</p> <p>09:30 – 11:00: Informal Presentations a. The evaluation of General Budget Support b. Experiences in Tanzania c. The view of the national consultant</p> <p>11:00 – 11:30: Tea/Coffee Break</p> <p>11:30 – 13:00: Breakout Sessions a. Benefits and challenges of joint evaluations</p>	<p>DAY TWO: OPTIONS FOR THE FUTURE (LOOKING FORWARD)</p> <p>09:00 – 09:30: Introduction a. Review of Day 1 and introduction to Day 2</p> <p>09:30 – 11:00: Informal Presentations a. The joint evaluations context b. Presentation by joint evaluations consultant c. The direction in Vietnam d. The direction in South Africa</p> <p>11:00 – 11:30: Tea/Coffee Break</p> <p>11:30 – 13:00: Breakout Sessions a. Ways forward and options for the future</p>
	Buffet Lunch (13:00 – 14:00)	Buffet Lunch (13:00 – 14:00)
PM	<p>14:00 – 15:00: Breakout Reporting</p> <p>15:00 – 15:30: Tea/Coffee Break</p> <p>15:30 – 17:00: Plenary a. Plenary discussion on the benefits and challenges of joint evaluations</p>	<p>14:00 – 15:00: Breakout Reporting</p> <p>15:00 – 15:30: Tea/Coffee Break</p> <p>15:30 – 17:00: Plenary a. Plenary discussion on ways forward and options for the future b. Workshop conclusions and next steps</p>
	Informal Dinner (19:30)	

Participants List

NAME	TITLE
Amr Aljowaily	First Secretary, Permanent Mission of Egypt, to the UN in Geneva (Personal Capacity)
Paschal B. Assey	Acting Director, PED, Vice President's Office, Tanzania
Oumoul Khayri Ba Tall	Consultant, Mauritania
Judith Bakirya	Development Advisor, Uganda
Horst Breier	Consultant , Germany
Lars Elle	Deputy Head of Evaluation , Denmark
Juan Carlos Gutierrez	Fiscal Affair Director, Nicaragua
Shafiqul Islam	Joint Secretary, Economic Relations Directory, Ministry of Finance, Bangladesh
Wambui Kimathi	Kenya Commission on Human Rights
Sebastian Ling	OECD/DCD Evaluation Section,
Hans Lundgren	Head, OECD/DCD Evaluation Section,
Joyce Mapunjo	Commissioner, Treasury, Tanzania
Nagaraju	Deputy Secretary, Ministry of Finance, India
Sharmala Naidoo	Director, Project Planning and Institutional Development, Treasury Republic of South Africa
Karen Odhiambo	Director, Kenya Evaluation Network
John Okidi	Executive Director, EPRC, Uganda
Kwesi Abbey Sam	Chairman PPB, Ghana
Vu Dai Thang	Senior Expert, Vietnam
Wilna van Zyl	Senior Policy Analyst, Treasury, Republic of South Africa
Professor Sam Wangwe	Economic and Social Research Foundation, Tanzania
Debazou Yantio	M&E Officer, Cameroon

ANNEX 4 – LIST OF PEOPLE MET

AUSTRIA

Arbeitsgemeinschaft Entwicklungszusammenarbeit (AGEZ)	Elfriede Schachner, Director
Austrian Development Agency (ADA)	Robert Zeiner, Director, Programmes and Projects
Austrian North-South Institute for Development Cooperation	Norman Spitzegger, Director
Care Austria	Peter Franz Kögler, Environment & Development Daniel Seller, Program Director Reinhard Trink, Emergency Coordinator
Federal Ministry for Foreign Affairs (BMAA)	Anton Mair, Dep. Director General, Section VII
Horizont 3000 – Austrian Organisation for Development Co-operation	Gerda Daniel, Directress, Projects and Programs

BELGIUM

ADE Consulting Services S.A.	Anne-Claire Luzot, Socio-Economist
Ministère des Affaires étrangères, Commerce extérieur et Coopération au Développement	Paul Avontroodt, Directorate-General for Development Cooperation Dominique de Crombrughe, Evalueur Spécial Anne-Marie Lambert, Directorate-General for Development Cooperation

CANADA

Baastel	Alain Lafontaine, Vice-President
Canadian International Development Agency (CIDA), Evaluation Division, Performance Review Branch	Goberdhan Singh, Director Françoise Mailhot, Pradip Shastri
Goss Gilroy Inc.	Sheila Dohoo Faure, Managing Partner Ted Freeman, Partner
International Development Research Centre (IDRC)	Fred Carden, Director, Evaluation Unit
The Cornucopia Group, Inc.	Diana McLean, Partner

DENMARK

COWI A/S	Niels Eilshov Olesen, Development Planning
----------	--

	Division
Danish Institute for International Studies	Steen Folke, Senior Researcher, Development Research
Danish Red Cross, International Department	Gitte Gammelgaard, Development Advisor Anders Ladekarl, Head Jytte Roswall, Health Advisor
Danish Refugee Council	Ann Mary Olsen, Deputy Head, International Department
Euro Health Group Consultants	Benedikte Lillebaek, Finance & Administrative Director
Ministry of Foreign Affairs (Danida), Evaluation Department	Niels Dabelstein,, Head Lars Elle, Deputy Head Esther Lønstrup
Nordic Consulting Group A/S	Sven Nilsson, Agro-Economist

EUROPEAN COMMISSION

EuropeAid Co-operation Office, Evaluation Unit	Jean-Louis Chomel, Director True Schedvin Pieter van Steekelenburg
--	--

FRANCE

Agence Française de Développement	Anne-Marie Cabrit, Responsable de la Mission, Direction de la Stratégie
Ministère des Affaires étrangères (DGCID)	Aude de Amorim, Chef du Bureau de l'évaluation, Michel Ruleta, Bureau de l'évaluation
Ministère de l'Economie, des Finances et de l'Industrie	Daniel Kamelgarn, Responsable d'Unité Évaluation, Direction du Trésor, Vice Chair, DAC Network on Development Evaluation

GERMANY

Federal Ministry for Economic Co-operation and Development (Evaluation Division)	Klaus Krämer Achim Mortier Michaela Zintl, Head Lioba Weingaertner, Independent Consultant
Ministry of Foreign Affairs	Eduard Westreicher, Counsellor, Permanent Mission of Germany to the OECD

INTER-AMERICAN DEVELOPMENT BANK

Office of Evaluation and Oversight

Stephen A. Quick, Director

INTERNATIONAL FUND FOR AGRICULTURAL DEVELOPMENT (IFAD)

Office of Evaluation

Luciano Lavizzari, Director

INTERNATIONAL MONETARY FUND

Independent Evaluation Office

Martin D. Kaufman, Senior Economist
Shinji Takagi, Advisor,

UN Representation of IMF

Reinhard Munzberg, Special Representative of the
IMF to the United Nations

INTERNATIONAL PLANNED PARENTHOOD FEDERATION (IPPF)

Med Bouzidi, Deputy Director General

JAPAN

Japan Bank for International Cooperation (Paris
Office)

Monotori Tsuno, Chief Representative
Eigo Azukizawa, Representative

NETHERLANDS

Ministry of Foreign Affairs (Policy and Operations
Evaluation Department/ IOB)

Rob D. van den Berg, Director
Henri E. J. Jorritsma, Deputy Director
Ted J. Kliest, Evaluator
J. Hans Slot, Evaluator
Fred Ph. M. van der Kraaij, Evaluator
Gerard van der Zwan, Evaluator
Marijke Stegeman, Evaluator
Anton R. M. Schutte, Evaluator

DGIS/PM&E

Piet de Lange, Evaluation Development
Cooperation Project Team

Education and Research Department

Ronald Siebes, Cultural Cooperation

Jeroen Verheul, Deputy Permanent Representative
to the OECD, Paris

n(o)vib – Oxfam Netherlands

Yvonne Es, Senior Advisor, Quality and Control
Bureau

PLAN Nederland

Rene Schoenmakers, Evaluation Section

Saltet & van de Putte

Bert van de Putte, Monitoring and Evaluation
Specialist

NORWAY

International Peace Research Institute Oslo (PRIO)

Wenche Hauge, Researcher

Nordic Consulting Group (NCG)	Jens Claussen, Managing Partner Anders Wirak, Sociologist
Norwegian Agency for Development Cooperation (Norad), Evaluation Dept.	Björg Schonhowd Leite, Director Sigurd Endresen, Dep. Director Agnete Erikson, Senior Adviser Tor E. Gjerde, Senior Adviser
Norwegian Institute for Urban and Regional Research (NIBR)	Jon Naustdalslid, Director
Scanteam Analysts and Advisers	Karstein Haarberg, Partner Björn Lunöe, Senior Partner Anne Mossige, Senior Partner Erik Whist, Senior Partner

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD)

Development Assistance Committee	Richard Manning, Chairman
Development Co-operation Directorate (DCD)	Michael Roeskau, Director Richard Carey, Deputy Director
DCD Review and Evaluation Division	Sean Conlin, Principal Analyst Martina Kampmann, Principal Analyst Hans Lundgren, Head, Evaluation Section Sebastian Ling Michelle Weston

SWEDEN

Andante – Tools for Thinking AB	Kim Forss
Rädda Barnen	Alfhild Petrèn
Swedish International Development Cooperation Agency (Department for Evaluation and Internal Audit)	Eva Lithman, Head Stefan Molund, Deputy Head Stefan Dahlgren Susanna Lundström Joakim Molander Eva Lövgren, AFRA (Malawi and Zambia)

SWITZERLAND

Intercooperation	Isabel Dauner Gardiol, Finance-Enterprise-Market
State Secretariat for Economic Affairs (SECO)	Ivo Germann, Balance of Payments Operations and Debt Relief Mukul Kumar, Trade & Clean Technology Cooperation Davorka Rzehak, Investment Promotion

Swiss Agency for Development and Cooperation (SDC)	Peter Meier, Head, Evaluation & Controlling EZA Gerhard Siegfried, Head, Evaluation & Controlling Anne Bichsel, Evaluation & Controlling Samuel Waelty, Evaluation & Controlling Béatrice Ferrari, International Financial Institutions Christoph Graf, Section Asia I
--	---

WORLD BANK

Operations Evaluation Department	Ajay Chibber, Director Osvaldo Feinstein Patrick G. Grasso Martha Ainsworth Alain Barbu Victoria Elliot John Erikson Nils Fostvedt Fareed M. A. Hassan R. Kyle Peters, Jr. Ray Rist Klaus Tilmes
Office of the German Executive Director	Susanne Dorasil

UNITED KINGDOM

ALNAP	John Lakeman, Database & Website Manager
Department for International Development (DFID), Evaluation Department	Mike Hammond, Director Neil McKie, Deputy Project Manager Lynn Quinn Joe Reid, Programme Manager Kate Tench Nick York
Department for International Development (DFID), Human Resources Development	Arthur Fagan
Options	Dana Hovig, Managing Director
Overseas Development Institute (ODI)	Simon Maxwell, Director Kate Bird, Research Fellow
Performance Assessment Resource Centre (PARC)	Achim Engelhardt, Consultant & Information Manager

UNITED NATIONS DEVELOPMENT PROGRAMME (UNDP)

Evaluation Office

Nurul Alam, Deputy Director
Fadzai Gwaradzimba, Senior Evaluation Adviser

UNITED NATIONS CHILDREN FUND (UNICEF)

Evaluation Office

Jean Serge Quesnel, Director
Christian Privat

UNITED NATIONS POPULATION FUND (UNFPA)

Oversight and Evaluation Branch

Henna Ong, Former Director and Chief
Linda Sherry-Cloonan, Dep. Director and Acting
Chief

UNITED STATES OF AMERICA

USAID

Brian A. Frantz, International Economist, PPC/DP
Janet Ellen Kerley, Monitoring and Evaluation
Officer, AFR/DP/POSE
Joseph M. Lieberman, Economist
Cressida Slote, Monitoring and Evaluation
Specialist, E&E/PO
Janice M. Weber, Director, Office of South
American Affairs

ANNEX 5 - BIBLIOGRAPHY AND REFERENCES

- Active Learning Network for Accountability and Performance in Humanitarian Action (ALNAP), *Field Level Learning – ALNAP Review of Humanitarian Action in 2003*, London 2004
- Ashoff, Guido, *Client Survey Study of DAC Peer Reviews*, Report submitted to the DAC, DCD/DAC (2002)28 /ANN1, OECD Paris, 25 October 2002
- Balogun, Paul and DANIDA, Evaluation Department, *A New Approach to Assessing Multilateral Organisations' Evaluation Performance, Approach and Methodology*, Draft Paper, Copenhagen, December 2004
- Ba Tall, Oumoul Khayri, *DAC/OECD Workshop on Joint Evaluations, a View from a Southern Consultant*, Power Point presentation, Nairobi, 20 – 21 April 2005
- Binnendijk, Annette, *Donor Experience with Joint Evaluations – A Typology and Lessons Learned*, working draft submitted by USAID to the 30th meeting of the DAC Working Party on Aid Evaluation, OECD Paris 1998
- Booth, D. and A. Lawson, *Evaluation Framework for General Budget Support*, Report to the Management Group for the Joint Evaluation of General Budget Support, London, May 2004
- Borton, J., Buchanan-Smith, M. and R. Otto, *Learning from Evaluations of Support to Internally Displaced Persons*, Draft IDP Synthesis Report, Ohain, Belgium, 2004
- Borton, J., and J. Eriksson, *Assessment of the Impact and Influence of the 1996 Joint Evaluation of Emergency Assistance to Rwanda*, Copenhagen, June 2004
- Bundesministerium für auswärtige Angelegenheiten, *Dreijahresprogramm 2004 bis 2006 der österreichischen Entwicklungspolitik*, Fortschreibung, Vienna 2003
- Canadian International Development Agency (CIDA), *Review of Lessons from the Process of Evaluation of Multilateral Organisations*, Ottawa, December 1993
- Canadian International Development Agency (CIDA), *Canada making a difference in the world, A Policy Statement on Strengthening Aid Effectiveness*, Ottawa, September 2002
- Canadian International Development Agency (CIDA), *CIDA Evaluation Guide*, Ottawa, January 2004
- Carlsson, J., Erikson-Baaz, M., Fallenius, A. M., Lövgren, E., *Are Evaluations Useful – Cases from Swedish Development Co-operation*, Sida Studies in Evaluation 99/1, Stockholm 1999
- Chen Zhaoying, *Lessons Learned from a Joint Evaluation between Donor and Recipient: the Netherlands' Mixed Credit Programme in China*, Beijing, undated

- Dabelstein, Niels, Evaluating the International Humanitarian System: Rationale, Process and Management of the Joint Evaluation of the International Response to the Rwanda Genocide, in: Disasters, vol. 20, No. 4
- Dabelstein, N., and J. Eriksson, Rwanda and Darfur: Déjà Vu Plus, open letter to ???, Copenhagen and Silver Spring, MD, xxx 2004
- Dabelstein, Niels, Note on Joint Evaluations, Note to the 37th meeting of the DAC Working Party on Aid Evaluation, Room Document No. 2, Paris , March 2003
- DAC Expert Group on Aid Evaluation, Continued Progress in Joint Evaluation, Note by the Delegation of Japan, DCD/DAC/EV(91)10, OECD Paris, 30 September 1991
- DAC Working Party on Aid Evaluation, Review of the DAC Principles for Evaluation of Development Assistance, OECD Paris 1998
- DAC Working Party on Aid Evaluation, Effective Practices in Conducting a Joint Multi-Donor Evaluation, OECD, Paris 2000
- DAC Working Party on Aid Evaluation, Managing the Joint Evaluation of Ghana's Road Sub-Sector Programme: Some Lessons Learned, Note by Denmark, Room Document No. 1, 34th meeting, Paris, 22 – 23 May 2001
- DAC Working Party on Aid Evaluation, Glossary of Key Terms in Evaluation and Results Based Management, OECD, Paris 2002
- DAC, Development Co-operation Directorate, DAC Joint Country Assessment in Tanzania: Approaches to partnership in the aid programmes of Denmark, Finland, Ireland and Japan, DCD/DAC (2003)14/REV1, OECD Paris, 15 October 2003
- Danida, Evaluation Department, Joint Evaluation of the International Trade Centre, Final Terms of Reference, Copenhagen, August 2003
- Danida, Evaluation Department, Protecting Lives and Reducing Human Sufferings, Framework for a Common Approach to Evaluating Assistance to IDPS, Copenhagen, October 2003
- Danida, Evaluation Department, Joint Government-Donor Evaluation of Basic and Primary Education Programme Nepal II, Terms of Reference, Copenhagen, November 2003
- Development Assistance Committee (DAC), Shaping the 21st Century: The Contribution of Development Co-operation, OECD Paris 1996
- Development Assistance Committee (DAC), Development Co-operation 2003 Report, OECD Paris 2004
- Dugger, Celia, World Bank Challenged: Are Poor Really Helped? in: The New York Times, 28 July, 2004
- ECORYS, Follow-up Study of the Joint Evaluation of the Road Sub-Sector Programme Ghana 1996 – 2000, December 2002
- Eriksson, J., and J. Borton, An Assessment of the Impact and Influence of the 1996 Joint Evaluation of Emergency Assistance to Rwanda, draft paper for a Special Issue of "Den Ny Verden", Washington, 2004

- European Centre for Development Policy Management (ECDPM), Terms of Reference for Evaluating and Learning about Coherence, Coordination and Complementarity in the European Union's development co-operation policies and operations, Maastricht, 1 June 2004
- Faure, S., Freeman, T., Kliest, T. and Samoff, J., Consensus, Legitimacy, and Critique in Joint Evaluations, The Hague 2005
- Federal Ministry for Economic Co-operation and Development (BMZ), Addressing the Reproductive Health Needs and Rights of Young People since ICPD – The Contribution of UNFPA and IPPF, Synthesis Report, Bonn, September 2004
- Feinstein, Osvaldo, *Lessons Learned from Experiences in Joint Evaluation*, in: DAC Working Party on Aid Evaluation/Ministère de l'Économie, des Finances et de l'Industrie, *Partners in Development Evaluation: Learning and Accountability*, Workshop proceedings, Paris 2003
- Feinstein, O., and G. K. Ingram, *Lessons Learned from World Bank Experiences in Joint Evaluation*, Room Document No. 3, presented to DAC Working Party on Aid Evaluation, Paris, March 2003
- Forss, Kim, *Cost and benefits of diversity – a note on the culture of evaluation and the evaluation of cultures*, paper submitted to the European Evaluation Society Annual Conference 2002, Seville, October 2002
- Global Environment Facility, *Second Study of GEF's Overall Performance*, Terms of Reference, Washington, D.C. 2000
- Gutierrez, Juan Carlos, *Experiences of Joint Evaluations – Joint Evaluation of General Budget Support*, Power Point presentation, Nairobi, 20 – 21 April 2005
- Hassan, Fareed M. A., *Partnership in Joint Country Assistance Evaluations: A Review of World Bank Experience*, Operations Evaluation Department, The World Bank, Washington, D.C. 2005
- Horton, D., A. Alexaki et al., *Evaluating Capacity Development – Experiences from Research and Development Organizations around the World*, The Hague, Ottawa and Wageningen 2003
- International Fund for Agricultural Development (IFAD), *Independent External Evaluation of IFAD*, Terms of Reference, Rome, 15 July 2003
- International Monetary Fund, Independent Evaluation Office, *Evaluation of the IMF's Role in Poverty Reduction Strategy Papers and the Poverty Reduction and Growth Facility*, Washington, D.C., 2004
- ITAD Ltd. and Oxford Policy Management, *Evaluation du Fonds d'équipement des Nations Unies (FENU)*, Rapport de synthèse, Hassocks, Great Britain, August 1999
- Joint Evaluation Follow-Up Monitoring and Facilitation Network (JEFF), *The Joint Evaluation of Emergency Assistance to Rwanda: A Review of Follow-up and Impact Fifteen Months After Publication*, Copenhagen, 10 June 1997
- Kabell, D., and P. Balogun, *Selected Bilateral Donors' Current and Future Performance Information Needs from Multilateral Organisations: An Assessment*, Holte, Denmark, October 2004
- Kamelgarn, Daniel, *Evaluation et Démocratie: l'Évaluation comme Élément du Mouvement Social*, manuscript, Paris, November 2004

- Killick, Tony, *Politics, Evidence and the New Aid Agenda*, in: Development Policy Review, 2004, 22 (1): 5-29
- Liebenthal, A., O. Feinstein and G. K. Ingram, *Evaluation & Development – The Partnership Dimension*, World Bank Series on Evaluation and Development Vol. 6, New Brunswick, N.J., 2004
- Ling, Sebastian, *The Joint Evaluations Context*, Power Point presentation, Nairobi, 20 – 21 April 2005
- Mapunjo, Joyce K. G., *Joint Evaluation – Benefit and Challenges*, Power Point presentation, Nairobi, 20 – 21 April 2005
- Ministère des Affaires étrangères, Direction générale de la Coopération internationale et du Développement, *Guide de l'Évaluation*, Paris, June 2004
- Ministère des Affaires étrangères, Direction générale de la Coopération internationale et du Développement, *Évaluer la coopération internationale française - Bilan des évaluations 2002*, Paris 2003
- Ministère des Affaires étrangères, Direction générale de la Coopération internationale et du Développement, *Évaluer la coopération internationale française - Bilan des évaluations 2003*, Paris 2004
- Ministry of Foreign Affairs, Danida, *Joint Nordic Evaluation of the Beira Port Transport System (BPTS) Programme Mozambique*, Copenhagen 1995
- Ministry of Foreign Affairs, Danida, *Joint Evaluation of the Road Sub-Sector Programme Ghana 1996 – 2000*, Copenhagen 2000
- Montes, C. and S. Migliorisi, *EU Donor Atlas – Mapping Official Development Assistance*, Brussels, May 2004
- Naidoo, Sharmala, *Joint Evaluations; the South African Experience*, Power Point presentation, Nairobi, 20 – 21 April 2005
- Netherlands Ministry of Foreign Affairs, Operations Evaluation Department, *Solutions locales à des défis mondiaux: Vers un partenariat efficace en éducation de base – Etude des documents*, The Hague, September 2003
- Netherlands Ministry of Foreign Affairs, Operations Evaluation Department, *Local Solutions to Global Challenges: Towards Effective Partnership in Basic Education*, Final Report, The Hague, September 2003
- Netherlands Ministry of Foreign Affairs, Operations Evaluation Department, *Local Solutions to Global Challenges: Towards Effective Partnership in Basic Education, Key Findings of the Joint Evaluation of External Support to Basic Education in Developing Countries*, The Hague 2003
- Netherlands Ministry of Foreign Affairs, Operations Evaluation Department, *From Evaluation to Policy and Practice: Aid and Education*, Report on International Colloquium, The Hague, 2004
- Netherlands Ministry of Foreign Affairs, Operations Evaluation Department, *Country-led Joint Evaluation of the ORET/MILIEV Programme in China*, Terms of Reference, The Hague and Beijing, 16 March 2004

- Nordic Consulting Group, *Beira Port Transport System Programme – Round Table Meeting No. 3*, 30 November to 1 December 1995, Minutes of the Meeting, Copenhagen, February 1996
- Norwegian Agency for Development Cooperation (NORAD), *Guidelines for Norway's Provision of Budget Support for Developing Countries*, Oslo, July 2004
- Norwegian Institute for Urban and Regional Research, *Appraisal of the Local Government Reform Programme Tanzania*, Final Report by the Joint Appraisal Mission, Pretoria and Oslo, April 1999
- odcp consult gmbh, *SDC – OED Partnership Program for Development Effectiveness through Evaluation, Assessment of the SDC/BWI-WB/OED Partnership Program 2nd Phase (2000-2004)*, Final Report, Zurich, 18 September 2004
- Organisation for Economic Co-operation and Development, *DAC Principles for Evaluation of Development Assistance*, OCDE/GD (91)208, Paris 1991
- Organisation for Economic Co-operation and Development, *Development Co-operation 2003 Report*, OECD Paris 2003
- Organisation for Economic Co-operation and Development, *Highlights of joint evaluations – Members responses to request for information*, Room Document No. 4, submitted to the 37th meeting of the DAC Working Party on Aid Evaluation, Paris, March 2003
- Organisation for Economic Co-operation and Development, *Lessons Learned on Donor Support to Decentralisation and Local Governance*, DAC Evaluation Series, OECD Paris 2004
- Picciotto, Robert, *Economics and Evaluation*, Washington, D.C., 28 August 2000
- Picciotto, Robert, *The Global Policy Dimension of Development Evaluation*, paper submitted to the Symposium on the Internationalization of Evaluation, United Kingdom Evaluation Society Conference 2002
- Republic of Mozambique, Ministry of Agriculture and Rural Development, *Proagri Evaluation*, Vol. 1 Main Report, Maputo, May 2003
- Royal Ministry of Foreign Affairs, *Evaluation of Development Assistance*, Handbook for Evaluators and Managers, Oslo, November 1993
- Royal Ministry of Foreign Affairs, *Peace building – a Development Perspective*, Strategic Framework, Oslo, August 2004
- Scanteam and SPM Consultants, *A Report of the Joint Sida and Norad/MFA Review of the Development Co-operation with Malawi*, Oslo, June 2004
- Steering Committee for the Evaluation of the Joint Government and Humanitarian Partners Response to the 2002-03 Emergency in Ethiopia, *Evaluation of the Response to the 2002-03 Emergency in Ethiopia*, Addis Ababa, October 2004
- Steering Committee of the Joint Evaluation of Emergency Assistance to Rwanda, *The International Response to Conflict and Genocide: Lessons from the Rwanda Experience*, five volumes, Copenhagen, March 1996

- Sweden, Government of, *Shared Responsibility: Sweden's Policy for Global Development*, Government Bill 2002/03:122, Stockholm, 15-05-2003
- Swedish International Development Cooperation Agency, *Looking Back, Moving Forward*; Sida Evaluation Manual, Stockholm 2004
- Swiss Agency for Development and Cooperation (SDC) and National Treasury, Republic of South Africa, *Swiss-South African Development Cooperation Programme 2000 – 2003*, Joint Review, Pretoria, January 2004
- The World Bank, *Toward Country-led Development – A Multi-Partner Evaluation of the Comprehensive Development Framework*, Synthesis Report, Washington, D.C., 2003
- The World Bank, Operations Evaluation Department, *Partnership for Education in Jordan*, Précis No. 193, Washington, D.C., 2000
- The World Bank, Operations Evaluation Department, *The Drive to Partnership: Aid Coordination and the World Bank*, Washington, D.C., 2001
- The World Bank, Operations Evaluation Department, and African Development Bank Operations Evaluation Department, *Lesotho – Development in a Challenging Environment*, Abidjan and Washington, D.C., 2002
- The World Bank, Operations Evaluation Department, *Rwanda: Country Assistance Evaluation*, Washington, D.C., 2004
- The World Bank, Operations Evaluation Department, *Jordan: Supporting Stable Development in a Challenging Region*, Washington, D.C. 2004
- The World Bank, Operations Evaluation Department, *The Poverty Reduction Strategy Initiative, An Independent Evaluation of the World Bank's Support Through 2003*, Washington, D.C., 2004
- United Nations Development Programme (UNDP) and German Federal Ministry for Economic Cooperation and Development, *The UNDP Role in Decentralisation and Local Governance*, New York 2000
- United Nations Development Programme (UNDP), Evaluation Office, *Handbook on Monitoring and Evaluating for Results*, New York 2002
- United Nations Development Programme (UNDP), Evaluation Office, *Development Effectiveness Report 2003 – Partnerships for Results*, New York 2003
- United Nations Development Programme (UNDP), Evaluation Office, *Danish Trust Fund for Capacity Development*, Essentials No. 13, New York, November 2003
- United Nations Evaluation Group (UNEG), *UNDAF Evaluation*, Guidelines for Terms of Reference, Draft, New York, 6 August 2004
- United Nations Office for the Coordination of Humanitarian Affairs (OCHA), *Learning from experience: Lessons from an inter-agency evaluation process*, New York, 1 June 2002

Vu Dai Thang, *Directions for ODA Joint Evaluation in Vietnam*, Power Point presentation, Nairobi, 20 – 21 April 2005

Weingärtner, Lioba, *Joint Evaluation of the Effectiveness and Impact of the Enabling Development Policy of the World Food Programme – Lessons Learned Paper*, Rottenburg, June 2005

Yantio, Debazou Y., *Joint Program Evaluation in Cameroon: Whose Perspective Matters, What to Do?* Paper submitted to the DAC Workshop on Joint Evaluations, Nairobi, April 2005