

LESSONS LEARNED AND THE CONTRIBUTIONS OF THE PARIS DECLARATION EVALUATION TO EVALUATION THEORY AND PRACTICE

Niels Dabelstein
Danish Institute for International Studies
Copenhagen, Denmark

Michael Quinn Patton
Founder and Director, *Utilization-Focused Evaluation*
Saint Paul, Minnesota, USA

Abstract: The final event of the Paris Declaration Evaluation was a lessons-learned workshop. This article first highlights the lessons about joint evaluations identified by participants in that workshop with the resulting report being a model of how to bring closure to a major evaluation. The article then presents 10 contributions of the Paris Declaration Evaluation to evaluation theory and practice. The article closes by recognizing that the Paris Declaration Evaluation received the 2012 American Evaluation Association (AEA) Outstanding Evaluation Award, which noted, “The success of the Evaluation required an unusually skilled, knowledgeable and committed evaluation team; a visionary, well-organized, and well-connected Secretariat to manage the logistics, international stakeholder meetings, and financial accounts; and a highly competent and respected Management Group to provide oversight and ensure the Evaluation’s independence and integrity. This was an extraordinary partnership where all involved understood their roles and carried out their responsibilities fully and effectively.”

Résumé : Le processus d'évaluation s'est achevé par un atelier consacré aux enseignements tirés. Cet article souligne les leçons identifiées par les participants concernant les évaluations conjointes, le rapport final étant un modèle de clôture d'une évaluation importante. Il présente ensuite les 10 contributions apportées par cette évaluation à la théorie et à la pratique en matière d'évaluation. Pour finir, il mentionne le prix d'excellence accordé en 2012 par l'American Evaluation Association, signalant : « Le succès de l'Évaluation de la Déclaration de Paris [. . .] a nécessité une équipe d'évaluation exceptionnellement chevronnée, compétente,

et totalement engagée; un Secrétariat visionnaire, bien organisé et bien réseauté afin de gérer la logistique, les réunions internationales des intervenants, et la comptabilité; et un Groupe de gestion hautement compétent et respecté pour assurer le suivi, ainsi que l'indépendance et l'intégrité de l'évaluation. De ceci a découlé un partenariat extraordinaire où tous les acteurs impliqués comprenaient leurs rôles, prenaient leurs responsabilités de manière complète et efficace. »

The Canadian Evaluation Society (CES) has pioneered “the CES Credentialed Evaluator (CE) designation [which] is designed to define, recognize and, promote the practice of ethical, high quality and competent evaluation in Canada through a program for professional designations” (CES, 2012). The Credentialed Evaluator must demonstrate competency in five domains: Reflective Practice, Technical Practice, Situational Practice, Interpersonal Practice, and Management Practice. It is this first area of competence that is the focus of this article.

Reflective Practice competencies focus on the fundamental norms and values underlying evaluation practice and awareness of one's evaluation expertise and needs for growth. (CES, 2010, p. 2)

REFLECTIVE PRACTICE

The final activity of the Paris Declaration Evaluation (PDE) was a lessons-learned workshop held in Paris in February 2012 to debrief the evaluation and generate lessons about the conduct of Joint Evaluations to inform future such endeavours and contribute to the evaluation profession's knowledge base about this type of evaluation. The workshop was held on the site where the Paris Declaration was endorsed in 2005. Over 50 participants engaged in reflective practice. They belonged to three categories: (1) national coordinators for the PDE studies in individual countries; (2) members of the PDE Management Group and the PDE International Steering Committee, and the PDE Core Evaluation Team; and (3) a number of independent consultants and experts with experience from large joint evaluations other than the Evaluation of the Paris Declaration.

In her opening speech at the workshop, Delphine D'Amarzit, the Assistant Secretary, Multilateral Affairs, Trade and Development Policies Department, Treasury General Directorate, of the French

Ministry of Finance, paid tribute to the work accomplished by conducting the Paris Declaration Evaluation and stressed how important it was to have completed this evaluation in time to inform the preparations for the Busan High Level Forum on Aid Effectiveness that had been held in November 2011. She also praised the extensive outreach and dissemination efforts made by those engaged in the Paris Declaration Evaluation as being critical to facilitating the use of the evaluation results.

The lessons-learned workshop included a number of presentations from evaluators with joint evaluation experience and commissioners of joint evaluations. Three themes surfaced across those presentations: (1) the importance of preparation for the evaluation, especially facilitating buy-in; (2) building trust and confidence; and (3) organizing the evaluation process to foster evaluation use. Here are excerpts from the lessons-learned report elaborating these three themes.

Preparation and Buy-In

It should go without saying that thorough *preparations* are important in all evaluations. Proper preparation is even more essential for joint evaluations where several, sometimes diverging interests must be met and where the background, knowledge, and capacity of the various partners engaged in the evaluation may differ considerably.

Managing *tensions* is part of joint evaluation work. Typical lessons include:

- Consult all stakeholders—to ensure buy-in on scope, agree on management structure, and establish common ground rules.
- Seek consensus—but do not compromise the integrity of the evaluation process and product.
- Have a strong and effective lead but also show openness and readiness to listen.
- Ensure evaluator independence to guarantee the integrity and credibility of the evaluation findings even while engaging diverse stakeholders to ensure relevance and use of the evaluation.
- Take the time necessary, but equally important, respect timelines in order to make a joint evaluation work.

Building trust and confidence

Another theme was the need to build confidence among the stakeholders commissioning large-scale joint evaluations. It must be acknowledged that experiences and points of view may differ, and agreement must be reached on all principal issues while it may not be necessary or possible to agree on every little detail. Thus authority to work out the details may be delegated to a smaller group of stakeholders tasked with managing the evaluation (e.g., an evaluation management group).

A way to gain trust is, 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.

Evaluation process and use

A third theme was the need to be *process oriented*. It is important to allow ample time for preparation in order to allow sufficient buy-in and understanding of what the evaluation is about as well as time for continuous consultations throughout the evaluation. The attention to use should be present during the whole process in order to focus the work and justify the investment in time and money. One aspect of the process is that every joint evaluation has an element of *capacity building* and this must be recognised and effectively handled. (Paris Declaration Evaluation Secretariat, 2012, pp. 2–4)

These three themes give some flavour of the nature of the reflective practice deliberations. A lessons-learned report was generated from the workshop and widely disseminated as part of the core commitments to transparency and use (Paris Declaration Evaluation Secretariat, 2012). On the final day of the workshop, participants reiterated two main conclusions that stood out from the engaged discussions:

One was that joint evaluations are useful and worthwhile especially for assessing complicated policy-related changes that span several countries, in spite of the often considerable investment required in time and money.

The other was that the broad network of evaluators and evaluation managers established during the implementation of the Paris Declaration Evaluation (PDE) and the accumulated experience of the network members should be preserved in one way or another after the finalization of the PDE and the closing down of its Secretariat. Such a network would be particularly well suited to promote and facilitate joint evaluations. (Paris Declaration Evaluation Secretariat, 2012, p. 1)

This lessons-learned workshop and report is a model of how to bring closure to a major evaluation. (As with all other documents related to Evaluation of the Paris Declaration, this lessons-learned report has been posted on the OECD/DAC website: Paris Declaration Evaluation Secretariat, 2012.) The remainder of this article presents and discusses contributions of the Paris Declaration Evaluation to evaluation theory and practice.

CONTRIBUTIONS OF THE PARIS DECLARATION EVALUATION TO EVALUATION THEORY AND PRACTICE

There are 10 reasons why the Evaluation of the Paris Declaration should be relevant to any evaluation practitioner, theorist, trainer, scholar, professional, policy maker, or funder.

1. The evaluation's substance is of *global importance*.
2. *Joint evaluation* is a specific type of and approach to evaluation that has significant and widespread implications for how evaluations are both funded and conducted.
3. *Evaluating a set of principles* offers important insights into new and emergent objects of evaluation (evaluands) beyond the traditional focus on projects and programs.
4. The evaluation demonstrates a rigorous approach to *synthesis analysis* in which a number of separate evaluations are examined and findings integrated to generate general, triangulated conclusions.
5. The evaluation exemplifies *authentic engagement of a large number of diverse stakeholders* in ways that increased the evaluation's credibility and utility.
6. The governance structure, stakeholder processes, management mechanisms, and methods exemplify how to *maintain evaluator independence* while undertaking a highly political and high stakes evaluation.

7. The evaluation adheres systematically, intentionally, comprehensively, and visibly to the new international standards for evaluation quality (Development Assistance Committee [DAC], 2010), demonstrating how to take a *standards-based approach* to the conduct of an evaluation.
8. The evaluation demonstrates multiple dimensions of *process use*, and how and why process use can be as important as findings use.
9. This is an *exemplar of utilization-focused evaluation*, that is, making attention to intended use by intended users the focus of the evaluation from beginning to end.
10. The evaluation demonstrates the role of *meta-evaluation* to enhance credibility, accountability, and utility.

We'll briefly preview each of these contributions to evaluation practice and link the items to the specific articles in this special issue where each contribution was discussed.

1. The Evaluation's substance is of *global importance*.

The Paris Declaration on Aid Effectiveness was endorsed in 2005 by over 150 countries and organizations including (a) the more developed aid-donor countries such as the United States, Canada, European countries, Australia, and New Zealand; (b) developing countries from around the world, including larger countries such as Indonesia, Nigeria, and Argentina as well as smaller countries like the Solomon Islands, Jamaica, and The Gambia; and (c) international development institutions such as the World Bank, the United Nations Development Group, and the Organisation for Economic Co-operation and Development (OECD). Thus, regardless of where you live and work, whether in an aid-giving country or an aid-receiving country, the effectiveness of development aid affects all of us. It is a matter of general global significance for anyone who cares about and is paying attention to what is going on in the world. All readers of the *Canadian Journal of Program Evaluation* would seem likely to be included as among those who are interested in such matters of global import.

Moreover, as the evaluation report notes in establishing the context for and significance of the Paris Declaration, it was considered a landmark international agreement and the culmination of several decades of attempts to improve the quality of aid and its impact on development. That history is traced in the opening article on the history and significance of the Paris Declaration (Dabelstein & Patton, 2012).

2. *Joint evaluation* is a specific type of and approach to evaluation that has significant and widespread implications for how evaluations are both funded and conducted.

The Paris Declaration Evaluation was organized and conducted as a joint international initiative. *Joint evaluations* involve multiple agencies and diverse international participants. Traditionally, international aid agencies such as the Canadian International Development Agency (CIDA) and USAID commissioned and conducted their own individual agency evaluations, and that remains the dominant form of evaluation. But for major initiatives that involve multiple donors and recipients acting collaboratively, this new form of *joint evaluation* has emerged. Dabelstein and Kliet (2012) discuss the merits and challenges of joint international evaluations and note that there is not universal agreement on the benefits of joint evaluations.

The first major global joint evaluation involving more than 40 donors, agencies, and civil society organizations was the Joint Evaluation of Emergency Assistance to Rwanda. The 1994 genocide and the ensuing relief operations provoked an unprecedented international collaborative evaluation process. The Joint Evaluation was initiated under the auspices of the Danish government's aid agency, Danida, but ultimately involved a number of other international agency participants. (For details about and a review of this Joint Evaluation, see Danida, 1996; Borton, 2004.) In 2008 the Active Learning Network for Accountability and Performance in Humanitarian Action (ALNAP) published a meta-evaluation of 18 joint evaluation reports entitled *Meta-evaluation: Joint evaluations coming of age? The quality and future scope of joint evaluations* (Beck & Buchanan-Smith, 2008). The report opens with commentary on the international trend toward joint evaluations:

[T]here has been a growing trend towards “jointness” in the aid world, and joint evaluations (JEs) of humanitarian action. Originally championed by donor governments, joint evaluations have been the focus of recent and growing interest and engagement from NGOs and UN agencies. In 2005/06 the second-ever systemwide evaluation of humanitarian action took place, of the international response to the Indian Ocean tsunami of 2004, through the Tsunami Evaluation Coalition (TEC)....

The sector is on a steep learning curve in terms of how to do joint evaluations, including how best to manage and organise them, when they are appropriate, and with whom. This is accompanied by an active debate about the pros and cons of joint evaluations, and how they relate to single-agency evaluations—for example, can they replace them? (Beck & Buchanan-Smith, 2008, pp. 84–85)

This relatively new type of evaluation poses particular challenges and, when successful, offers potentially significant advantages not only for international evaluators but also for domestic evaluations that could involve joint endeavours across ministries, or across government agencies, nonprofits, universities, and philanthropic foundations. The Paris Declaration Evaluation, as an exemplar of a joint evaluation, can serve as a potential model for this type of evaluation generally, including a wide variety of collaborative and multi-partner evaluations.

We would go even further and suggest that the joint evaluation model constitutes a “prescriptive” theory.

Evaluation theories offer active propositions that have an “in order to achieve *x*, do *y*” structure, and an organized set of such propositions constitutes a prescriptive theory. (Christie, 2012, p. 9)

Prescriptive theories deal with values and what one believes to be the best course of action. In that regard, joint evaluation constitutes a prescriptive theory for organizing, funding, and implementing complex, multi-stakeholder, high stakes evaluations.

3. *Evaluating a set of principles* offers important insights into new and emergent objects of evaluation (evaluands) beyond the traditional focus on projects and programs.

The profession of evaluation has been heavily oriented toward evaluation of projects and programs, but the Paris Declaration is neither. It is a set of principles and political commitments. As the Evaluation Final Report notes:

The object of the Evaluation—an agreed set of principles and commitments to improve aid effectiveness—is not a project or programme, the more normal objects of

development evaluation. In a broad sense it is more like a strategy, a domain where evaluation is beginning to be tested, but the Declaration campaign has less-clear boundaries than most strategies. Interesting common elements can also be found in the growing experience in evaluating policy influence. (Wood et al., 2011, p. 3)

This broad and innovative approach to the *evaluand*, the unit of analysis for evaluation, is part of what makes this evaluation an exemplar of its type. In a *New Directions for Evaluation* volume on *Evaluating Strategy* (Patrizi & Patton, 2010), the editors and authors argued that it is critically important that the profession demonstrate the relevance of *evaluative thinking* to the full variety of types of interventions aimed at social betterment. Evaluating the implementation of formally agreed-on principles is one such new frontier—and this evaluation exemplifies and provides a model for how to undertake an evaluation of such a complex evaluand.

How large is the scope of potential evaluations of international agreements? The University of California (2006) maintains an archive of treaties and international agreements that number in the tens of thousands. The importance and challenges of monitoring compliance with international declarations was the focus of a presentation and roundtable discussion at the African Evaluation Association conference in Niamey, Niger, in 2007. The session began by discussing ways of monitoring and evaluating the Solemn Declaration on Gender Equality in Africa and the Protocol on Women's Rights in Africa and then broadened the discussion to all kinds of international agreements (Campbell-Patton & Patton, 2007). The organizing questions were:

- What is the role of civil society and NGOs in the process of monitoring and evaluating progress made with regards to declarations and protocols that are not legally binding and have few mechanisms to ensure accountability by governments or international agencies?
- What role does monitoring and evaluating play in making governments and international agencies more accountable for their international commitments?

Participants agreed that a crucial area of development is adherence to international commitments. However, evaluation had not yet developed methods, mechanisms, and legitimate roles for monitoring

progress made in complying with and implementing declarations and protocols that are not legally binding and have few mechanisms to ensure accountability. This was viewed by the participants as a critical area for developing evaluation in the future, especially in an African context.

The Evaluation of the Paris Declaration points the way in this new and crucial evaluation frontier. The methodological challenge was how to evaluate a political statement. The political challenge was how to execute the evaluation in accordance with the Paris Declaration Principles by engaging donors and developing countries as equal partners. The Evaluation derives its significance and relevance from the importance of the Paris Declaration Principles. The Paris Declaration Evaluation was and remains, itself, a landmark and groundbreaking evaluation of a different kind of evaluand.

4. The Evaluation demonstrates a rigorous approach to *synthesis analysis* in which a number of separate evaluations are examined and findings integrated to generate general, triangulated conclusions.

The overall purpose of the Evaluation was to assess the relevance and effectiveness of the Paris Declaration and its contribution to aid effectiveness and ultimately to development effectiveness. The evaluation of the implementation of the Paris Declaration was carried out in two phases (2007–2008 and 2010–2011). Together the two phases comprised 22 country-level evaluations that were designed within a common evaluation framework to ensure comparability of findings across countries while allowing flexibility for country-specific interests. Each of these evaluations was conducted by independent evaluation teams managed by the respective partner country under the oversight of a national reference group and with technical support from the core evaluation team and the PDE Secretariat. The country-level evaluations were supplemented by 18 donor and multilateral development agency studies that assessed how the Paris Declaration was represented in the policies, strategies, and procedures of these donors and agencies. These studies mainly consisted of document reviews supplemented by interviews with key actors at headquarters level and in field offices. The studies were conducted by independent teams managed by the respective agencies' evaluation departments. A literature review on aid effectiveness theory and practice, particularly at country level, was conducted and used as a knowledge base to inform the evaluation questions, design, and data collection framework. In addition several thematic studies were commissioned

covering diverse subjects such as the Developmental Effects of Untying of Aid; Support to Statistical Capacity Building; the Applicability of the Paris Declaration in Fragile Situations; Development Sources Beyond the Current Reach of the Paris Declaration; and the Relationship Between the Paris Declaration, Aid Effectiveness, and Development Effectiveness.

The final evaluation process and report provided a synthesis of all component evaluations and thematic studies, including the Phase 1 evaluation. As a fully Joint International Evaluation, the Final Report summarized, analyzed, and synthesized more than 50 studies. Part of what makes this an evaluation exemplar is the exceptional clarity, coherence, and relevance of the final report in integrating and synthesizing mountains of evidence from diverse sources and methods to systematically address complex evaluation questions. The evidence is artfully summarized and cross-referenced across questions and carefully linked to the resulting conclusions and recommendations. (For details see Wood & Betts, 2012; Betts & Wood, 2012.)

5. The Evaluation exemplifies *authentic engagement of a large number of diverse stakeholders* in ways that increased the evaluation's credibility and utility.

No tension more deeply permeates evaluation than the admonition that evaluators should work closely with stakeholders to enhance mutual understanding, increase relevance, and facilitate use while maintaining independence to ensure credibility. Managing this tension was at the centre of the way the Evaluation of the Paris Declaration was structured, administered, and governed.

- At the country level, a national reference group in each country selected the evaluation team, based on published terms of reference and competitive processes; country evaluation teams were to operate independently in carrying out the evaluation.
- At the international level, the Secretariat administered the evaluation, with an Evaluation Management Group that selected the core evaluation team (through open international tender) and provided oversight and guidance, while the Core Evaluation Team conducted the evaluation independently.
- The International Reference Group (IRG) of country representatives, donor members, and international organiza-

tion participants provided a stakeholder forum for engaging with the Core Evaluation Team and the evaluation's draft findings.

Thus, at every level, the evaluation was structured to separate evaluation functions from political and management functions, while also providing forums and processes for meaningful interaction. Such structures and processes are fraught with tensions and risks. Yet, in the end, the credibility and integrity of the evaluation depends on doing both well: engaging key stakeholders to ensure relevance and buy-in while maintaining independence to ensure the credibility of findings.

In the International Reference Group meetings, the collaborative process allowed input on and reactions to all aspects of the synthesis. In its December 2010 meeting in Bali, IRG members had an opportunity to react to emerging findings. The Core Evaluation Team responded systematically to how they would deal with the feedback they received, including presenting appropriate changes in language to reflect the responses and reactions offered. This process continued and intensified at the IRG meeting in Copenhagen in April 2011. Every chapter of the draft report was reviewed by small groups. The groups reported their reactions in plenary. The Core Evaluation Team responded to the feedback they were offered, being clear about what changes they could make based on the evidence, and always expressing appreciation for the group's input and reaffirming that, in the end, the Core Evaluation Team would determine and own the findings, conclusions, and recommendations. An example of how this played out in practice occurred near the end of the Copenhagen IRG meeting when a pragmatic solution was negotiated that allowed one final review of the report by IRG members while ensuring the evaluation report could be produced on time with its integrity and independence maintained.

Members of the International Reference Group were heard. Their inputs were taken seriously. And in the end the Core Evaluation Team took responsibility for the findings, conclusions, and recommendations based on the team's reading of the evidence. The independent meta-evaluation (Patton, 2012) affirmed the accuracy of how the Core Evaluation Team described the collaborative process at the beginning of the Final Report:

As the Synthesis Report and its Technical Annex show, this large and complex joint international evaluation

has been a fully transparent and participatory exercise throughout. The designers and champions of the process deserve great credit for being prepared to act in the spirit of the Paris Declaration ideals. There is ample evidence that the participants have invested and benefited greatly, learning and building together. The intensive collaboration has extended from the earliest design and regional workshops all the way through to the final improvements on this Synthesis Report. (Wood et al., 2011, p. x)

Country ownership by partner countries was made a priority to ensure full participation and engagement. Country representatives participated in designing the evaluation matrix. Quality control was based on peer reviews and international participant feedback. Providing sufficient support to make the evaluation a trilingual exercise—English, French, and Spanish—was aimed at ensuring full participation in and access to all aspects of the evaluation.

Balancing evaluator independence and stakeholder engagement

The Paris Declaration Evaluation demonstrates that assuring evaluator independence and genuine stakeholder involvement can be mutually reinforcing in support of enhanced evaluation quality and credibility. Stakeholder involvement assures the relevance of evaluations, and evaluator independence ensures credibility. But stakeholder involvement is sometimes seen as undermining independence while processes to assure independence often limit stakeholder involvement. The Evaluation of the Paris Declaration has demonstrated how to strike the appropriate balance by creating structures and processes that provide meaningful and authentic stakeholder involvement without impinging on evaluator independence. Clearly delimited roles, skilled facilitation, and shared commitment to high-quality evidence on all sides are factors that support balance and enhance the potential for use. A degree of ongoing, mutual monitoring between the international and national governance structures and evaluation teams contributed to establishing a contextually appropriate balance between evaluator independence and stakeholder involvement, with early intervention when potential difficulties arose.

6. The governance structure, stakeholder processes, management mechanisms, and methods exemplify how to *maintain evaluator independence* while undertaking a highly political and high stakes evaluation.

As noted at the beginning of this article, the Canadian Evaluation Society has pioneered the CES Credentialed Evaluator (CE) designation, which “is designed to define, recognize, and promote the practice of ethical, high quality, and competent evaluation in Canada through a program for professional designations” (CES, 2012). The Credentialed Evaluator must demonstrate competency in five domains: Reflective Practice, Technical Practice, Situational Practice, Interpersonal Practice, and Management Practice. It is this last area of competence that is often overlooked but is clearly critical, especially for such a large, multi-faceted, global evaluation as that of the Paris Declaration.

Management Practice competencies focus on the process of managing a project/evaluation, such as budgeting, coordinating resources, and supervising. (CES, 2010, p. 2)

The importance of competent management as a critical factor in high-quality evaluations has been documented in research on essential competencies (Patton, 2012, Chapter 2; Stevahn, King, Ghere, & Minnema, 2005). That’s why a full presentation and review of management and governance issues has been included in this special issue of the *CJPE* (Dabelstein & Kliet, 2012). The participants in the lessons-learned workshop held in early 2012 in Paris also emphasized the critical role of competent and effective management:

[A] typical joint evaluation is generally managed at three levels. A *Reference Group/Steering Committee*, which comprises all or most stakeholders, takes the strategic decisions such as adopting the Terms of Reference, criteria for selection of the evaluation team, and endorsement of the final report in terms of its quality and credibility. A smaller *Management Group* elected by the Reference Group is involved with day-to-day issues and provides guidance and advice to the third level, the *Evaluation Manager/Coordinator*, who is responsible for the daily, practical work and execution of decisions taken by the steering bodies as well as direct contact with the evaluation team as need be. This was considered a vi-

able mechanism of governing complex joint evaluations.
(Paris Declaration Evaluation Secretariat, 2012, p. 4)

Dabelstein and Kliet (2012) describe in detail the governance and management of the Paris Declaration Evaluation. Those governance and management structures and processes were critical to ensuring the evaluators' independence while supporting meaningful and appropriate stakeholder engagement. But while the governing arrangements were aimed at ensuring the Evaluation's independence, they were also geared toward making the Evaluation a truly joint exercise. Moreover, the evaluation was conducted in accordance with the Paris Declaration principles.

In-country evaluation team procurement was the responsibility of and followed partner country processes. Indeed, from the beginning, the evaluation was conceptualized, designed, and implemented to adhere to—and support the implementation of—Paris Declaration principles themselves. This was not always easy and resulted in administrative challenges and, in some cases, delays in implementation, but the evidence gathered in the meta-evaluation supports the conclusion that the evaluation was a genuine and authentic joint endeavour.

It is worth re-emphasizing this point: *the Evaluation was conducted according to and in compliance with the Paris Declaration principles* to ensure that it was genuinely and authentically a *joint* endeavour that fully respected the interests and concerns of partner countries and donor agencies, as well as the broader international community. The evaluation process was participatory and consultative among partner countries, donors, and international organization participants.

Evaluation financing and resources

It is clear that an international joint evaluation of this magnitude and scope will be costly. One of the elements that make this evaluation an exemplar is that the organizers used their great experience, credibility, and global contacts to raise sufficient funds to conduct a high-quality evaluation. An underresourced evaluation would lead to a poor evaluation in every respect, including poor quality and poor use. Moreover, it is yet another hallmark of this evaluation that the financing is transparent. The total cost of the two phases of the Paris Declaration Evaluation is estimated at €10,433,200 (US\$12,977,500)

as of June 30, 2012. (See Dabelstein & Kliet, 2012, Exhibit 4, for financial details.) Undoubtedly some will consider this very expensive. However, the subject evaluated—the way development cooperation is functioning—is broad, complex, and of the essence for the future configuration of international development cooperation, and the evaluation covered 22 countries and 18 donors/agencies directly and many more indirectly. As important as the overall resources garnered to support the evaluation is the way in which multiple sources of funds and diverse funding mechanisms were used. Effectively and efficiently managing these funds and funding mechanisms was essential to the evaluation's implementation and ultimate success.

7. The Evaluation Adheres systematically, intentionally, comprehensively, and visibly to the new OECD/DAC standards for evaluation quality, demonstrating how to take a *standards-based approach* to the conduct of an evaluation.

The Development Assistance Committee (DAC) of the OECD finalized quality standards for development evaluation in 2010 based on draft standards tested from 2006 to 2010. The DAC Network on Development Evaluation supports implementation of the standards. The DAC quality standards for evaluation provide relevant, appropriate, and useful guidance for conducting a complex, strategic, and principles-focused evaluation like the Paris Declaration Evaluation. The meta-evaluation (Patton, 2012) found that three fourths (76%) of country participants reported that the DAC standards were “very useful” or “somewhat useful” to the partner country evaluations. They also proved applicable, appropriate, and useful for the evaluation of the evaluation (Dabelstein & Patton, 2012; Patton, 2012). The Evaluation Final Report Technical Annex includes a detailed analysis of how the DAC standards informed the evaluation (Wood et al., 2011, pp. 218–221). This itself is exemplary in that the Evaluation systematically assessed itself against a set of evaluation standards.

8. The Evaluation demonstrates multiple dimensions of *process use*, and how and why process use can be as important as findings use.

One of the major conceptual developments in scholarship on evaluation use of the past decade has been attention to *process use*. Process use concerns the impacts of an evaluation that go beyond the findings. It includes what is learned by those involved, the capacity for evaluation that is built through doing evaluation, and the variety

of impacts that result from the fact that the inquiry is even being undertaken (Cousins, 2007; Patton, 2012, Chapter 6). The Paris Declaration Evaluation is an exemplar of the power and importance of process use.

Focusing attention on and inquiring deeply into the Paris Declaration and its effects

For five years the evaluation engaged a dedicated, knowledgeable, and diverse group of people in thinking about, gathering data on, and interpreting the effects of the Paris Declaration and Accra Agenda for Action. A comprehensive and in-depth evaluation process involves defining terms, clarifying and focusing priority questions, operationalizing key concepts, formalizing the theory of change, establishing boundaries around the inquiry, building partnerships, constructing data collection protocols, building capacity to gather and analyze data, facilitating relations, and focusing attention on what is being evaluated: the implementation of the Paris Declaration and its effects on development aid. The final product, *The Evaluation of the Paris Declaration Phase 2 Final Report* (Wood et al., 2011), was the culmination of all this effort. But behind that report, both making it possible and extending its impact, were a large number of people around the world who were deeply engaged in thinking about these issues and bringing evaluative thinking and methods to bear on critical questions of effectiveness, efficiency, and impact. It seems fair to suggest that the large and diverse group of people engaged in the many activities and distinct studies that, together, make up the whole of the evaluation, have thought as deeply about these issues as any group anywhere in the world. An evaluation process is not just talking about something. It means studying the thing being evaluated. At its best, as in this case, it involves deep intellectual and rigorous methodological inquiry. Thus, a major example of process use was the very process of focusing in-depth attention on and inquiring into the Paris Declaration.

Expanding knowledge of and engagement with the Paris Declaration principles

Process use includes the in-depth knowledge and understanding created among those engaged in the evaluation in participating countries and agencies. Another dimension of process use is the increased awareness of and engagement with the Paris Declaration, as well as issues of aid effectiveness, well beyond what would

have occurred without the evaluation. Policy makers, ministers, civil servants, agency directors and staff, NGOs, and others who were interviewed as part of country and donor studies became more aware of and thought more deeply about the Paris Declaration as a result of being interviewed. The Evaluation, therefore, served a dissemination function even as the inquiry gathered data about the degree of knowledge about and engagement with the Paris Declaration. Members of national reference groups, country evaluation teams, and donor study groups have all reported that the evaluation contributed to greater knowledge about and, in many cases, greater commitment to and implementation of the Paris Declaration. Attention to process use highlights the fact that evaluation inquiries have an impact on the focus of inquiry quite apart from and well before the findings are reported.

Evaluation capacity-building as process use

Capacity-building is yet another dimension of process use. The workshops for country teams, work on understanding and applying the evaluation framework, technical assistance to country teams, and feedback on country reports are examples of capacity-building supported by the core evaluation team and Paris Declaration Evaluation Secretariat throughout the evaluation process. As the Final Evaluation report asserts:

Adequate resources for specific technical support are an essential part of the Evaluation design; they should not be seen as “additional” or to be drawn on “if/where required.” (Wood et al., 2011, p. 218)

9. This is an *exemplar of utilization-focused evaluation*, that is, making attention to intended use by intended users the focus of the Evaluation from beginning to end.

The OECD/DAC Standards include a section on follow-up, use, and learning. DAC Standard 4.1 on *timeliness, relevance and use of the evaluation* begins: “The evaluation is designed, conducted, and reported to meet the needs of the intended users” (DAC, 2010, p. 15). That, in essence, is the overarching criterion of utilization-focused evaluation (Patton, 2012). It goes on: “Conclusions, recommendations, and lessons are clear, relevant, targeted, and actionable so that the evaluation can be used to achieve its intended learning and accountability objectives.” These are basically *potential* utility criteria, as

is the standard that “the evaluation is delivered in time to ensure optimal use of the results.”

Then standard 4.2 moves from potential utility to actual use:

DAC Standard 4.2. Systematic response to and follow-up on recommendations. Recommendations are systematically responded to and action taken by the person(s)/body targeted in each recommendation. This includes a formal management response and follow-up. All agreed follow-up actions are tracked to ensure accountability for their implementation. (DAC, 2010, p. 15)

This standard creates an interaction around the evaluation findings and recommendations. Here, again, the Paris Declaration Evaluation is exemplary. The Evaluation was designed and implemented to contribute to the 4th High Level Forum on Aid Effectiveness, which met in Busan, Korea, in December 2011. That forum involved presidents, foreign ministers, international agency heads, and global policy makers. The delegates to the High Level Forum, and those preparing and supporting the delegates, were the primary intended users of the evaluation. The evaluation questions focused on the priority issues that the High Level Forum was expected to discuss. The timing of the Evaluation, including very tight deadlines that were rigorously adhered to, were based on when the evaluation report had to be produced to contribute to the deliberations of the High Level Forum. The recommendations are differentiated for and aimed at specific intended users. Three recommendations are aimed at policymakers in partner countries; three are directed specifically to policymakers in donor countries; and five are offered to policymakers in both partner countries and donor countries and agencies. This exemplifies being utilization-focused. And, as already noted earlier, key stakeholders connected to primary intended users were involved in all stages of the Evaluation, in partner country case studies, in development agency reviews, and in the International Reference Group. In his opening remarks in Busan, the Chair of the OECD Development Assistance Committee made special mention of the importance of the Independent Evaluation.

To emphasize the global importance of the High Level Forum, a list of opening speakers is informative: OECD Secretary-General Angel Gurría, Korean President Lee Myung-bak, United Nations Secretary-General Ban Ki-Moon, Rwandan President Paul Kagame, Queen Rania of Jordan, United States Secretary of State Hillary

Clinton, and Chair of the Better Aid Co-ordinating Group Antonio Tujan Jr. The Chairs of the International Reference Group and Management Group as well as the Evaluation Team Leader and the Head of the Paris Declaration Evaluation Secretariat participated in the Busan High Level Forum and made presentations on the evaluation findings and the evaluation process. Several National Coordinators representing their countries were also present.

While the Paris Declaration Evaluation produced findings and recommendations that could be and were used in the deliberations at the 4th High Level Forum, the most important use came in the preparations for and build-up to the forum in Busan rather than at the forum itself. In truth, an international forum of this kind involves a lot of pro forma ceremony, political posturing, speech-making, and window dressing. Knowing this, the evaluation dissemination process was designed, implemented, and orchestrated to attain maximum use in the agenda setting and planning processes that led up to the Forum. That's when the real substantive use occurred. Thus, dissemination of findings was anticipated at the design stage, budgeted for, scheduled, and facilitated as described in detail in Wood & Dabelstein (2012).

10. The Evaluation demonstrates the role of *meta-evaluation* to enhance credibility, accountability, and utility.

The major addition to the Joint Committee on Standards for Educational Evaluation in 2010 was the addition of "Evaluation Accountability Standards" focused on metaevaluation (Joint Committee, 2010; Yarbrough, Shulha, Hopson, & Caruthers, 2010). *Meta-evaluation* is evaluation of evaluation. "Metaevaluation is a professional obligation of evaluators. Achieving and sustaining the status of the profession requires subjecting one's work to evaluation and using the findings to serve clients well and over time to strengthen services" (Stufflebeam & Shinkfield, 2007, p. 649).

The Paris Declaration Evaluation is an exemplary model in taking meta-evaluation seriously as described in detail in Patton (2012). The evaluation of the Evaluation was commissioned in time for the meta-evaluators to observe how the Evaluation was conducted and to complete the meta-evaluation report shortly after the release of the final evaluation report. Thus, readers of the evaluation report also had immediate access to the meta-evaluation as well as a summary of the meta-evaluation findings included in the preface to the final

report. In these ways, the meta-evaluation itself was utilization-focused.

OUTSTANDING EVALUATION AWARD AND LESSONS LEARNED

The Paris Declaration Evaluation received the 2012 American Evaluation Association (AEA) Outstanding Evaluation Award. The AEA award announcement states:

The success of the Paris Declaration Phase 2 Evaluation required an unusually skilled, knowledgeable, and committed evaluation team; a visionary, well-organized, and well-connected Secretariat to manage the logistics, international stakeholder meetings, and financial accounts; and a highly competent and respected Management Group to provide oversight and ensure the Evaluation's independence and integrity. This was an extraordinary partnership where all involved understood their roles, carried out their responsibilities fully and effectively, and respected the contributions of other members of the collaboration. Thus, the 2012 AEA Outstanding Evaluation is for all major participants in the Joint Evaluation Collaboration. (AEA, 2012)

In accepting the award at the annual conference of the AEA in Minneapolis on behalf of the joint collaboration, Bernard Wood, the Core Evaluation Team Leader, paid special tribute to the many and diverse participants in the evaluation process and, especially, the skillful management of the evaluation. The insights he offered and lessons he emphasized provide a fitting conclusion to this article, so we quote him at some length.

You can't see them right now, but there are more than 200 people standing beside me as I have the honour of accepting this award. They are the people from more than 50 countries around the world who played key parts in this unique evaluation....

And what better year for the award to this evaluation, when the AEA's theme is "Evaluation in Complex Ecologies"? The ecology around this evaluation was so complex that many said it couldn't be done. And they were almost proved right at a number of points.

It was tough—intellectually and methodologically, in terms of building relationships of trust across cultural and political divides, and not least organizationally and logistically. I am very proud of what our core evaluation team was able to do:

- first, solving the riddle of a coherent and manageable methodological framework in a highly participatory way;
- second, working with far-flung teams to apply the comparative design while drawing out the central issues in their own contexts and reinforcing some of their evaluation capacities; and
- third, bringing together the synthesis picture in a compelling, credible, and policy-relevant report that was shaped and timed to be useful to a critical international policy debate.

But most of all, I want all of us here as evaluators to pay tribute to the enabling environment that was created for this evaluation by the management group....

They showed enormous courage, skills, integrity, and perseverance in conceiving, resourcing, designing, and sustaining such a complex and sensitive operation and protecting its independence.

In closing, I want to stress to this important forum of evaluators that our tribute to the “best evaluation managers in the world” is not just a polite or sentimental gesture. It is a challenge to all evaluation managers and supervisors. All of us have worked before and since on other important evaluations where the work was dragged down by the lack of organizational and mobilizing skills, respect for evaluation standards, or true independence on the part of evaluation managers. Like evaluators, they have a special and difficult calling, and we must salute and learn from the best. (Wood, 2012)

THE FUTURE OF JOINT EVALUATIONS

The lessons-learned workshop held as the final reflective practice activity of the Paris Declaration Evaluation included closing remarks by Niels Dabelstein, the Head of the Paris Declaration Evaluation

Secretariat (and co-author of this article), on the status and future challenges of joint international evaluations. The lessons-learned report concluded with his remarks, which we reproduce here.

[T]here is a strong interest in joint evaluations while all recognise their limitations and constraints. Large-scale international evaluations like the PDE are the exception rather than the rule and the focus should be on country-led joint evaluations at country level (or on a smaller scale). Several partner countries have an institutional framework for evaluation in place, but many still need support to build capacity and to improve institutional arrangements, including quality awareness. Evaluation capacity building is much about empowerment, and donors should use country systems and procedures to the greatest extent possible. If evaluation is to develop into a joint management and accountability tool, the donors need to be flexible, to adjust to the planning cycle of the developing countries, to enable the partner countries to take the lead in evaluations, and to be accountable not only to their own constituencies but also their clients—the people of the developing countries. (Paris Declaration Evaluation Secretariat, 2012, p. 7)

What Is At Stake?

This concluding article opened with the results of the lessons-learned workshop held as the closing activity of the Paris Declaration Evaluation. Building on the findings of both that workshop and the meta-evaluation, we have identified and discussed 10 contributions of the Paris Declaration Evaluation to evaluation theory and practice. These are contributions that result from examining the Paris Declaration Evaluation as an exemplary model for joint international evaluations. These contributions to evaluation practice and theory are above and beyond the primary intended use of the Evaluation, which was to inform the 4th High Level Forum deliberations on development effectiveness in Busan, Korea, in December 2011.

We close this article and, indeed, this special issue, by reminding readers what was at stake in this evaluation, for that is what motivated those involved to do all they could to ensure the evaluation's quality, credibility, and use. That is also what justifies the substantial costs of the evaluation and the enormous resources expended, both

financial and human. The preface to the Final Report addresses what is at stake and so we close where the report begins:

The Evaluation is important both for accountability and to point the way for future improvements. The underlying stakes are huge: better lives for billions of people (reflected in the approaching Millennium Development Goals for 2015); hundreds of billions of dollars expended; vital international relationships; and growing demands to see results from development aid. (Wood et al., 2011, p. xii)

However, the arguments put forward by the Evaluation to keep the momentum of and reinforce the Aid Effectiveness agenda spelled out in the Paris Declaration and the Accra Agenda for Action were not entirely reflected in the final Busan proceedings and document. One may say that the Aid Effectiveness agenda lost out to the political agenda of engaging new partners (i.e., “emerging donors”) in the development partnership. Thus, the future of both the Paris Declaration principles and joint international evaluations are in doubt.

The sad thing is that the Paris Declaration was derailed in Busan. The final document focused on getting the BRIC countries [Brazil, Russia, India and China] on board, but without committing themselves to the Paris principles. Old signatories affirmed their commitment, but when we have the BRICs not accepting it, others will question why they have to pay attention to it either. (Niels Dabelstein, Head of the Paris Declaration Evaluation Secretariat, quoted in *Development Today*, 2012)

Final Reminder Lesson

Evaluators and evaluation reports can present findings and make recommendations, but policy makers and other decision makers ultimately determine whether findings are taken seriously and recommendations are followed. While evaluators may sometimes, even often, wish that they could mandate follow-through on evaluation findings and recommendations, that is not the evaluators' role. John Milton famously wrote (1655/1940), “They also serve who only stand and wait.” We would say: “They also serve who independently present findings and recommendations.” Evaluators should work diligently to inform and facilitate use (not just stand and wait), and the find-

ings and recommendations of the Paris Declaration Evaluation were actively disseminated and promoted. But evaluators are, ultimately, not the primary users. And so it has ever been.

REFERENCES

- American Evaluation Association. (2012). *2012 Outstanding Evaluation Award*. Retrieved from www.eval.org
- Beck, T., & Buchanan-Smith, M. (2008). *Joint evaluations coming of age? The quality and future scope of joint evaluations*. London, UK: ALNAP. Retrieved from <http://www.alnap.org/pool/files/7rha-Ch3.pdf>
- Betts, J., & Wood, B. (2012). The Paris Declaration Evaluation process and methods. *Canadian Journal of Program Evaluation*, 27(3), 69–100.
- Borton, J. (2004). The joint evaluation of emergency assistance to Rwanda. *Humanitarian Exchange Magazine*, 26. Retrieved from <http://www.odihpn.org/humanitarian-exchange-magazine/issue-26/the-joint-evaluation-of-emergency-assistance-to-rwanda>
- Campbell-Patton, C., & Patton, M. Q. (2007, January). *The importance and challenges of monitoring compliance with African Union or other regional and international declarations*. Presentation at the meeting of AfrEA IV in Area 4: Governance, evaluation and development challenges in Africa, Niamey, Niger.
- Canadian Evaluation Society. (2010). *Competencies for Canadian evaluation practice*. Ottawa, ON: Author. Retrieved from http://www.evaluationcanada.ca/txt/2_competencies_cdn_evaluation_practice.pdf
- Canadian Evaluation Society. (2012). *Professional designations program*. Ottawa, ON: Author. Retrieved from <http://www.evaluationcanada.ca/site.cgi?en:5:6>
- Christie, C. A. (2012). Advancing empirical scholarship to further develop evaluation theory and practice. *Canadian Journal of Program Evaluation*, 26(1), 1–18.
- Cousins, J. B. (Ed.). (2007). *Process use in theory, research, and practice*. *New Directions for Evaluation*, No. 116.
- Dabelstein, N. (1996). Evaluating the international humanitarian system: Rationale, process and management of the joint evaluation of the

- international response to the Rwanda genocide. *Disasters*, 20(4), 287-294. Special Issue: Multi-donor evaluation of Rwanda. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.1467-7717.1996.tb01044.x/abstract>
- Dabelstein, N., & Kliet, T. (2012). Preparing, governing, and managing the Paris Declaration Evaluation. *Canadian Journal of Program Evaluation*, 27(3), 37-68.
- Dabelstein, N., & Patton, M. Q. (2012). The Paris Declaration on Aid Effectiveness: History and significance. *Canadian Journal of Program Evaluation*, 27(3), 19-36.
- Danida (Danish Ministry of Foreign Affairs). (1996). *The international response to conflict and genocide: Lessons from the Rwanda experience* (Synthesis report). Copenhagen, Denmark: Danida Evaluation Department. Retrieved from <http://www.oecd.org/dataoecd/32/7/50189495.pdf>
- Development Assistance Committee. (2010). *Quality standards for development evaluation* (DAC Guidelines and Reference Series). Paris, France: Organisation for Economic Cooperation and Development. Retrieved from <http://www.oecd.org/dataoecd/55/0/44798177.pdf>
- Development Today*. (2012). International recognition for Niels Dabelstein. October 12, p. 6. Retrieved from http://www.development-today.com/magazine/2012/dt_12/news/international_recognition_for_niels_dabelstein
- Joint Committee on Standards for Educational Evaluation. (2010). *The program evaluation standards*. Retrieved from <http://www.jcsee.org/program-evaluation-standards>
- Milton, J. (1940). On his blindness. In A. Quiller-Couch (Ed.), *The Oxford book of English verse: 1250-1900*. New York, NY: Oxford University Press. (Original work published 1655)
- Paris Declaration Evaluation Secretariat. (2012). *Workshop on lessons learned from international joint evaluations*. Retrieved from <http://www.oecd.org/dataoecd/52/39/49892763.pdf>
- Patrizi, P., & Patton, M. Q. (Eds.). (2010). Evaluating strategy. *New Directions for Evaluation*, No. 128.

- Patton, M. Q. (2012). *Essentials of utilisation-focused evaluation*. Thousand Oaks, CA: Sage.
- Stevahn, L., King, J., Ghere, G., & Minnema, J. (2005). Establishing essential competencies for program evaluators. *American Journal of Evaluation*, 26(1), 43–59. Retrieved from <http://aje.sagepub.com/cgi/content/abstract/26/1/43tills>
- Stufflebeam, D., & Shinkfield, A. (2007). Metaevaluation: Evaluating evaluations. In *Evaluation theory, models, and applications*. San Francisco, CA: Jossey-Bass.
- University of California. (2006). Treaties and international agreements. Berkeley, CA: University of California Library. Retrieved from http://www.lib.berkeley.edu/doemoff/govinfo/intl/gov_trtygde.html
- Wood, B. (2012, October). 2012 Outstanding Evaluation Award acceptance speech. Presented at the annual conference of the American Evaluation Association, Minneapolis, MN.
- Wood, B., & Betts, J. (2012). Results of the Paris Declaration Evaluation. *Canadian Journal of Program Evaluation*, 27(3), 101–126.
- Wood, B., Betts, J., Etta, F., Gayfer, J., Kabell, D., Ngwira, N., ... Samaranyake, M. (2011). *The evaluation of the Paris Declaration, final report*. Copenhagen, Denmark: Danish Institute for International Studies. Retrieved from www.oecd.org/dac/evaluationnetwork/pde
- Wood, B., & Dabelstein, N. (2012). Dissemination and early use of the Paris Declaration Evaluation. *Canadian Journal of Program Evaluation*, 27(3), 127–144.
- Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2010). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd ed.). Thousand Oaks, CA: Sage.

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 continued as Vice-Chairman until 2005. He led the drafting of the DAC Principles for Aid Evaluation in 1991, the Guidelines for Evaluating Humanitarian Assistance in 1998, the DAC Evaluation Glossary in 2002, the DAC Guidance for Joint Evaluation in 2005, and the DAC Evaluation Quality Standards in 2006. He was also one of the “founding fathers” of the International Development Evaluation Association (IDEAS) in 2002. Most recently (2007–2012) he managed the Joint Evaluation of the Implementation of the Paris Declaration.

Michael Quinn Patton is founder and director of *Utilization-Focused Evaluation*, Saint Paul, Minnesota. He is former president of the American Evaluation Association and recipient of both the Alva and Gunnar Myrdal Award from the Evaluation Research Society for “outstanding contributions to evaluation use and practice” and the Paul F. Lazarsfeld Award for lifetime contributions to evaluation theory from the American Evaluation Association. The Society for Applied Sociology honored him with the Lester F. Ward Award for Outstanding Contributions to Applied Sociology. He is author of six major evaluation books including *Essentials of Utilization-Focused Evaluation* (2012) and *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use* (2011). He conducted the metaevaluation of the Paris Declaration Evaluation.