PREPARING, GOVERNING, AND MANAGING THE PARIS DECLARATION EVALUATION

Niels Dabelstein
Danish Institute for International Studies
Copenhagen, Denmark

Ted Kliest
Ministry of Foreign Affairs
The Netherlands

Abstract: Joint or multi-partner evaluations are evaluations of development cooperation policies, programs, and projects in which different donors, development agencies, and partner countries participate. The complexities of a large number of diverse stakeholders and multiple units of analysis in joint evaluations pose significant governance and management challenges to ensure the evaluation’s independence, credibility, quality, and utility. This article reports how governance and management were structured and operated to facilitate the evaluation of the Paris Declaration. A common evaluation framework was established to facilitate synthesis. The integrity of national evaluations had to be ensured, including capacity-building and support as needed. National and international reference groups were established to ensure the engagement and buy-in of different stakeholder groups, including input and feedback to the core team that synthesized the results of the national evaluations. The article concludes with three important lessons about complex joint evaluations.

Résumé : Les évaluations conjointes ou à plusieurs partenaires portent sur des politiques, des programmes, et des projets de coopération au développement auxquels participent différents donateurs, agences de développement, et pays partenaires. La complexité liée au grand nombre de parties prenantes et d'unités d'analyse pose de sérieux défis en matière de gouvernance et de gestion pour garantir l'indépendance, la crédibilité, la qualité, et l'utilité de l'évaluation. Cet article présente la façon dont ces défis ont été relevés dans le cadre de l'Évaluation de la Déclaration de Paris. Une grille d'évaluation commune a été établie en vue de faciliter la synthèse. Il a fallu veiller à ce que les normes fondamentales soient suivies dans les évaluations nationales, si besoin en renforçant les capacités ou en apportant d'autres formes d'appui.

Corresponding author: Niels Dabelstein, Ostbanegade 21, 3tv, 2100 Copenhagen O, Denmark; <dabelstein@telia.com>
The Canadian Journal of Program Evaluation

Des groupes de référence nationaux et internationaux ont été formés afin de garantir l’engagement et l’adhésion des différents groupes de parties prenantes et de s’assurer de leur apport et de leur rétroaction à l’équipe chargée de synthétiser les résultats des évaluations nationales. En conclusion, l’article tire trois leçons cruciales concernant les évaluations conjointes complexes.

INTRODUCTION

Joint or multi-partner evaluations of international development cooperation policies, programs, and projects pose significant governance and management challenges to ensure the evaluation’s independence, credibility, quality, and utility. This article presents details on the governance and management structures and operations of the evaluation of the Paris Declaration. Our roles in the evaluation involved heading the Secretariat for the Evaluation of the Paris Declaration on Aid Effectiveness (Dabelstein) and co-chairing the evaluation’s Management Group (Kliest). In these roles, we observed firsthand how the different arrangements (International Reference Group, Evaluation Management Group, Evaluation Secretariat, and national reference groups established for the different components of this joint evaluation) functioned. Thus, this article is based on our experiences, observations, and reflections. Much of the article is descriptive in factually presenting how the management and governance structures were organized. However, we also offer reflections and judgements about the effectiveness and efficiency of the evaluation’s management and governance processes. In rendering such judgements, we are far from unbiased observers. Thus, this article is written in the tradition of reflective practice in which we offer our insiders’ views of how the evaluation was organized and the implications, from our perspective, for joint evaluations generally. Independent validation of what we present here, including independent judgement about the importance and effectiveness of the governance and management structures and processes employed, can be found in the independent evaluation of the Paris Declaration Evaluation (Patton, 2012, in this volume; Patton & Gornick, 2011).

IMPORTANT ASPECTS OF JOINT EVALUATION

Joint or multi-partner evaluations are evaluations of development cooperation policies, programs, and projects in which different donors, development agencies, and partner countries participate (Organisation for Economic Development and Cooperation/Develop-
The Network on Development Evaluation, hereafter referred to as Evaluation Network, is a subsidiary body of the Development Assistance Committee (DAC) of the Organisation for Economic Co-operation and Development (OECD). Its purpose is to contribute to improved development effectiveness by supporting robust, informed, and independent evaluation. Among other things, the Evaluation Network encourages harmonization and standardization of evaluation programming and facilitates joint evaluations. It has been promoting more joint evaluation work in line with the DAC’s broader agenda on enhancing donor coordination and collaboration between donors, development agencies, and partner countries in the developing world.

Joint evaluations have been a feature of international development cooperation for the last three and a half decades. Their significance and frequency has increased since the mid 1990s. In 2000 the DAC Working Party on Aid Evaluation, the predecessor of the current Evaluation Network, published *Effective Practices in Conducting a Joint Multi-Donor Evaluation*. This guidance booklet primarily focused on one type of joint evaluation, that is, those involving multi-donor collaboration. It did not provide specific guidance to other types of joint evaluations: those involving donor/development agency-partner country collaboration or evaluations characterized by multiple stakeholder collaboration.

In his study *Joint Evaluations: Recent Experiences, Lessons Learned and Options for the Future*, Breier (2005) reviewed more than 50 joint evaluations of development cooperation published between 1990 and 2005 with the aim of throwing light on their benefits, challenges, and good practices. He distinguished three types of joint evaluation. The classic multi-partner evaluation is open to all stakeholders who participate and contribute actively and on equal terms. The qualified multi-partner evaluation, where participation is open to those who qualify on the basis of certain criteria, for example, a joint evaluation conducted or commissioned by partners engaged in a certain sector-wide program. The hybrid multi-partner evaluation covers a range of complex ways of working together. For instance, the responsibility for the evaluation may be delegated to some stakeholders while others take the role of “silent partner.” Alternatively, some parts of the evaluation can be undertaken jointly while other parts are handled by individual stakeholders. Finally, other approaches establish similar but parallel evaluations or agree to a common evaluation framework that is used to guide evalua-
tions conducted by different partners. Breier encountered only a few examples of the “classic multi-partner evaluation.” He concludes, “Consequently, there is little sign of a functioning market place 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 … Instead, most multi-partner evaluations are generated by individual actors negotiating through their bilateral or group channels to solicit, support and find partners” (Breier, 2005, p. 17).

In general, joint evaluations provide potential benefits to all stakeholders involved by offering the opportunities to harmonize and align the evaluation process; to build ownership, partnership, and evaluation capacities; to share the burden of work and costs involved; and to lower the number of parallel evaluations, thereby reducing the administrative burdens and transaction costs on partner countries and donors. They also enable investigation of subjects that cannot be covered by single-actor evaluations such as joint programs, harmonized policies, and multi-donor/multi-agency development efforts—all activities whereby development goals are to be achieved through partnerships between donors, development agencies, and the partner country. On the other hand, joint evaluations pose specific challenges and difficulties. The partners (to be) involved may have different operational processes and evaluation cultures. They may wish to adhere to different evaluation approaches and methods. They may also have their own priorities and political agendas. All of this may result in lengthy and costly processes of building consensus regarding the objective, scope, approach, and methodology of a joint evaluation. In addition, the evaluation of subjects that are specifically suited to joint evaluation, such as co-financed program support at the sector or general budget level, and evaluation of the results or impacts of activities supported by and involving several actors that are expected to work in a harmonized or coherent manner, offer other methodological challenges compared to evaluations of stand-alone single-actor interventions (projects).

In their article, Winckler Andersen & Broegaard (2012) provide an interesting critique on joint evaluations and the motives for evaluation departments of individual donors and development agencies to engage in such evaluations. They argue that many joint-donor evaluations seem to run into management and coordination problems, have unforeseen high transaction costs for the evaluation departments involved and, more importantly, result only in general
conclusions and recommendations that may not be very useful for the individual partners engaged in the evaluation. They argue that the broad scope of many joint-donor evaluations, their high transaction costs, and their limited utility are related to differences in donor interests and incentive structures and an insufficient attention to the methodological challenges posed by such endeavours.

In 2006 the OECD/DAC guidance on joint evaluation was revised with the aim of covering all types of joint evaluation work (OECD/DAC, 2006). The revised booklet echoes the principles of the Paris Declaration on Aid Effectiveness. It reiterates the importance—if not necessity—that joint evaluations enable donors, development agencies, and partners to collectively assess the results and impacts of their combined efforts, which is vital at a time in which the international community is prioritizing more effective ways of managing aid and emphasizing the importance of mutual accountability in demonstrating results ... in a context where significant quantities of aid are now provided through joined-up modalities such as basket funds, sector-wide approaches and budget support and within the frameworks of coordinated and joint assistance strategies. (OECD/DAC, 2006, p. 3)

Importantly, the guidance underlines the need and importance of timely and genuine involvement of all partners in the programming of a joint evaluation; its objective, approach, and methodology; the governance and management of its implementation; and the manner in which the evaluation results may be utilized for policy and practice. These issues have also been pointed out by Samoff (2005) and Jensen & Eriksson (2009).

Important success factors of joint evaluations include (a) timely and inclusive planning to identify and engage evaluation partners; (b) generating commitment of all partners to the joint evaluation; (c) inclusive, consensus-based decision-making that, although sometimes cumbersome, time-consuming, and costly, is necessary to ensure the evaluation’s legitimacy; (d) agreed arrangements for dissemination, follow-up, and feedback, which are essential if a joint evaluation is to be useful for all involved; and (e) building on mutual trust and confidence, rather than tight bureaucratic control and predetermined administrative processes.
As the *Guidance for Managing Joint Evaluations* (OECD/DAC, 2006) points out, the evaluation should be managed by adopting common ground rules for its development and implementation by all partners. A commonly applied governance and management structure consists of a steering committee or reference group with a broad membership of all interested partners. This forum normally apportions the evaluation’s costs, provides broad guidance and direction to the evaluation, takes responsibility for key decisions such as approving the Terms of Reference for the evaluation, and releases the evaluation products including the final report after having ascertained the quality and credibility of these products. The steering committee or reference group normally appoints a smaller management group that is entrusted with the preparation of the evaluation—developing the Terms of Reference and the budget for the evaluation, and selecting and contracting the evaluation team(s) commissioned to conduct the actual evaluation. The management group takes care of day-to-day running of the evaluation process and has the role of effectively and efficiently liaising with the evaluators.

**DEVELOPING THE INITIAL APPROACH OF THE EVALUATION OF THE PARIS DECLARATION**

Besides its strong focus on monitoring progress, the Paris Declaration highlighted the importance of conducting an independent cross-country evaluation of the implementation of the Declaration’s principles and their results in terms of enhancing aid and development effectiveness. The evaluation was to provide a comprehensive understanding of how increased aid effectiveness contributes to meeting the development objectives of the partner countries without imposing additional burdens on these countries.

When the Paris Declaration was adopted at the High Level Forum in Paris in March 2005, the Evaluation Network, with the ample evaluation expertise of its collective membership and its independent position within the DAC, explored how it could contribute to developing the required evaluation. A number of Network members were willing to help design a proposal for a joint evaluation that—as was stipulated—needed to be developed with the participation of partner countries and in collaboration with the OECD/DAC Working Party on Aid Effectiveness. This forum was originally a subsidiary body of the OECD/DAC. Its membership (DAC donors) was expanded to include partner country representatives and it became a “freestanding committee” hosted by the OECD/DAC.
The Working Party on Aid Effectiveness, which was to commission studies on aid and development effectiveness, established the Joint Venture on Monitoring the Paris Declaration. This group representing donors, development agencies, and partner countries was made responsible for developing and implementing a system to periodically monitor the implementation of the principles of the Paris Declaration. Collaboration with the Working Party on Aid Effectiveness and the Joint Venture on Monitoring the Paris Declaration would ensure the complementarity of the proposed evaluation with the monitoring process.

In September 2005, the Evaluation Network contacted the Joint Venture on Monitoring the Paris Declaration to seek this collaboration. The initiative for the evaluation was welcomed, and the Evaluation Network was asked to develop “a cohesive cross-country evaluation process.” The evaluation was to deliver its findings quickly and go as far down the results chain as possible.

Following a discussion of the purpose of the proposed evaluation at the meeting of the Working Party for Aid Effectiveness in October 2005, it was agreed that the Evaluation Network would take the first step and commission an approach paper elaborating options for developing an evaluation framework. In November 2005, the Network on Development Evaluation agreed on the terms of reference for the development of an options paper, which was to detail the possibilities for the design, management, and implementation of the evaluation. The British Overseas Development Institute (ODI) was commissioned to write the paper. Members of the Joint Venture on Monitoring, the Working Party on Aid Effectiveness, the Evaluation Network, and representatives of civil society organizations were invited to provide inputs to the options paper through interviews with the ODI consultants. The paper was also to draw on a comprehensive literature review. The options paper provides a qualified positive assessment of the evaluability of the Paris Declaration commitments, outlines a possible evaluation framework, and details the various trade-offs to be considered in designing the evaluation process (Booth & Evans, 2006).

More particularly, it proposed the development of a common evaluation framework articulating the logic of the Paris commitments and the implied linkages between aid effectiveness and development effectiveness. This framework was to include
1. a number of country evaluations building directly on the results of the monitoring activities to be undertaken and other available knowledge about the management and effects of aid. These evaluations were to be initiated and led by the partner countries;

2. a series of donor country and development agency case studies, which would look at the way in which the Paris Declaration was operationalized across a sample of donors and development organizations; and

3. a medium- to long-term program of analytical work designed to draw together and critically assess findings from a variety of sources. This program was to be coordinated with the activities of the Working Party on Aid Effectiveness and the Joint Venture on Monitoring the Paris Declaration to avoid duplication and to ensure the value added of any study conducted in connection with the evaluation.

The draft options paper was presented separately to the Joint Venture on Monitoring and the Evaluation Network, both meeting in March 2006. Although the Evaluation Network decided to support the proposed joint evaluation, some members cautioned against starting an evaluation too early because the Paris Declaration was just being implemented and results could not yet be expected.

It was agreed to consult with the Working Party on Aid Effectiveness to seek guidance from its partner country members on the timing of the evaluation, the willingness of partner countries to join the evaluation, and the prospects among donors, development agencies, and partner countries to actually undertake evaluation studies.

The Optimal Timing for the Evaluation

It was argued that an evaluation would only be meaningful once donors, development agencies, and partner countries had implemented the Paris Declaration commitments for some time and sufficient monitoring data would be available to feed into the evaluation process. However, delaying the start of the evaluation for too long would make it less relevant and not timely enough to feed its results into the ongoing international debate on aid and development effectiveness. A key question, therefore, was whether the Working Party on Aid Effectiveness would prefer either a formative evaluation focusing on early lesson learning while the aid effectiveness agenda (the Paris Declaration) was being implemented, or a summative evaluation to post-hoc assess the final results and impacts of the Paris Declaration.
This “either/or” choice was not considered useful, and it was proposed to follow a two-phase approach consisting of a formative evaluation, which was to start immediately, to be followed some years later by a summative evaluation. The first phase of the evaluation should focus on if and how the principles of the Paris Declaration were implemented. It should produce a report synthesizing the findings of the formative country-level evaluations and donor/agency studies in time to feed into the preparations for the third High Level Forum of the Paris Declaration planned to take place in the second half of 2008. The summative evaluation assessing the results and impacts of the Paris Declaration on aid and development effectiveness would subsequently be conducted, with a report becoming available for the next High Level Forum planned for 2011. This two-phase approach for the evaluation would require careful management of stakeholders’ expectations about what meaningful results the evaluation could deliver prior to the High Level Forum of 2008. It was communicated that the evaluation would focus at this stage on lesson learning and throw light on major issues such as “What have we learned so far from experiences with implementing the Paris Principles in different country contexts?” and “Under what conditions and why are envisaged benefits realized or not realized?” Specific evaluation questions and issues would be identified through a consultative process of developing the evaluation framework. Exhibit 1 presents a detailed timeline for major milestones in the evaluation of the Paris Declaration.

Willingness Among Partner Countries to Join a Task Team to Coordinate the Evaluation

*Guidance for Managing Joint Evaluations* (OECD/DAC, 2006) would be the key reference document for managing the evaluation process. Rather than opting for the “usual” architecture of a central evaluation with field studies among partner countries, donors, and agencies, it was proposed that the Paris Declaration evaluation should be conducted in a decentralized manner, but with a “sufficiently strong level of coordination” to ensure an effective and credible “cross-country evaluation process.” It was suggested that a small task team consisting of Evaluation Network members and partner country members of the Working Party on Aid Effectiveness would be responsible for developing an overall evaluation framework. This framework was to ensure comparability of findings from different country-level evaluations and donor/agency studies, adequate coordination and management of the joint evaluation process, and an effective synthesis of evaluation findings and recommendations.
Members of the Evaluation Network that showed an initial interest in joining the task team included Ireland, Spain, France, Japan, Sweden, Denmark, the United Kingdom, the United Nations Development Programme (also representing the United Nations Evaluation Group), and the World Bank (also representing the Evaluation Coordination Group of the multilateral development banks). Denmark offered to coordinate the evaluation process. Partner country members of the Working Party on Aid Effectiveness were invited to join the task team. Where possible, meetings would be organized back-to-back with meetings of the Working Party on Aid Effectiveness or the

Exhibit 1
Management Structure of the Evaluation of the Paris Declaration
Joint Venture on Monitoring the Paris Declaration to minimize travel cost and facilitate a close relationship with the evolving Medium Term Monitoring Plan developed by the Joint Venture.

The options paper suggested that the development of the evaluation framework should be organized through two to three regional workshops involving partner countries that would be invited to come up with specific questions and issues to be covered by the evaluation framework.

Prospects Among Donors, Development Agencies, and Partner Countries to Undertake Evaluation Studies

The options paper stressed that the evaluation should provide a balanced picture by assessing the implementation of the Paris Declaration and its effects at the country level and at the level of individual donors and development agencies. It further suggested that partner countries, donors, and development agencies should determine whether they would engage with the evaluation. Self-selection aimed to meet the Paris Declaration stipulation that the evaluation would not create unwanted additional burdens by being imposed on partners. As a result, the partner country members of the Working Party on Aid Effectiveness were invited to volunteer to conduct a country-level evaluation. In turn, the donor and agency members of the Evaluation Network were invited to indicate their willingness to commission their own studies and provide financial support for the evaluation and/or technical support to partner countries to enable them to implement the country-level evaluations.

During September and October 2006, representatives of the Evaluation Network presented the general evaluation concept at three regional workshops: the Sub Saharan Conference on Aid Effectiveness (Durban, South Africa); the regional workshop on harmonization and alignment for Latin America and the Caribbean (Santa Cruz, Bolivia); and the Asian Regional Forum on Aid Effectiveness: Implementation, Monitoring and Evaluation (Manila, the Philippines). At these occasions short timeslots were made available to present and discuss the evaluation concept. However, as always at such events, networking during breaks was important to answer queries, to explain details, and for canvassing and soliciting support. By the end of October quite a number of partner countries, donor countries, and development agencies indicated their interest in participating actively in the first phase of the evaluation as shown in Exhibit 2.
Exhibit 2
Evaluation Milestones

<table>
<thead>
<tr>
<th>Year</th>
<th>Month</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>March</td>
<td>Options Paper by DAC Network on Development Evaluation</td>
</tr>
<tr>
<td></td>
<td>Mar.–Dec.</td>
<td>Consultations with partner countries</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2nd Reference Group meeting, Copenhagen: Launch of Phase 1</td>
</tr>
<tr>
<td>2008</td>
<td>February</td>
<td>3rd Reference Group meeting, South Africa: Emerging Findings workshop</td>
</tr>
<tr>
<td></td>
<td>March</td>
<td>4th Reference Group meeting, Paris: draft Synthesis report</td>
</tr>
<tr>
<td></td>
<td>June</td>
<td>Phase 1 Synthesis Report published</td>
</tr>
<tr>
<td></td>
<td>September</td>
<td>3rd High Level Forum in Accra, Ghana: Completion Phase 1</td>
</tr>
<tr>
<td>2009</td>
<td>February</td>
<td>1st Reference Group meeting, Auckland: Phase 2 Approach approved</td>
</tr>
<tr>
<td></td>
<td>December</td>
<td>2nd Reference Group meeting, Paris: Launch of Phase 2</td>
</tr>
<tr>
<td>2010</td>
<td>December</td>
<td>3rd Reference Group meeting, Indonesia: Emerging Findings workshop</td>
</tr>
<tr>
<td>2011</td>
<td>April</td>
<td>4th Reference Group meeting, Copenhagen: Phase 2 draft Synthesis report</td>
</tr>
<tr>
<td></td>
<td>June</td>
<td>Phase 2 Synthesis report published</td>
</tr>
<tr>
<td></td>
<td>November</td>
<td>4th High Level Forum in Busan, Korea: Completion Phase 2</td>
</tr>
<tr>
<td>2012</td>
<td>February</td>
<td>Workshop on lessons learned from the evaluation process</td>
</tr>
</tbody>
</table>

Ultimately, all partner countries except Morocco and Mali, which did not get their evaluations organized, became involved in the first phase of the evaluation. Zambia started the preparations for its evaluation, but the process lost momentum and its country evaluation was completed during the second phase of the overall evaluation. All donors and development agencies took part in Phase 1 and conducted studies. It is noteworthy that, on the donor side, four major players did not actively take part in Phase 1: the European Commission, Japan, the USA, and the World Bank. The USA and Japan, however, did actively participate in the evaluation in Phase 2 and commissioned their own studies, in contrast to the European Commission and the World Bank. The Independent Evaluation Group of the World Bank argued that an evaluation process where partner countries, donors, and development agencies were to commission and manage their own evaluations or studies would not be sufficiently independent, but it nevertheless became involved as a “silent member,” following the evaluation process at milestone moments. Despite several appeals, the European Commission’s evaluation department showed no interest in becoming engaged even as a silent member or observer.

In December 2006, a small group of evaluation experts and evaluation managers from the African Evaluation Association (AfrEA), the evaluation department of Danida (Ministry of Foreign Affairs of Denmark), the National Treasury of South Africa, the Ministry of
Planning and Investment of Vietnam, and an independent evaluation expert from the USA developed the first draft of the evaluation framework.

This framework, building on the options paper and taking into account the issues discussed at the three regional conferences, (a) specified the “architecture” of the evaluation: two distinct phases delivering outputs in 2008 and 2010 (later changed to 2011 as the 4th High Level Forum was delayed) with a focus on country-level evaluations and donor and development agency headquarters studies; (b) foresaw “supplementary studies” in which the subject was to be defined at a later stage; (c) indicated that each partner country, donor, and development agency was to commission its own evaluation or study, applying “generic” terms of reference that were to be developed by the Evaluation Management Group (see below) for the country evaluations and headquarters studies; (d) pointed out that the results of these evaluations and studies were to be synthesized by a team of independent evaluators—the “synthesis team”; and (e) defined the governance structure with a broad-based International Reference Group and a small Evaluation Management Group supported by a Secretariat. It was also pointed out that each country evaluation and donor and agency headquarters study was to be guided by a reference group to be instituted by the country, donor, or agency at issue. The governance and management structure is depicted in Exhibit 3.

Exhibit 3
Partner Countries, Donors and Agencies Showing an Initial Interest in the Evaluation (Phase 1)

<table>
<thead>
<tr>
<th>Partner countries</th>
<th>Donor countries and agencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bangladesh</td>
<td>Asian Development Bank</td>
</tr>
<tr>
<td>Bolivia</td>
<td>Australia</td>
</tr>
<tr>
<td>Cambodia</td>
<td>Denmark</td>
</tr>
<tr>
<td>Mali</td>
<td>Finland</td>
</tr>
<tr>
<td>Morocco</td>
<td>France</td>
</tr>
<tr>
<td>Senegal</td>
<td>Germany</td>
</tr>
<tr>
<td>South Africa</td>
<td>Luxemburg</td>
</tr>
<tr>
<td>Sri Lanka</td>
<td>The Netherlands</td>
</tr>
<tr>
<td>The Philippines</td>
<td>New Zealand</td>
</tr>
<tr>
<td>Uganda</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>Viet Nam</td>
<td>UNDP (UN Development Group–UNDG)</td>
</tr>
<tr>
<td>Zambia</td>
<td></td>
</tr>
</tbody>
</table>
The draft evaluation framework was circulated among interested partner countries, donors, and development agencies and discussed at the first meeting of the International Reference Group (IRG) in Paris in March, 2007. The IRG would be co-chaired by a representative from a partner country and a representative from a donor country and was composed of (a) partner-country representatives of the Working Party for Aid Effectiveness plus representatives of other partner countries undertaking country-level evaluations; (b) members of the Network on Development Evaluation representing individual donor countries and development agencies including the Evaluation Office of the United Nations Development Program (UNDP) representing the United Nations Evaluation Group, and the Asian Development Bank; (c) a representative of Reality of Aid/Ibon representing civil society organizations in partner countries, and a representative of EURODAD/Bond representing civil society organizations in donor countries; and finally (d) a representative of the African Evaluation Association (AfrEA).

The IRG appointed a small Evaluation Management Group (EMG) consisting of evaluation experts/managers from two partner countries, two donor countries, and a multilateral development agency. In view of the magnitude and complexity of the evaluation, a two-person Evaluation Secretariat was funded by the evaluation department of the Danish Ministry of Foreign Affairs and hosted by the Danish Institute for International Studies.

International Reference Group and Evaluation Management Group

The IRG provided overall strategic guidance for the evaluation, worked to ensure “buy-in” to the evaluation process from the different stakeholder groups (partner countries, donors and development agencies, and civil society), and ensured that adequate resources were generated. Specific tasks included (a) oversight of the EMG; (b) endorsing the overall framework for the evaluation; (c) reviewing Terms of Reference for the partner country evaluations, for the donor and development agency headquarters studies, for the supplementary thematic or cross-cutting studies, and for the study that would synthesis the findings of the component evaluations and studies; (d) reviewing the draft synthesis report, endorsing the evaluation’s dissemination strategy, and participating in dissemination activities.
The IRG convened at “milestone moments” during the first phase of the evaluation, for example, when emerging findings could be shared and discussed and the points in time when the draft final synthesis report could be reviewed. To the extent possible the IRG meetings were planned back-to-back with other international events in which its members would need to participate (meetings of the Evaluation Network on Development Evaluation, the Working Party on Aid Effectiveness, or the Joint Venture on Monitoring the Paris Declaration).

The EMG fulfilled the executive role of coordinating the overall evaluation process and managing the evaluation’s components to ensure progress according to the agreed time schedule and within budget. Tasks included (a) keeping the IRG and other relevant stakeholders appropriately involved with and informed about the evaluation’s progress; (b) developing all Terms of Reference for various components (country evaluations, headquarters studies, thematic and cross-cutting studies, and the synthesis) and criteria for selecting consultants as needed in these efforts; (c) providing advice and support to partner countries, donors, and development agencies regarding the planning and implementation of their evaluation activities; (d) commissioning the thematic and cross-cutting studies (upon approval of their selection by the IRG), the synthesis report, and other consultancies as necessary; (e) ensuring that mechanisms for resolving disputes were developed and communicated to all involved; (f) developing a communication and dissemination strategy; (g) facilitating the establishment of the overall evaluation budget, which was to include direct financing by donors and agencies of their own studies and financing by individual donor members of the International Steering Group of evaluations to be commissioned by partner countries, and setting up and accounting for an Evaluation Trust Fund administered by the Evaluation Secretariat; and (h) developing a plan outlining the structure, approach, and methodology for the Phase 2 summative evaluation.

How Did It Work in Practice?

In general the evaluation’s governance structure functioned as foreseen. It was not overly bureaucratic, and its constituent parts provided the necessary guidance and advice to the partner countries, donors, and development agencies that undertook individual evaluations and studies. The management structure was also successful in ensuring the independence of the overall evaluation and its components by
putting in place mechanisms to foster and ensure the credibility and quality of the evaluation work. The IRG, the EMG, and the reference groups instituted by partner countries and donors/agencies to guide component evaluations or studies proved able to review the draft evaluation products as to their quality, credibility, and clarity. There was no evidence of political interference with the findings. The final report synthesizing the findings of the different components of the evaluation was adopted by the IRG, applying the same criteria of quality, credibility, and clarity. The findings, conclusions, and recommendations in the final report as well as those in the reports of the component evaluations and studies were the responsibility of the respective evaluation or study teams. In view of the very tight time schedule for the evaluation, the partner countries, donors, and development agencies that commissioned evaluations and studies were constantly reminded to deliver their products on time. This was essential to enable a proper synthesis process in order to deliver the final evaluation report in time to prepare for the third High Level Forum of the Paris Declaration in Accra in September 2008. Unfortunately, delays occurred that negatively impacted on the way the synthesis team could conduct its analysis and produce the final report. However, the final report of the first phase of the evaluation was produced in time to be utilized for the preparations of the High Level Forum.

Meetings of the International Reference Group

The IRG convened another four times following its first meeting described above. The second meeting took place in Copenhagen in June 2007 in connection with the launch workshop of the evaluation. At this workshop, partner country representatives and representatives from donor countries and development agencies who were to manage evaluation or studies familiarized themselves with the generic terms of reference. The discussions were at times quite heated, and it was not easy for some of the participants to familiarize themselves with the generic terms of reference they were to apply in their own evaluation or study.

The IRG convened for the third time on 31 January and 1 February 2008 in Parys, South Africa. This meeting coincided with an evaluation workshop involving the evaluation and study managers, the leaders of the evaluation or study teams, the two peer reviewers who had scrutinized the quality and logic of the different draft reports, and the team tasked with producing the synthesis report. Its objective was to discuss initial findings and issues emerging from
the different components of the evaluation. The workshop proved to be instrumental for the synthesis work. It also was the scene of intensive debates among all participants regarding the credibility and quality of the evaluation as a whole. Ultimately these debates resulted in a sufficient level of confidence among the members of the IRG that Phase 1 of the evaluation could be finalized in time and produce a credible report.

The fourth meeting was held in March 2008 in Paris to discuss the draft report. Taking into account advance written comments and points raised at the meeting, the synthesis team produced a revised draft that the Evaluation Management Group reviewed for quality, credibility, and clarity. Following this review and a final revision by the synthesis team, the EMG adopted the final report. The Evaluation Secretariat arranged publication and dissemination of the report in English, French, and Spanish, including a CD-ROM with all component evaluations and studies.

Following the High Level Forum in Accra it was time to take stock of the utilization of the report of Phase 1 of the evaluation and to decide on the direction and approach for Phase 2. Even before the first phase report was finalized, the EMG undertook steps to explore possibilities to shape this second phase by commissioning an evaluability study. The next meeting of the IRG, which was its first gathering in connection with the second phase of the evaluation, took place in February 2009 in Auckland, New Zealand. (More details on this session are presented below.)

In sum, the IRG proved to be instrumental during Phase 1 of the evaluation by adhering to its mandate that had been carefully crafted in order to ensure a high level of independence of the evaluation. In particular, the group’s role was to endorse the overall evaluation framework but only to comment on the generic terms of reference for the country evaluations, the headquarters studies, and supplementary studies, the reports of the different components of the evaluation, and the draft synthesis report. In essence, substantial authority for decision-making was delegated to the Evaluation Management Group.

Meetings of the Evaluation Management Group

The EMG met nine times between April 2007 and November 2008—five involved telephone conferences, while four were back-to-back
meetings with evaluation workshops and meetings of the IRG. In general, the EMG supported by the Evaluation Secretariat functioned well. Its tasks in managing the evaluation process proved time-consuming, resulting in unequal burdens on those members who were able to spend sufficient time in view of their regular tasks in their own institutions. The decision to establish a separate Evaluation Secretariat proved right, as the head of the Secretariat and his support staff were able to handle the bulk of the administrative work involved in the day-to-day management of this complex evaluation with its different components. Coordination was increased across governance responsibilities by having the head of the Secretariat also co-chair the IRG and EMG.

On the whole the EMG members worked in a harmonious manner. However, conflicts did arise at times regarding the way in which the evaluation had to be managed to ensure its progress and quality. This proved to be particularly the case at the time of the above-mentioned emerging findings workshop. The members’ backgrounds and experiences stemming from different evaluation cultures (bilateral donor, multilateral agency, and developing country) resulted in a disagreement concerning the expectation as to what should reasonably constitute a synthesis report in view of the quality of the (then) still preliminary drafts of the component evaluation and study reports. There was also considerable difference of opinion as to the way in which the work of the independent synthesis team should be guided, with some members suggesting that the EMG should be complemented by a separate team of “eminent evaluation experts” tasked to conduct a quality review of the draft synthesis report. The differences between the EMG members were the subject of sometimes heated and tense debates and led to an uneasy relationship between the EMG and the synthesis team. Eventually, both agreed on a pragmatic way forward that fully respected the independence of the synthesis team whilst ensuring the quality of its work.

The EMG’s ninth and final meeting during Phase 1 of the evaluation took place in November 2008 to reflect on the implementation of Phase 1. Attention was paid to the challenges encountered during the evaluation, in particular those connected with assuring the quality of its components to ensure the quality of the synthesis report, to create and maintain a sufficient level of active engagement of partner countries, donor countries, and development agencies in the evaluation (a partial failure as some important donors and development agencies did not wish to take part), and to manage the evaluation process in
order to deliver its final report in time. Recognizing these challenges and weaknesses, the EMG concluded that Phase 1 of the evaluation had been successful. The final report had been very well received by the Working Party on Aid Effectiveness and provided an important input to the preparatory consultations for the High Level Forum in Accra. The evaluation’s key findings were presented and discussed at these consultations, giving substance to the preparations of the High Level Forum’s agenda, which was highly appreciated by partner countries. The evaluation’s final report was often referred to at the High Level Forum, both in the Chair’s opening speech indicating the value added of the evaluation in relation to the results of the monitoring survey and during plenary discussions and those at the Thematic Round Tables. In order to facilitate the utilization of the evaluation, the synthesis team had produced concise thematic briefing notes on the evaluation’s findings and conclusions. These “two-pagers” proved to be an effective feedback mechanism. Last but not least, the final report was used to back arguments during the negotiations to reach agreement on the High Level Forum’s outcome document—the Accra Agenda for Action.

PHASE TWO OF THE EVALUATION

The Working Party on Aid Effectiveness met in November 2008 to reflect on outcomes of the Accra High Level Forum and to chart the way toward the next High Level Forum. At that meeting the Evaluation Secretariat presented a brief note on the approach to the second phase of the evaluation. Partner countries, donors, and development agencies were again invited to participate in the evaluation and help shape its process. During and following the meeting of the Working Party on Aid Effectiveness, many partner countries, donors, and development agencies expressed their interest in participating in the second phase. Ultimately 22 partner countries and 18 donors and development agencies actively engaged in Phase 2, including those who also participated in Phase 1.

Realizing the challenge of assessing the Paris Declaration’s contribution to aid effectiveness and development results, the EMG had already in January 2008 commissioned an evaluability study to throw light on the (theoretical) linkages between the principles of the Paris Declaration, aid effectiveness, and development effectiveness (Stern et al., 2008). This Linkages Study formed the basis for a draft approach paper for Phase 2, which the Evaluation Secretariat, in close consultation with the EMG and supported by a consultant, prepared.
The approach paper for the next phase of the evaluation was very detailed and precise in describing the evaluation process and the evaluation’s governance structure—details on the evaluation’s approach and methodology are provided elsewhere in this journal. It addressed two weaknesses identified in the initial phase of the evaluation. During that phase, the synthesis team was contracted late in the evaluation process, had not been engaged in design of the evaluation, and played no role in assuring the quality of the reports of country evaluations and the donor and agency studies. To mitigate this weakness, the synthesis team, now called the Evaluation Core Team (hereafter referred to as the Core Team) was to be contracted early in Phase 2 to be involved in developing the approach and methodology of the country evaluations and the synthesis. The methodology to be applied in additional donor and agency headquarters studies would be similar to that in Phase 1. A second shortcoming of the first phase was the lack of technical support provided to the partner countries, donors, and agencies during the implementation of their evaluations and studies. As a result these undertakings suffered delays and proved to be rather uneven in quality. In order to address this issue, the Core Team, in consultation with and supported by the EMG, focused on developing and refining the evaluation’s approach and methodology prior to the start of the component evaluations or studies. This very intensive process took place from September 2009 to June 2010. It involved two rounds of regional workshops with participation of the national evaluation coordinators, members of national reference groups, and the team leaders of the planned country evaluations.

Regional workshops took place in Cambodia, Colombia, Malawi, Benin, New Zealand, Bangladesh, Tunisia, Vietnam, and Bolivia. They focused on defining the key principles of the methodology, shaping the evaluation questions, confirming the methodology, working out how to operationalize the methodology, finalizing the evaluation compo-
nents, establishing the evaluation matrix, and setting out guidance on methods for country evaluations and donor and agency studies.

A separate workshop was held in the United Kingdom for evaluation managers of donors and agencies commissioning headquarters studies during Phase 2.

The Governance Structure for Phase 2

The evaluation’s governance and management structure was similar to that of Phase 1 of the evaluation.

The International Reference Group (IRG) consisted of a representative of each entity (partner country, donor, or aid organization) having a strong interest in the evaluation or actively participating in it, either through overseeing and coordinating participation in evaluation activities (country evaluations and/or donor and agency studies) or by providing a financial contribution to the evaluation as a whole or to separate components. This implied that every country commissioning a country evaluation was represented, as were donors and other organizations participating in and contributing to the evaluation. In addition, two umbrella organizations representing Northern and Southern civil society organizations were represented. In total, 52 governments and organizations were represented in the IRG, which was co-chaired by Malawi and Sweden.

The IRG was tasked with overseeing the evaluation and reviewing its products, including (a) the evaluation approach paper, the terms of reference for the Core Evaluation Team, generic terms of reference for the country evaluations and the donor and agency studies, the evaluation’s inception report produced by the Core Team, which included a paper detailing the evaluation’s methodology, and approach papers (and terms of references) for supplementary studies; (b) draft country evaluations, donor/agency studies, and supplementary study reports for quality, clarity, and credibility; and (c) draft and final synthesis report for quality, clarity, and credibility.

While the IRG needed to approve the products mentioned in item (a) above, it was to review and provide comments only on the products mentioned in (b) and (c). In order to protect the independence and credibility of the evaluation, the International Reference Group was not tasked to approve the various reports produced by the independent evaluation and study teams.
The IRG met at four milestone moments during Phase 2: in New Zealand (February 2009) to discuss and agree on the overall approach for the Phase 2 evaluation; in France (November 2009) to discuss and validate the Terms of Reference for the country evaluations and the donor and agency studies; in Indonesia (December 2010) to review the emerging findings of the (initial) synthesis of the component studies; and in Denmark (April 2011) to validate the draft synthesis report on quality, clarity, and credibility.

*The Evaluation Management Group* (EMG) comprised seven members, six representing partner countries and donors (Colombia, Malawi, the Netherlands, Sweden, the USA, and Vietnam) and the Evaluation Secretariat, which continued to support both the EMG and IRG. The Evaluation Secretariat was funded by the Policy and Operations Evaluation Department of the Netherlands Ministry of Foreign Affairs and remained located at the Danish Institute for International Studies. The EMG was co-chaired by the Netherlands and Vietnam. The group reported to the IRG, but was separately charged with the responsibility to assure the quality and safeguard the independence of the evaluation.

Responsibilities included (a) overseeing and guiding the work of the Evaluation Secretariat and acting as the tender committee for contracts; (b) reviewing all products by the Core Team at draft stage (generic terms of reference for the country evaluations and donor/agency studies, methodology paper, terms of reference for supplementary studies, inception report and progress reports of the Core Team, and its draft and final synthesis report, as well as other products that emanated from the evaluation); (c) ensuring that the Core Team gave full consideration and response to substantive comments from both the EMG and the IRG; and (d) assessing and approving all Core Team products in terms of their adherence to the terms of reference guiding the team’s work, their quality using the OECD/DAC Evaluation Quality Standards as a reference, and readability.

The EMG met eight times during Phase 2, with four meetings held back-to-back with those of the IRG. Between meetings, there were frequent e-mail and telephone communications.

Reference Groups for Component Evaluations and Studies

As in Phase 1, it was suggested to each of the partner countries commissioning its evaluation to establish a country reference group.
Again, each donor and development agency was to establish a reference group to guide the headquarters study.

The Country Reference Groups consisted of key government and non-government development stakeholders in a given partner country, including but not confined to the central government and key donors. Each group was to closely work together with the national evaluation coordinator. To foster ownership of the evaluation and interaction with the evaluation team at issue, each Reference Group had a more elaborated set of tasks. These included (a) approving the design of the country evaluation that comprises a common set of evaluation questions applicable to all country-level evaluations and a module with country-specific evaluation questions; (b) deciding on selection criteria for the country evaluation teams; (c) selecting the members of country evaluation teams, consistent with the selection criteria and national competitive procurement or tender rules; (d) serving as a resource and providing advice and feedback to the country evaluation teams; and last but not least (e) reviewing (but not approving) the draft products of the country evaluation in terms of their quality, clarity, and credibility.

The Reference Groups for donor and agency studies had similar tasks.

How Did It Work in Practice?

As was the case in Phase 1, the evaluation’s governance structure functioned in general as foreseen. But there were major differences from Phase 1 in how the structure functioned. First, a much more intensive interaction occurred between the EMG and the Core Evaluation Team as a result of their joint tasks to develop and refine the evaluation and its methodology. There was also a very strong involvement of individual EMG members as observers and resource persons in the regional workshops that focused on developing and clarifying the scope and methodology of the country-level evaluations. Third, the Core Evaluation Team interacted very intensively with the national evaluation coordinators and evaluation teams involved in the country-level evaluations in order to foster the quality of these evaluations. This involvement, however, did not fully ensure quality. Finally, there was much stronger involvement of the Country Reference Groups in guiding the country-level evaluations. There were signs of attempted interference in some cases, but diligent attention to the credibility and integrity of the evaluation ensured the ultimate independence of these evaluations.
The IRG played an important role in ensuring a high level of independence of the evaluation by endorsing the overall evaluation framework and reviewing and validating (approving) the generic terms of reference for the country evaluations, the headquarters studies, and supplementary studies, and the Core Team’s inception report. By attending the workshop during which the emerging findings of the (initial) synthesis of the component studies were discussed, the IRG members provided very useful input in further guiding the synthesis process. At the group’s final meeting, its members engaged in a thorough discussion of the draft final report, thereby playing an important role in validating the draft report in terms of its quality, clarity, and credibility. Again, the IRG proved able to delegate substantial authority for decision-making to the EMG.

Similar to Phase 1, the development, conduct, and finalization of the Phase 2 evaluation was not without tensions but, as mentioned elsewhere in this issue, approaches were developed to solve them without compromising the independence and credibility of the evaluation.

FINANCING

The evaluation of the Paris Declaration did not come cheap; its total cost amounted to €10.4 million. Undoubtedly some will consider this very expensive. However, the evaluation covered 22 countries and 18 donors/agencies directly, and many more indirectly. Moreover, the subject evaluated—the way development cooperation is functioning—is broad, complex, and of the essence for the future configuration of international development cooperation.

The costs of country evaluations and donor/agency studies are based on a combination of reported actual costs and estimated costs for several countries and donors that did not report their costs. The core costs comprised consultants conducting synthesis, thematic, and other studies, costs of the meetings of the International Reference Group and workshops, and dissemination. The costs for the Evaluation Secretariat covered staff salaries and travel. See Exhibit 4 for an overview of evaluation costs.

The core costs and several country studies were financed through a trust fund held by the Evaluation Secretariat. With amounts ranging from €50,000 to €790,000, 19 donors contributed to this fund. Many donors were very forthcoming in adjusting their original contributions when costs of workshops and dissemination exceeded the budget.


Exhibit 4
Cost of the Evaluation

<table>
<thead>
<tr>
<th></th>
<th>Phase 1 (Euros)</th>
<th>Phase 2 (Euros)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country Evaluations</td>
<td>580,700</td>
<td>1,794,700</td>
</tr>
<tr>
<td>Agency Headquarters Studies</td>
<td>824,300</td>
<td>1,256,900</td>
</tr>
<tr>
<td>Core Costs</td>
<td>1,781,200</td>
<td>2,835,400</td>
</tr>
<tr>
<td>Evaluation Secretariat</td>
<td>649,200</td>
<td>710,800</td>
</tr>
<tr>
<td>Total</td>
<td>3,835,400</td>
<td>6,597,800</td>
</tr>
</tbody>
</table>

Total Phase 1 & 2 10,433,200

It was anticipated that all country evaluations would be funded through the trust fund. This proved impossible, however, as many donor agencies’ evaluation departments had difficulties in providing funds for the core budget; they could fund country evaluations directly. As a result the country evaluations were financed in different ways: from the trust fund—some earmarked, some not; through direct funding from a donor agency evaluation department; by the donor office in the partner country; by the countries themselves—funds that were drawn from existing donor funding pools. Because such pooled funds have alternative uses, such funding is considered “self-financed.”

In particular, the direct donor funding of country evaluations—which could be labelled “tied aid”—goes against the principles of the Paris Declaration. Some countries whose evaluations were funded through the trust fund experienced difficulties in channeling the funds through the “normal” channel, that is, the Ministry of Finance. This resulted in some cases in delays in contracting evaluation teams.

Regardless of the source of funding, in all but a few cases the country evaluation teams were contracted in accordance with national procedures for procurement of goods and services. The exceptions were cases in which the national tender procedures were prohibitively slow and cumbersome, leading to donors contracting evaluation teams directly, but always in consultation with the partner country at issue. In two countries the evaluation teams were selected through national tender procedures and contracted by the country authorities, with invoicing of the teams being arranged by the Evaluation Secretariat.

In sum, in most respects the evaluation of the Paris Declaration was organized, managed, and financed in observance of the Paris
Declaration principles; the exception was using country financial and procurement systems—a principle that has proved difficult to implement in general.

CONCLUSION

The evaluation of the Paris Declaration was one of the evaluations discussed in a workshop aimed to distill lessons from recent international joint evaluations focusing on how these evaluations were developed, governed, and managed. The workshop was held in Paris in February 2012 and involved some 30 evaluation experts and evaluation managers from donors, development agencies, and partner countries.

Three important issues discussed at the workshop will be highlighted: (a) preparation of the evaluation and allowing “buy-in” of stakeholders, (b) building trust and confidence among the stakeholders, and (c) organizing the evaluation process and use.

Preparation and Buy-In

Thorough preparations are important in all evaluations, but are even more essential for joint evaluations where, as mentioned in the first section of the article, the motives for evaluation departments of donors and development agencies to engage in joint evaluations can be different, and many joint evaluations seem to run into management and coordination problems. It was very clear when the evaluation of the Paris Declaration was prepared and developed that sometimes diverging interests needed to be met and that background, knowledge, and capacity of the various partners who might become engaged in this evaluation differed considerably. Consequently, in setting up the governance and management structure for the evaluation, attention was paid to the following aspects:

1. Consultation was held with all stakeholders in order to ensure buy-in on the scope of the evaluation, to agree on the governance and management structure, and to establish common ground rules making this structure work.
2. The approach was to seek consensus without compromising the integrity of the evaluation process and product.
3. Care was taken to have a strong and effective lead (the Evaluation Secretariat and the co-chairs of both the Inter-
national Reference Group and the Evaluation Management Group) but also be open and ready to listen to the members of both groups as well as to those managing and conducting the individual evaluations and studies.

4. The rules of the game as implemented by the bodies forming the evaluation governance and management structure were developed to ensure evaluator independence in order to guarantee the integrity and credibility of the evaluation findings. Engaging diverse stakeholders to ensure relevance and use of the evaluation did not compromise this.

5. Those involved in governing and managing the evaluation clearly understood that it takes time to develop, implement, and finalize such a complex undertaking, but endeavoured to point out that timelines had to be respected in order to produce a report that would be relevant for policy makers.

It goes without saying that implementing the common ground rules did not happen without tensions arising. However, as outlined in the final article in this issue, these proved to be manageable because sufficient trust and confidence was built among the different stakeholders represented in the governance and management structure.

Building Trust and Confidence

As indicated in the 2006 OECD/DAC guidance on joint evaluations, it is very important to build confidence among the stakeholders commissioning large-scale joint evaluations. At the same time—as outlined in the first section of this article—it should be acknowledged that stakeholders’ experiences and points of view will differ, and although agreement must be reached on all principal issues, it may not be necessary or possible to agree on every little detail. Hence, in the evaluation of the Paris Declaration, the authority to work out the details was delegated to the small Evaluation Management Group, which worked in a very transparent fashion. The establishment of an independent, dedicated, and well-resourced Evaluation Secretariat, not attached to (or embedded in) any stakeholder organization, was another major factor contributing to stakeholders’ confidence in the evaluation process.

Ways to gain trust are, of course, to engage with all important stakeholders—be they influential policy makers or interest groups—and to relate to established evaluation or feedback mechanisms in the country/organization.
Evaluation Process and Use

A third important aspect to be considered by those involving in governing and managing the evaluation is the need to be process-oriented. It is not only important to allow ample time for preparation in order to allow sufficient buy-in and understanding of what the evaluation is about, it is equally important to allow for continuous consultations throughout the evaluation and to define milestone moments for decision-making. By engaging all stakeholders through the different levels of governance in the evaluation, it was possible to continuously focus the work and justify the investment in time and money as well as pay attention to utility and use of the evaluation. This avoided an important pitfall of joint evaluations: their limited utility for each of the individual stakeholders mentioned by Winckler Andersen & Broegaard (2012). In the final article of this issue, Michael Quinn Patton describes utility and use of the evaluation in more detail. A final aspect that was served by engaging the stakeholders in the evaluation’s governance and management structure relates to capacity building. Every evaluation, and especially joint evaluation involving a wide variety of stakeholders, has an element of capacity building that must be effectively handled. The central players in governing and managing the evaluation of the Paris Declaration clearly envisaged using the evaluations governance and management structure to function as a platform conducive to engaging the stakeholders in the design of the evaluation, the methods and approaches to be applied, the ways in which the evaluation teams and the evaluation Core Team were to be selected, and the interpretation of findings through review and discussion of the draft evaluation and study reports and draft final synthesis report in terms of their quality, clarity, and credibility.

Finally, in terms of utility and use, the evaluation’s governance and management arrangements resulted in the timely delivery of the final report, which could be used by aid administrators and policy makers during their preparations for the High Level Forum on Development Effectiveness in Busan.

To sum up, the way in which the governance and management structure for the evaluation of the Paris Declaration was developed resulted in tangible benefits for all stakeholders involved by providing opportunities for a harmonized evaluation process; building ownership, partnership, and evaluation capacities; sharing the burden of the work and the costs; and minimizing the transaction costs for
the individual donors, development agencies, and partner countries involved in this complex undertaking. It also managed to overcome the challenges and difficulties connected with joint evaluations that involve partners with different evaluation cultures and expertise, wishing to pursue different evaluation approaches and methods, and having their own priorities and political agendas.

REFERENCES


Niels Dabelstein was Head of the Evaluation Department of Danida, Ministry of Foreign Affairs of Denmark, from 1988 to 2007. During his tenure the department managed and published some 200 evaluations. In 1994 he initiated and led the Joint Evaluation of the International Response to Genocide in Rwanda. He chaired the Steering Committee comprising 39 donor, UN- and international agencies. As a follow-up he promoted the creation of Active Learning Network for Accountability and Performance in Humanitarian Action (ALNAP) and was the first chairman of the board of the Humanitarian Accountability Project (now the Humanitarian Accountability Partnership). He was also instrumental in initiating and managing the Tsunami Evaluation in 2006. From 1997 to 2002, he was Chairman of the OECD/DAC Working Party on Aid Evaluation and con-

Ted Kliest is senior evaluator in the Policy and Operations Evaluation Department of the Netherlands Ministry of Foreign Affairs, where he has been involved in several joint evaluations. He chaired the Evaluation Management Group for the Joint Evaluation of External Support to Basic Education in Developing Countries and the Evaluation Management Group of the Multi-Donor Evaluation of Support to Conflict Prevention and Peace-Building Activities in Southern Sudan. He is a member of the OECD Development Cooperation Directorate’s Network on Development Evaluation. He was co-chair of the Management Group for the Evaluation of the Paris Declaration on Aid Effectiveness, Phase 2.