THE PARIS DECLARATION EVALUATION PROCESS AND METHODS

Julia Betts
IOD PARC Ltd.
Edinburgh, UK

Bernard Wood
Bernard Wood & Associates Ltd.
Ottawa, Ontario, Canada

Abstract: The Paris Declaration Evaluation faced the challenge of assessing the effects of a policy compact between nations, an area in which evaluation thinking and practice is still emergent. It also confronted the challenge of attempting to link development results to the implementation of a broad reform agenda in countries of widely differing circumstances. This article describes how the Core Evaluation Team, in consultation with the management and governance structures of the Evaluation and the evaluation teams involved, addressed these challenges. It describes the development of the evaluation questions, the methodological framework, and how the methodology worked in practice. Highlights include ensuring the use of evidence through evaluation teams; the analysis and synthesis processes; and processes for ensuring independence, integrity, quality, and ethics. Limitations and risks encountered are discussed. Finally, the article sets out lessons learned for complex transnational studies from the Core Evaluation Team perspective.

Résumé : L’Évaluation de la Déclaration de Paris a confronté le défi de mesurer les effets d’une convention de politiques d’évaluation entre nations, un domaine où la théorie comme la pratique de l’évaluation sont encore émergentes. Il en est de même de la tentative d’établir un lien entre les résultats en termes de développement et la mise en œuvre d’un large programme de réformes dans des pays aux situations très diverses. L’article décrit comment l’Équipe principale d’évaluation a relevé ces défis, en concertation avec les structures de gestion et de gouvernance ainsi qu’avec les équipes d’évaluation concernées. Il relate l’élaboration des questions d’évaluation, le cadre méthodologique, et son fonctionnement dans la pratique, en détaillant notamment la nécessité de garantir l’utilisation de preuves par les équipes d’évaluation, les processus d’analyse et de synthèse, ainsi que les

Corresponding author: Julia Betts, IOD PARC, 16-26 Forth St, Edinburgh, Scotland, EH1 3LH; <julia@iodparc.com>

PURPOSE

This article describes the process and methods of the Paris Declaration Evaluation, conducted from 2009 to 2011. Dabelstein and Patton (2012) and Dabelstein and Kliest (2012) have elsewhere described the preparatory activities, governance, and management arrangements of the evaluation. In particular, the latter have discussed the major efforts undertaken to ensure the evaluation’s independence, credibility, quality, and utility, and to facilitate the engagement of the multiple stakeholder groups involved. Building on these previous discussions, this article describes the evaluation’s basic processes and methods in light of the preparatory work undertaken and working within the governance and management structures and processes of the evaluation. The article will also offer our perspective on lessons for any similar exercises in the future.

With a main expected readership of practicing evaluators, this article is designed to give a working account of the evaluation’s journey in terms of its implementation. This includes establishing and refining the evaluation’s design and developing a robust methodological framework for the evaluation of a highly complex object—a political compact between nations—built on sound evaluation practice. We will discuss the cross-national teamwork within complex governance structures and tackling time and other pressures encountered. We will address the development of the methodological framework, the operationalization of the methodology, and how the methodology worked in practice, whilst highlighting how the use of evidence was ensured through guidance and support to evaluation teams. We will also examine the approaches adopted for analysis and synthesis, how independence, integrity, and ethics were addressed, and how quality assurance was embedded in the evaluation review processes. Limitations and risks encountered will be acknowledged and discussed. Finally, this article sets out our lessons learned for complex transnational studies from the Core Evaluation Team perspective.

et al., 2011b). The full final evaluation report (Wood et al., 2011a) and a summary of findings (Wood & Betts, 2012) present the results of the methods discussed in this article. Additional documents produced during the evaluation process are available online (Paris Declaration Evaluation, 2011).

ESTABLISHING AND REFINING THE APPROACH

As Dabelstein and Kliest (2012) explain, a 2006 Options Paper for the Evaluation had concluded that, subject to some limitations, the Declaration was broadly “evaluable” and that an optimal approach would be a set of four connected but loosely integrated evaluation activities: developing a common framework, country-led evaluation initiatives, a set of thematic case studies across donors, and a medium- to long-term program of analytical work (Booth & Evans, 2006). The same authors also set out how the various component studies (21 country, 7 donor, and a range of supplementary studies) of the overall four-year evaluation exercise fit together. For the Phase 2 study, on which this article concentrates, some aspects of the Evaluation were initially fixed or predetermined. These included:

- **the joint nature of the Evaluation**—it would be conducted across donor agencies and partner countries, and within a governance structure that reflected this.
- **the management and governance structure** at national and international level, which is described by Dabelstein and Kliest (2012) and was devised to ensure the joint nature of the process, as well as the participatory and consultative approach adopted, and solid quality assurance.
- **the voluntary basis of participation**, which meant that inclusion in the sample of countries and agencies was essentially on the basis of self-selection.
- **the primacy of the country** as the main arena for evaluation, to allow a focus on how aid effectiveness reform has played out in practice.
- **for donor headquarter studies**, the application of the same Terms of Reference as those from Phase 1 (carried out over 2007–2008) to enhance comparability.
- **a trilingual exercise**, in order to serve the participants and intended users of the Evaluation as fully as possible, the process was organized to operate throughout in English, French, and Spanish.
The Core Evaluation Team

Phase 2 commenced in September 2009 when the Core Evaluation Team contract was finalized, following a competitive procurement process, to International Organisation Development Limited (IOD PARC), a UK-based development consultancy. Members of the Team included experienced development professionals from seven countries: Canada, the United Kingdom, Nigeria, Denmark, Sri Lanka, Peru, and Malawi. All had expertise in specialist areas related to aid effectiveness.

Regional Workshops: Round 1

To help develop and refine the Evaluation’s approach and methodology, as well as to put into practice the participatory spirit of the process, a set of Regional Workshops were held in Siem Riep (Cambodia), Bogota (Colombia), Lilongwe (Malawi), and Cotonou (Benin) during 2009. Participants for this first round were mainly national evaluation coordinators and members of National Reference Groups. Discussion focused on the Evaluation’s principles, core questions, and approach, as well as on the governance structure, intended process, and quality assurance at the national level. As well as valuable substantive input, the events helped generate common understandings and a common “language” among participants. A sense of coming together on a shared journey was created, and new lines of communication opened up.

Principles Applied

The Evaluation was intended by its designers to be summative and formative. It needed to allow judgements to be made about Declaration achievements in terms of aid effectiveness and development results, whilst also supporting forward-looking policy development and improvement among a wide range of different constituencies and stakeholders.

Drawing on the experience of Phase 1 as well as early inputs such as the Approach Paper, the Core Team identified some key principles to apply, endorsed by the Management Group. These were:

- **Utility:** Above all, the evaluation processes, the findings, and the ways these were reported needed to be credible, rele-
vant, and accessible. This was a fundamental principle of the Evaluation.

- **Jointness and transparency:** Given the large numbers of participating countries and agencies involved in the process, a system of full and balanced engagement was needed to ensure the credibility and utility of results.

- **The critical importance of context:** The centrality of context to aid effectiveness was reflected in its ranking as one of the three main evaluation questions in the Framework for Phase 2, and in the conceptualization of the role and limits of aid and aid reform in overall development processes (below).

- **Comprehensiveness:** The Evaluation was intended to address all five main principles of the Declaration and its 56 specific commitments, since it had been conceived and negotiated as a package.

- **Comparability and relevance:** Phase 1 had shown the importance of a robust common basis for research and analysis if evaluation at an aggregate level was to take place with confidence, but also the need for country-specific issues to be treated.

- **Country ownership:** The centrality of partner countries and their experience within the Evaluation meant ensuring the fullest possible participation of all teams in the Evaluation design and having their agreement on a common approach, language, and methodology across studies.

- **Focus on partnership at country level:** The workings of country-donor partnerships and development outcomes at country level needed to be put at the centre.

- **Knowledge-sharing and capacity development:** The Evaluation aimed to build capacity in part through a process of ongoing sharing of learning, support, peer review, and quality assurance.

These principles helped guide the Evaluation’s conduct throughout, including its governance processes, methodology design, implementation of country and donor studies, and analysis and synthesis.

**Refining the Core Evaluation Questions**

Initial thinking about the Evaluation’s Core Questions had been outlined in the approach paper and refined by the first round of Regional Workshops (see above). Their final formulation was shaped by two main concerns:
1. The *centrality of context* and particularly the context pre-2005, reflected in the need to recognize that the 2005 Declaration itself built on some reform efforts and initiatives that had been underway in different countries well before 2005. A “mountain streams” diagram was developed by the Core Team to illustrate this, and is reproduced as Exhibit 1 in Dabelstein and Patton (2012, this volume).

2. The recognition that *aid is only one part of the many different elements contributing to the processes of development and growth*. Its contribution varies in different contexts according to its scale relative to other important factors (such as other resource flows and drivers or obstacles to development). Exhibit 1, “Aid Reform in Perspective,” reflects this.

Exhibit 1
Aid Reform in Perspective

Once refinement of the Core Evaluation Questions had taken place through the first round of Regional Workshops, they were presented in a draft Evaluation Framework and Workplan for validation by the full 52-member International Reference Group in December 2009. The final Core Questions were as follows:

Core Q1: What are the important factors that have affected the relevance and implementation of the Paris Declaration and its potential effects on aid effectiveness and development results? (“The Paris Declaration in Context”)
Core Q2: To what extent and how has the implementation of the Paris Declaration led to an improvement in the efficiency of aid delivery, the management and use of aid and better partnerships? (Process and intermediate outcomes)

Core Q3: Has the implementation of the Paris Declaration strengthened the contribution of aid to sustainable development results? How? (Development outcomes)

These questions were supported by a number of subquestions (see the Synthesis Report’s Technical Annex for more information, Wood et al., 2011b). Through the same consultative process, a framework for conclusions, intended to guide analysis and drafting of country and synthesis reports, was also finalized.

Defining Terms

From a very early stage, different understandings were apparent around some of the basic working terms of the Evaluation, including “aid,” “capacity,” “social capital,” and “division of labour.” A glossary was consequently developed and disseminated to teams. This applied OECD/DAC definitions where available. Specific guidance on more complex issues such as institutional capacity and social capital and on the use of the term “aid” in the Evaluation, was also made available (see “Ensuring the Use of Evidence: Guidance and Support” below).

DEVELOPING THE METHODOLOGICAL FRAMEWORK: THE USE OF PROGRAM THEORY AND CONTRIBUTION ANALYSIS

The evaluation methodology for Phase 2 had to meet the challenge of assessing the effects of a very unusual evaluation object—that of a broad reform agenda, expressed in a political declaration and which is being applied to both partner and donor countries with widely differing circumstances. Essentially, the Declaration is a policy compact between nations. Rather than the more usual development evaluation object of a project or program, therefore, the Evaluation had more in common with the evaluation of policies or a strategy—areas where evaluation thinking and practice is only beginning to be tested (e.g., Patton & Patrizi, 2010; Jones, 2011). It was agreed to apply a tested evaluation technique for assessing complex change processes.
The Options Paper of 2006 had found that the set of hypotheses that give the Declaration its logic, or program theory, had not been fully articulated. Nonetheless, the evaluation would need to apply a “theory-based” approach—that is, bring to the surface implicit theories of change.

To resolve this challenge, the Core Team returned to the source of the evaluation object, the Declaration text itself. It drew from this the Declaration’s implicit “Program Theory,” in the classic sense set out in Chen (1990). This had set out the underlying logic, including the desired goals expressed (intended outcomes), both intermediate (in improved aid effectiveness) and longer-term (in contributions to improved development results), and the description of how these goals would be generated (programmatic actions). This program theory, illustrated in Exhibit 2, also built on the expected sources in the literature, that is, “prior theory and research, implicit theories of those close to the program, observations of the program, and exploratory research to test critical assumptions” (Donaldson, 2001, as cited in Coryn, Noakes, Westine, & Schröter, 2011).

To situate the program theory in its real-world context, and to highlight the assumptions implicit within it, the Core Team also illustrated some of the “complex pathways to change” from development objectives to development results in a schematic, as Exhibit 3 illustrates. This depicts the many other powerful influences at work on development in different contexts, and the consequent potential and limits on the role of aid in contributing to development results.

By May 2010, a review and collation of key literature sources was also developed, keyed to the three Core Evaluation Questions. This built on an earlier major review, in order to (a) support country and donor evaluations and (b) ensure that the Evaluation was informed by the up-to-date body of relevant research. Digests were produced and circulated, and the review was continually updated thereafter.

Methodological Framework

It was a challenging task to integrate the principles of the Evaluation, the conceptual base, and the approach envisaged in the early inputs into a single methodological framework to respond to the Core Questions. The final Evaluation Methodology recognized that reaching the intended changes of the Declaration—and reflected in its program theory—should be conceived as a journey. This was likely to
### Exhibit 2
The Program Theory of the Paris Declaration on Aid Effectiveness

<table>
<thead>
<tr>
<th>INTENDED DRIVERS</th>
<th>PROGRAMMATIC INPUTS/ ACTIONS</th>
<th>PD OUTPUTS</th>
<th>INTENDED INTERMEDIATE (AID EFFECTIVENESS) OUTCOMES</th>
<th>INTENDED LONGER TERM (DEVELOPMENT EFFECTIVENESS) OUTCOMES</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘Continued high-level political support’</td>
<td>‘Agreed political commitment to change’</td>
<td>56 commitments Deliverables relating to changes in working practice by: Partner Countries Donors Donors &amp; Partner countries</td>
<td>1. “Stronger national strategies and operational frameworks” 2. Increased alignment of aid with partner countries’ priorities, systems and procedures, help to strengthen capacities 3. Defined measures and standards of performance and accountability of partner country systems in public financial management, procurement, fiduciary standards and environmental assessments 4. Less duplication of efforts and rationalised, more cost-effective donor activities 5. Reformed and simplified donor policies and procedures, more collaborative behaviour 6. More predictable and multi-year commitments on aid flows to committed partner countries 7. Sufficient delegation of authority to donors’ field staff, and adequate attention to incentives for effective partnerships between donors and partner countries 8. Sufficient integration of global programmes and initiatives into partner countries’ broader development agendas 9. Stronger partner countries’ capacities to plan, manage and implement results-driven national strategies 10. Enhanced respective accountability of countries and donors to citizens and parliaments 11. Less corruption and more transparency, strengthening public support and supporting effective resource mobilisation and allocation”</td>
<td>‘Increase the impact of aid in: 1. Reducing poverty 2. Reducing inequality 3. Increasing growth 4. Building capacity 5. Accelerating achievement of MDGs’</td>
</tr>
</tbody>
</table>
Exhibit 3
The Context for Implementing the Paris Declaration: Complex Pathways to Change

Increase the impact of aid in:
1. Reducing poverty
2. Reducing inequality
3. Increasing growth
4. Building capacity
5. Accelerating achievement of MDGs'
be travelled by different actors in different ways, in different periods, and at different speeds (depending on their particular contexts, including the pre-2005 starting point and other drivers of development or obstacles). Consequently, the methodological approach focused on assessing the direction of travel toward the intended goals of the Declaration, and then the pace and distance travelled so far.

The evaluation also required a means of addressing the difficult issues of causality and attribution. Paris Declaration implementation is a multidimensional, multilevel process, affected by many factors, which can change direction, emphasis, and pace at different times and in response to many different influences. In finding a framework that could describe any relationship between Declaration implementation and improved development results, the following realities had to be reflected: much of the change process toward aid effectiveness started in many countries before the Declaration was formally signed and implemented (i.e., pre-2005); as Exhibit 3 makes clear, aid is only one of many potential contributory factors to development, and achieving either aid reform or development results is too complex and context-specific a process to be readily reduced to quantitative indicators or relationships; and other features of a context (e.g., governance/political changes/economic crisis/institutional capacity) can have far greater effects on development results than aid reform or aid as a whole.

For all these reasons, traditional linear approaches to evaluation that would aim to causally “attribute” change to the Declaration were not appropriate here. As the Synthesis Report stresses, a political statement cannot by itself cause change; rather, what the Evaluation aimed to research is whether the operational commitments, relevant actors, and motivational elements that it helped bring together from many sources had actually contributed to the intended improvements. It was therefore agreed with the International Reference Group that, owing to the diffuse nature of this Evaluation object, the use of any simple “counterfactuals,” or indeed econometric models, would be methodologically inappropriate. Attempting to extend linear lines of attribution between Declaration implementation and development results (in sectors or in the form of poverty indicators, for example) would be even more flawed.

Consequently, the Core Team opted to steer a course in favour of contribution rather than direct attribution, drawing on the work of Mayne (2001) and others. This was conceptualized as the seeking
out of the “plausible contributions” of the Declaration in individual contexts to development results; the evidencing of any changes and connections observed; and the identification of any other plausible explanations. Further discussion below explains how this was applied in practice.

Finally, the agreed approach and common methodology were clarified to the evaluation’s stakeholders in documents including the Evaluation Framework and Generic Terms of Reference for Country Studies of December 2009, and the Inception Report of May 2010. With the design in place, the next phase of implementation could begin.

OPERATIONALIZING THE APPROACH AND METHODOLOGY

In the next phase of the evaluation, the relatively complex approach and methodology described above needed to be translated into an operational tool that could be applied across the 21 diverse contexts in which country studies would be carried out for Phase 2 (donor studies are described separately below). It had to allow for flexibility, while providing the consistency and commonality needed for synthesis level. Above all, it had to be feasible for country teams to implement. Consequently, the Operational Matrix was developed by the Core Team. This was described as the “spine” of the Evaluation. It was designed to help country teams generate a robust and comprehensive response to the Core Questions, whilst providing a clear common framework for comparative analysis at synthesis level. The full matrix was necessarily lengthy, but Exhibit 4 provides an example section.

Within the matrix, Core Question 1, on context, is explored through a set of proposed indicators and evidence aligned with the subquestions. Core Questions 2 and 3 were more complex in terms of their means of assessment, particularly given likely data paucity issues at country level.

Core Question 2, on the effects of the Declaration on aid effectiveness, had proven a challenge for all the early design and approach work. A major breakthrough, and the turning point in enabling the Evaluation to systematically assess the effects of the Declaration on aid effectiveness, was the proposal by the Core Team in 2009 to accept the Declaration’s own definition of aid effectiveness as the framework for assessment. Specifically, this meant applying the 11 “expected outcomes,” extracted directly from the opening paragraphs
### Exhibit 4
**Example of Matrix Results Logic**

<table>
<thead>
<tr>
<th>PD expected outcomes</th>
<th>Progress markers</th>
<th>Potential indicators of change / milestones</th>
<th>Methods / Forms of Analysis</th>
<th>Judgement on progress, especially since 2005</th>
<th>Key reasons &amp; explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>B Building more inclusive and effective partnerships for development</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>iv. Less duplication of efforts and rationalized, more cost-effective donor activities</td>
<td>• Increased use of donor comparative advantage (relative strengths / complementarity) led by government</td>
<td>Clear views/strategy by Government on donors comparative advantage and how to achieve increased donor complementarity Evidence of reprogrammed aid according to statement of relative strengths</td>
<td>A, C, D</td>
<td></td>
<td>Substantial Some Little None Regression</td>
</tr>
<tr>
<td></td>
<td>• Increased ‘division of labour’ at country / sector level</td>
<td>Mapping process conducted / maintained Number and type/theme of formal Division of Labour arrangements Reprogrammed aid according to Division of Labour agreements / arrangements Co-operative / joint work between agencies within e.g. project modalities</td>
<td>A, C</td>
<td></td>
<td>Substantial Some Little None Regression</td>
</tr>
</tbody>
</table>

*Note.* Letter codes in the Methods/Forms of Analysis column correspond to individual methods proposed for country studies. These are fully listed in Individual Methods section of this article. Those in this excerpt are as follows:
A. Literature and documentation review
C. Survey instruments
D. Interviews and focus groups with key stakeholders

For a full discussion of the matrix, see Wood et al. (2011b, pp. 205–207).
of the Declaration text. This decision had two major advantages: first, the Declaration’s diagnosis of the main problems with aid and the improvements needed had been internationally accepted, as a result of decades of experience. No other useable definition of aid effectiveness had emerged. And second, the ability to use the Declaration’s own explicit statement of intended outcomes as a basis for an evaluation of what it had actually achieved provided the Evaluation with a very legitimate framework.

Core Question 3, on development outcomes, is the main specific location within the evaluation where contribution analysis has been applied. This question had four subquestions, derived from the Declaration’s goals and selected through the regional workshops. These were:

1. Were there contributions to development results in specific sectors, using health as a common tracer sector? (plus one other sector of the country’s choosing)
2. Were there contributions to the prioritization of the needs of the poorest, including women and girls?
3. Were there contributions to increases in social and institutional capacity?
4. Were there contributions to an appropriate mix of aid modalities (e.g., budget support, programs, and projects)?

Contribution analysis was applied through following a three-step sequence of questions: (a) What development results were achieved in this area? (e.g., reduced infant mortality) (b) Did aid contribute? Where and how? (e.g., through donor-funded immunization programs) (c) Did aid reforms plausibly strengthen the aid contribution? (e.g., through the use of sector-wide approaches within the health sector). Teams were asked to place a very strong emphasis on evidencing the pathways of contribution between aid contributions and results, and between aid reforms and the aid contribution. They were also asked to explicitly identify and cite other plausible explanations for results.

To assess direction, distance, and pace of travel, teams were asked to make progress judgements along a broad scale (mainly substantial/some/little/none/regression, with slightly different scales where relevant). A column for key reasons and explanations was included, in order that judgements could be properly explained and evidenced.
The matrix also took into account the OECD/DAC criteria for the evaluation of development assistance, namely relevance, effectiveness, efficiency, impact, and sustainability (see the Synthesis Report’s Technical Annex for a detailed account of the conduct of the evaluation against the full set of OECD/DAC Standards for Development Evaluation [2010]); and commitments under the Accra Agenda for Action, which were included as progress markers/indicators where feasible. Full screening and quality assurance took place for Declaration and Accra Agenda for Action commitments on gender and exclusion.

Finally, the matrix was refined and agreed to through a second series of regional workshops, in consultation with national evaluation teams and national reference groups, as well as the Evaluation Management Group and International Reference Group.

Individual Methods

In advising on specific methods to apply, the Core Team had to strike a balance between providing guidance to ensure consistency and respecting the independence of teams and the diversity of different country contexts. Capacity and resource constraints were also serious concerns. As Exhibit 4 illustrates, the Operational Matrix set out proposed methods against each evaluation question and subquestion, and particularly against the intermediate outcomes for Questions 2 and 3. This included specific methods for assessing contributions to results. To support triangulation and validation, as well as consistency at synthesis level, teams were expected to adopt a multimethod approach, and if appropriate to apply additional methods beyond those identified (in the event, none did). Proposed methods and their associated letter codes in the Operational Matrix included:

A. Literature and documentation review
B. Quantitative/statistical analysis
C. Survey instruments
D. Interviews and focus groups with key stakeholders
E. Stakeholder analysis: especially relevant to Question 1 on context, as well as to the pre-2005 environment. Key types of individual stakeholder/groups/institutions were suggested.
F. Case studies: where relevant to explore or illustrate specific themes or sectors.
G. Additional methodological approaches/forms of analysis: including Outcome Mapping, decision analysis, appreciative enquiry, and Most Significant Change analysis.
Country Survey Tool

To ensure consistency, a common Country Survey tool was developed and piloted, based around the key lines of enquiry of the Matrix. The tool was comprehensive and intended to be used as a basis for teams to develop their own specific instruments. It was widely used, with some teams adapting it to their own specific studies, and others applying it in its entirety.

Donor Headquarter Studies

Donor studies were intended to be conducted in parallel with country studies. A preparatory workshop on Agency HQ Evaluations in September 2009 confirmed that the Phase 2 donor studies would be working to the original terms of reference for Phase 1. The Core Team subsequently developed and proposed a Generic Terms of Reference for Donor Studies, based around the intent of the same terms of reference for Phase 1 plus expansion in line with the methodology for Phase 2. A Donor Headquarter Study Matrix was also developed to operationalize the terms of reference, and also to try to integrate key elements reflected in the country Operational Matrix, to support triangulation. Donors were also requested to provide brief, structured updates to Phase 1 reports.

Supplementary Studies

During Phase 1, a number of thematic studies had been commissioned to deepen knowledge in particular areas of importance for the evaluation. These included work on the “untying” of aid (Clay, Geddes, Natali, & te Velde, 2008); on the applicability of the Declaration in fragile and conflict-affected situations (Oxford Policy Management/the IDLgroup, 2008) and on statistical capacity-building (Oxford Policy Management, 2009). To further expand the evidence base, other supplementary studies commissioned in the second phase of the evaluation included a survey on Paris Declaration implementation across the Latin America and Caribbean region (carried out in collaboration with the Organization for American States) and a study on “Development Resources Beyond the Current Reach of the Paris Declaration” (Prada, Casabonne, & Bezanson, 2010). A number of internal working papers were also developed on existing knowledge on aid effectiveness in situations of fragility, climate change financing, the evolving position and work of civil society organisations vis-à-vis the Declaration and the Accra Agenda, and the knowledge base
on managing for development results. These studies added depth to the substantive knowledge available on issues of aid effectiveness in countries and across themes; they were especially applied at synthesis level.

Ensuring the Use of Evidence: Guidance and Support

Phase 1 had highlighted the importance of clear evidence in substantiating findings. This was a major concern for Phase 2, both for credibility of individual component evaluations (country and donor) and for facilitating robust analysis at synthesis level. The issue was addressed in four main ways. First, recognizing the breadth and scope of the matrix, as well as the complexity of analysis required, a number of guidance papers were produced and disseminated. Topics included the glossary; a clarification of “What is ‘Aid’ in the Paris Declaration?”; guidance on contracting country teams; guidance on issues of attribution and contribution; guidance on questions relating to institutional capacity and social capital, and for specific cases, guidance on handling the Declaration Phase 2 Evaluation in fragile situations. These were all disseminated via the Evaluation’s own Internet site, the Extranet (Paris Declaration Evaluation, 2011).

Second, a specific “Guidance Note on the Use of Evidence” was issued, which asked teams to ensure sufficiency of evidence in their reports, using a specified understanding of the term (Danida, 2006) and weighing the relevance, reliability, and validity of the evidence. This guidance illustrated the potential to use some of the results of the Paris Declaration Monitoring Surveys as forms of triangulation and verification (a separate Guidance Note set out the differences and commonalities between the Monitoring Survey and the Evaluation).

Third, a second round of Regional Workshops was held during 2010 in Wellington (New Zealand), Dhaka (Bangladesh), Tunis (Tunisia), Ha Long Bay (Vietnam), Bogotá (Colombia), and La Paz (Bolivia). Participants at this stage included most of the team leaders of country evaluation teams who had by then been contracted. Discussion focused on the Evaluation Matrix and the particular methods proposed, with valuable suggestions being provided on specific indicators. As well as providing substantive direction, this process continued the participatory process described above, with those involved in the common journey coming together for a second time at a critical milestone.
Finally, but perhaps most importantly, the Core Team was mindful of the parallel aim of the Evaluation to help develop capacity, as well as guiding teams through the relatively complex approach and methodology. A targeted program of support to individual studies was therefore developed. This had to strike the balance between recognizing the autonomy of individual teams, and engaging sufficiently to allow for a free flow of information.

Support therefore centred on the interpretation and use of the Matrix; clarification on the tools, approaches, mechanisms, and methodologies to be applied; and pointing teams toward relevant sources from the literature review that might benefit their evaluations. For the donor headquarter studies, one dedicated Core Team member kept abreast of progress, provided a resource function, and responded to specific requests for advice. For the country evaluations, a primary resource person from within the Core Team took responsibility for providing advice and guidance on the technical aspects of the Evaluation. Comprehensive feedback was issued on the methodologies proposed in inception reports, and on draft reports as they were issued. At least one in-person meeting was held with all evaluation teams except for four, where interaction took place virtually in the first instance.

THE METHODOLOGY IN PRACTICE

Whilst the process of study design had worked smoothly, the implementation phase showed some unsurprising bumps in the road. These were generally minor, but had to be managed, with mitigation strategies as follows.

Country Studies

Most country teams started work from April 2010. By June-September of that year, the majority of studies were well underway. Staggered starts and rates of progress were inevitable, however; one country had completed its report by August 2010, while another was still negotiating national procurement procedures in October of the same year. Country teams managed this unevenness largely through a compressed process where necessary; from the Core Team and Evaluation Management Group perspective, it was handled through extending very slightly the deadline for submission of reports, though conveying firmly to national reference groups the finality of the cut-off point. Any studies received after the final deadline in
January 2011 would simply not be included in the synthesis process. In the event, all studies were submitted by deadline.

Donor Studies

Despite the efforts toward consistency within the country studies described above, parallel consistency proved unfeasible in the implementation of the donor studies. In some cases the planning of particular studies was already well advanced by the time country studies got underway, with the design simply replicating the Phase 1 approach. This limited the potential use of the “mirror questions” suggested within donor studies or even of the Phase 2 Donor Matrix itself. Mitigation measures included the dedicated support function providing guidance to teams on specific areas that the Phase 2 study would like to address and, through the integration of additional research and information at synthesis stage, to try to supplement the gap (see synthesis discussion below).

Methods

Although the Evaluation Matrix and supplementary guidance had provided a wide range of methods for possible use, studies concentrated on those most traditionally used in development evaluation, namely documentation review, including of internal, grey literature; interviews and focus groups with key stakeholders (including civil society and academia); and quantitative/statistical analysis. The survey instrument was widely used, and stakeholder analysis was adopted in some studies, particularly in relation to policy formulation in individual sectors. For case studies, the sector studies under Question 3 were the only ones applied, mostly from a longitudinal perspective. It was noticeable that whilst some inception reports envisaged the use of innovative methods such as Outcome Mapping and Most Significant Change, these had in practice fallen away when draft reports were submitted at Emerging Findings stage.

Guidance and Support

The ongoing process of guidance and support proved unexpectedly valuable from the Core Team and Evaluation Management Group perspective. As well as supporting country teams, it provided ongoing insight into methods being used, challenges arising in applying the matrix (mostly around applying contribution analysis), and the use
of the ratings scales provided (a source of hesitancy for some of the teams). This accompaniment both supported the evaluation process and helped continue the approach of two-way dialogue around which the evaluation process was built.

ANALYSIS AND SYNTHESIS PROCESSES

A key lesson emerging from Phase 1 was the need for clear direction on the synthesis process right from the start. This is particularly important where an evaluation brings together findings across multiple and varied contexts, even where the evaluation framework has been applied specifically to enable cross-case synthesis. For Phase 2, the synthesis exercise faced two main tests:

1. designing a sufficiently robust approach to bring together diverse material from a disparate range of contexts for comparability at aggregate level, and
2. ensuring the validity and reliability of the findings and conclusions at synthesis level.

We now describe how these issues were tackled.

Emerging Findings

To meet the deadline of an Emerging Findings note for validation by November 2010, first draft reports were needed by the end of October 2010. Given the tight timeframe plus some late starts, the Core Team anticipated unevenly developed drafts. This proved the case, with material arriving in highly varied form by the Emerging Findings deadline.

Analytical Process

To analyze the material within the tight timeframe, and to help meet the quality and rigour tests involved, the Core Team prepared a detailed filter template for analysis. Each report was analyzed by designated Team Members along the parameters of the Core Questions and subquestions of the matrix. Simultaneously, to ensure that the findings being sifted out were adequately supported by evidence (another lesson learned from Phase 1), reviewers would both rate the quality of the evidence presented (plus identify type of sources) and assess the clarity of analysis, as follows:
For findings | For conclusions
---|---
Data transparency and coverage | Extent to which questions were answered
Data reliability and accuracy | Clarity of analysis

The combination of a fairly prescriptive matrix, plus the effort devoted to targeted support and guidance, showed clear results in the data emerging at the Emerging Findings stage. A far firmer and more consistent evidence base than was available under Phase 1 was apparent in the promising “data reliability and accuracy” assessments, and reports were geared around the structure of the matrix, including the application of the subquestions and indicators and the use of the Framework for Conclusions.

Initial findings were recorded in filter templates. However, the quality and state of preparedness of material received was highly varied, and many drafts were embargoed for quotation or citation. Also, the ratings scales provided had not been comprehensively applied—only some teams had applied these systematically, with others holding back from judgements.

**Aggregating Emerging Findings Material**

Once complete, each filter was verified/quality assured by another team member. All findings across the evaluations were assembled into a compilation by question and subquestion, and then analyzed and categorized, with the framing of each aggregate finding calibrated by the strength of the underlying evidence. The use of the matrix as the spine of the evaluation was further validated at this stage.

**Emerging Findings Note Validation**

Validation took place through the presentation of the Emerging Findings note, including the main findings (without quotation or citation) to members of the International Reference Group and evaluation teams at a workshop hosted by the government of Indonesia in December 2010. To maximize the time available, especially given the late and partial inputs, plus the embargo constraints on use, a structured process of comment and validation was applied, where individual representatives validated the draft Emerging Findings text against their own reports.
The Core Team then presented for final commentary a refined version of the Emerging Findings note, plus the key limitations emerging, challenges highlighted at this stage (see below), some major themes emerging from the workshop, a list of areas of agreement and disagreement, and some nuances/additions proposed. The workshop succeeded in both validating the evidence to date and helping motivate all teams and coordinators to finalize their reports as quickly as possible. The presence of the meta-evaluators helped solidify the shared sense of direction for the home stretch. The Colombian and South African teams also collaborated to start drawing out some common features from five “middle income” country evaluations, for input to the final synthesis.

Two major course corrections emerged at this stage:

1. The reference to the yardstick of “direction of travel” (rated as positive throughout) was felt to convey too positive an appraisal of progress (although it had been stressed that pace and distance of travel were more significant). A few possible disagreements also emerged, which the Core Team undertook to work through as part of the synthesis process.
2. Donor/agency coverage was clearly insufficient. Further steps would be needed to integrate donor/agency findings of Phase 2 studies and updates into the main Evaluation Matrix, and broaden and deepen coverage of donor aid effectiveness responses to the extent possible.

These challenges were tackled in the synthesis phase, outlined below.

SYNTHESIS REPORT

The Synthesis stage required the Core Team to “[synthesize] the results of all the component evaluations of Phase 2, together with Phase 1 and supplementary materials, in a major policy-oriented synthesis report in time to feed into the [Korea] High Level Forum” (Terms of Reference).

Analysis

Evaluation teams submitted their final reports in January 2011. Without exception, and as at the draft stage, these were structured around the “spine” of the Evaluation Matrix and had applied the indicators and other progress assessment markers suggested. A three-
person group from the Core Team analyzed and synthesised revised reports following the same steps above and applying or re-applying the same steps and controls as in the emerging findings stage. Additional evidence, including that from sources including Birdsall and Kharas (2010) and Knack, Rogers, and Eubank (2010), was reviewed and used where merited to help fill the gaps in donor coverage. This was an especially intensive phase of the evaluation.

By mid February 2011, a composite evidence base had been developed from which key themes could be identified for the synthesis narrative. This evidence base was drawn from the findings from all country evaluations and donor studies against the matrix and provided the main source of evidence for the synthesis report. Bearing in mind the principle that the country should be the main arena for research, the vast weight of the evidence was drawn from country study reports, and particularly those from Phase 2.

Analytical Pathways

Conscious of the interest in identifying possible categories or trends (whether geographical, thematic, or findings-related) in Declaration implementation and results, the Core Team looked carefully for any patterns arising, particularly from the evidence around Core Question 1 on different contexts. Factors considered included region, income and development status, scale of aid flows, length of engagement with the aid effectiveness agenda, extent of Declaration ownership, governance context, experience of natural disasters or political upheaval, among others.

The only two categories to emerge from the evidence as cohesive enough to merit specific treatment were those of “fragility” and “middle-income” status. They were consequently reflected in the final synthesis report. Methodologically, the Evaluation concluded that applying predetermined hypotheses or categories across a broad range of country contexts risks constructing arbitrary and artificial groupings that do not reflect the complex realities of implementing a political compact across a diverse set of nation-states.

Integration of Phase 1 and Phase 2

The Synthesis phase included mapping the findings from Phase 1—especially on aid effectiveness—against the Phase 2 findings, explor-
ing in detail any differences, divergences, or variations, noting any expansions or clarifications, and then revisiting these areas in depth. Phase 1 findings in particular were revisited at the level of donor performance, especially given the paucity of information available from the country studies under Phase 2.

Other Sources

At this stage, additional information was also plotted in from the literature review—particularly supporting evidence surrounding country aid reforms, donor performance, and evidence on specific development trends. The supplementary studies were also heavily drawn upon at this stage, particularly the study on “Development Resources Beyond the Current Reach of the Paris Declaration,” and the working paper on aid effectiveness in situations of fragility.

Report Drafting Process

The drafting process for the main synthesis report involved the following steps:

1. Against the composite evidence base (including Phase 1 studies, Phase 2, and all supplementary forms of evidence), extracting key themes for findings.
2. Tracking back to ensure the logical derivation of the themes from the evidence.
3. Once themes had been verified, drawing these together in narrative form.
4. Once the findings narrative was in place, drawing out conclusions.
5. Tracking back to ensure that conclusions were logically derived from the findings.
6. Distilling key messages and recommendations.

Three other elements, prominent at the November 2010 Emerging Findings workshop, were also reinforced:

1. Emphasizing the pace and distance of change over direction, since it was apparent from Emerging Findings stage that direction was essentially uniform (forward).
2. Rating the degree of difficulty of achieving the different outcomes against specified criteria, since the Emerging Findings discussions brought home even more powerfully than
the reports the varying difficulty in achieving the different intended outcomes.

3. Integrating different starting points as well as degrees of difficulty, since the importance of the pre-2005 context was strongly emphasized in both the material received at Emerging Findings stage and the Workshop itself.

In terms of aggregate judgements on the pace of change and the distance remaining to achieve the intended outcomes for the synthesis report, the Declaration’s authors had clearly understood that these would not be fully achieved in five years. Instead, they specified expected levels of achievement for the selected monitoring indicators. At synthesis level, therefore, the Evaluation’s standard of judgement applied on the intended outcomes was a dual one, blending relative and absolute standards. In relative terms, if reports on some countries or donors showed that they had been able to substantially achieve the end condition in the intended outcome, this was taken as a measure of the possible. The pace and distance remaining in other cases were then assessed accordingly. If there was no such case of basic completion, the assessment applied a standard—partly suggested by the Timeframes for the Monitoring indicators—that if by 2010 the end condition has been at least half achieved, this would implicitly merit a “fast pace” rating and “little” distance remaining.

The final drafting stage involved raising the level of the report, to ensure that the text of the Evaluation was appropriately policy-oriented—adding another level of analysis—in line with the Core Team’s Terms of Reference. Meetings of the International Reference Group, including the Emerging Findings session and a final validation meeting in Copenhagen in April 2011, played a vital role in ensuring that the expected level had been reached.

The draft text, up to but not including conclusions and recommendations, was circulated for wider Core Team commentary and validation in late February 2011. A second round of comments was elicited on a complete draft, including conclusions and main recommendations, in early March. A draft version of the report was then submitted to the Evaluation Secretariat on March 9, 2011.

A full validation meeting of the International Reference Group then took place in Copenhagen in May 2011. This provided commentary and feedback on the quality, credibility, and clarity of the Draft Report. In terms of quality and credibility, the report was judged to have met requirements. However, there were requests for more clarity in
parts of its presentation. The Core Team took note of the comments, and following redrafting, a final version of the report, in the three languages of the Evaluation, was presented to the Management Group on May 16, 2011.

INDEPENDENCE, INTEGRITY, AND ETHICS

The Phase 2 Evaluation’s status as a joint international process meant that independence and integrity were critical to ensuring its credibility. This took place in the following ways:

- Through the Evaluation’s governance processes at both the national and international levels, aimed at assuring the independence of Evaluation reports, described elsewhere in this issue.
- Through screening reports for the underpinnings of independence and transparency as part of the synthesis analysis process.
- Through the application of the OECD/DAC Quality Standards for Development Evaluation, which were emphasized and reported upon throughout at both national and international levels.
- Through the constant focus on a participatory and consultative approach, both within individual studies and at an international synthesis level, as described above.
- Through ensuring transparency, reflected in the sharing of key documents and drafts for scrutiny by external stakeholders at appropriate milestones, supported by the use of a web-based platform (the Extranet) for information-sharing.
- Through a focus on meeting required ethical standards during the conduct of the evaluation, including the recognition of gender considerations and the securing of informed consent of, for example, interviewees through the assurance of anonymity and confidentiality, and so on.
- Through the integration in the Evaluation Framework and Matrix of aspects of gender equality, human rights, and social inclusion, principles that are prominent in both the Declaration and the Accra Agenda for Action.

QUALITY ASSURANCE

Quality assurance was given considerable explicit attention throughout the process of the Phase 2 Evaluation. A paper setting out the relevant provisions was approved by the International Reference
Group and disseminated in May 2010. This clearly distinguished at all levels that whilst approvals of the framework documents for the evaluation would be required, individual evaluation reports would be reviewed and commented upon for quality, credibility, and clarity, but would not be subject to approval, in order to safeguard the independence of the evaluation. Following is a summary of how the quality assurance arrangements were handled at various levels.

At the country level, each evaluation process was required to establish internal quality assurance and control systems. The National Evaluation Coordinator was responsible for quality-assuring evaluation reports before submission to the Core Team for inclusion in the synthesis. The terms of reference for the National Reference Group explicitly stated that the quality should be assessed against national, regional, or international Evaluation Quality Standards.

In the Donor Studies, Evaluation Coordinators were responsible for quality-assuring the study reports before submission to the Core Team for inclusion in the synthesis. The Core Team was tasked to support the quality assurance at national/donor level by engaging with and supporting the country evaluations and, if requested, the donor studies. Quality assurance of the work of the Core Team had three levels:

1. Internal quality control and assurance of the team’s processes through the appointment as Quality Manager of a senior staff member of the contracted institution (IOD PARC) from outside the project team.
2. Quality assurance by the Evaluation Management Group, who were tasked to review all products by the Core Team and to assess and sign off on quality.
3. Quality assurance by the International Reference Group, who were tasked to oversee evaluation products, including the key documents produced by the Core Team.

These arrangements were comprehensive and are considered by the parties involved to have worked well.

LIMITATIONS AND RISKS

The main limitations and risks of the evaluation were recognized from an early stage of the process. They were specified in the Inception Report and confirmed at the Emerging Findings workshop in Indonesia. They included:
• The unusual character of such a broad reform program and political declaration as an evaluation object and the resulting limits on applying standard evaluation methodologies that imply more linear causality. This was tackled through the application of a theory-based approach and the use of contribution analysis.

• The breadth and complexity of the goals of the Declaration and the wide variety of contexts and actors involved, meaning that very few robust analytical categories, groups, or trends were found to apply. Those that were relevant were identified, and analysis presented accordingly.

• The paucity of data in some areas, including the limited participation by multilateral actors that has meant that very few robust conclusions could be drawn in relation to them. This was not mitigated, and this data paucity was reflected in the final analysis.

• The limited time elapsed since the Declaration was endorsed in 2005 (and even less since Accra in 2008), meaning that little scope existed for some of the fundamental changes expected to have been implemented. As above, contribution analysis enabled those changes that had arisen to be captured.

• The expectation of breadth and comprehensiveness needing to be balanced with rigour and depth: the matrix was broad and extensive, and it tested considerably the capacity of evaluation teams. Guidance and targeted support played a major role here.

• The self-selected nature of participating countries and agencies, which limited the representative basis of the component studies—the distribution among partner countries was in the end reasonably representative.

• The eventually unrealized hope of achieving sufficient coverage of donors’ and agencies’ policies and actions on the ground through the country evaluation reports, resulting in limited intersections between donor headquarter studies and country evaluations. Additional sources had to be applied here to triangulate as far as feasible.

• The fact that the limited methodology for donor/agency headquarter studies from Phase 1 was carried over into Phase 2—and that a number of key donors did not carry out such studies at all—also meant that the evaluation lacked equally targeted instruments to assess donor/agency performance as those for country evaluations, meaning, as
above, that supplementary sources had to be used to narrow the gap.

- Some significant timing failures and delays, particularly around evaluation team procurement and the approval processes for both country and donor reports. Although the Core Team had anticipated this eventuality and proposed mitigation measures of special catch-up support where needed, material/reports were inevitably delayed both at Emerging Findings and Synthesis stages. This meant that the Core Team was working at various milestone points with material at very different stages of development, resulting in an especially intensive synthesis phase.

LESSONS LEARNED BY THE TEAM

The experience of the Paris Declaration Evaluation has provided some valuable lessons on the complexities of conducting multi-country and multi-agency studies at the level of a political declaration and program of change. Other articles in this special issue will yield further lessons for future such evaluations, but some of the most critical identified from the Core Team’s perspective are summarized here.

1. Evaluators and managers need to campaign for evaluable frameworks of intended outcomes in the upfront design of programs and policies, but remain wary of crude and oversimplified indicators for such complex changes. Accept and embrace the need for rigorous qualitative evaluations of complex realities.


3. Keep the working language as clear and non-technocratic as possible, minimizing jargon—especially, but not only, in multilingual and multicultural evaluation processes. Carry this through to reports to maximize ultimate dissemination and use.

4. Recognize genuinely participatory design and validation as not just desirable but integral to the ownership of the process and the ultimate quality and utility of the evaluation. Build in the “careful planning, structure, execution, and facilitation” implied.
5. **Recruit highly competent teams** early to play a major role, together with evaluation managers and stakeholders, in design.

6. *(Perhaps) be prepared to impose selectivity,* even among vital questions, in order to have a manageable challenge across the body of cases.

7. **Prepare for complex and uneven processes in multi-site evaluations** but set and keep the deadlines necessary to maintain momentum and deliver timely results.

8. **Expect uneven capacities and delivery among varied teams,** even while working to strengthen them. Be ready to reinforce, but if necessary abandon or sideline results where they are found weak against transparent standards. Ensure in advance that an adequate base will remain for reasonable overall validity after some “dropouts.”

9. Recognize that **written component and synthesis reports are only part of the contribution of the evaluation,** alongside benefits from the process and building a community of shared understanding and trust.

10. Set and consistently apply **rules to protect teams’ independence** within agreed evaluation frameworks and arrangements for quality assurance and validation.

11. **Be realistic about the candour to be expected** in assessments of other actors’ performance as well as self-assessments.

12. **Calibrate the strength of particular synthesis findings, conclusions, and recommendations** according to the relative strength of evidence in the body of cases.

**CONCLUSION**

As noted in the introduction, the approach and methodology for the Paris Declaration Evaluation faced the challenge of assessing the effects of a high-level policy compact between nations, an area where evaluation thinking and practice is only just beginning to be tested. We have reviewed the development of the evaluation’s design and methodological framework, and discussed the cross-national team-working within complex governance structures and tackling time and other pressures encountered. We have discussed the operationalization of the methodology and how the methodology worked in practice while highlighting how the use of evidence was ensured through guidance and support to evaluation teams. We have examined the approaches adopted for analysis and synthesis; how independence, integrity, and ethics were addressed; and how quality assurance was
embedded in the evaluation review processes while facilitating the engagement of the multiple stakeholder groups. Limitations and risks encountered have been acknowledged and discussed. Finally, we set out lessons learned for complex transnational studies from the Core Evaluation Team perspective.

In all of these elements, the Core Team and the Evaluation’s management and governance structures strove overall to ensure the Evaluation’s independence, credibility, quality, and utility. Ultimately, of course, methods are not an end in themselves but a means to producing high quality findings that are credible, useful, and relevant to the policy debates they seek to inform. The results of the Paris Declaration Evaluation are summarized in Wood and Betts (2012), quality and credibility are assessed in Patton (2012) and Dabelstein and Patton (2012), whilst the utility of the Evaluation’s findings is reviewed in Dabelstein and Wood (2012).

REFERENCES


Paris Declaration Evaluation. (2011). *Background documents.* Retrieved from [http://www.oecd.org/document/60/0,3343,en_21571361_34047972_38242748_1_1_1_1_00.html](http://www.oecd.org/document/60/0,3343,en_21571361_34047972_38242748_1_1_1_1_00.html)


Julia Betts, Ph.D., is a principal consultant at IOD PARC Ltd, an international development consultancy based in Edinburgh, UK. Formerly of the UK’s Department for International Development, she worked from 1996 to 2002 in countries including India, Malawi, and El Salvador. A consultant since 2003, Julia specializes in the evaluation of aid and development programs, policies, and strategies. Her work focuses particularly on the development of methodologies for complex or challenging evaluation objects, and on the role of gender, equity, and exclusion issues within development cooperation. Recent assignments have included assessments of development assistance in fragile and conflict-affected states, and reviewing donor use of new modalities in the changing aid environment to address policy concerns. Julia was a Core Team Member of the Paris Declaration
Evaluation Phase 2 and, jointly with Bernard Wood, co-wrote the Synthesis report.

**Bernard Wood** has over 40 years of experience as an evaluation specialist, senior multilateral manager, think tank director in development and security policy, and parliamentary advisor. He has worked in more than 60 countries. In the 1980s he pioneered taking evaluations of official programs of development assistance to the critical country level and deepening assessments of civil society assistance. He led the international core team for one the largest ever development evaluations—the Evaluation of the Paris Declaration. He has led some 50 complex evaluations, reviews, and organizational assessments at different levels. From 1993 to 1999 he served as the Director of Development Cooperation and head of the DAC secretariat at the OECD in Paris. Earlier he was Deputy Director of the Parliamentary Centre, then founding CEO of the North-South Institute for 13 years, and CEO of the Canadian Institute for International Peace and Security. He twice served as Personal Representative of the Prime Minister of Canada in the fight against apartheid. He holds an M.A. from the School of International Affairs at Carleton University, Ottawa (International Development). In 1992–93 he was a Center Fellow at the Center for International Affairs at Harvard University.