The DAC Network on Development Evaluation – 30 years of strengthening learning in development

June 2013
The DAC Network on Development Evaluation –
30 years of strengthening learning in development
The Organisation for Economic Co-operation and Development is a unique forum where the governments of 34 democracies work together to address the economic, social and environmental challenges of globalisation. The OECD is also at the forefront of efforts to understand and to help governments respond to new developments and concerns, such as corporate governance, the information economy and the challenges of an ageing population. The Organisation provides a setting where governments can compare policy experiences, seek answers to common problems, identify good practice and work to co-ordinate domestic and international policies.

In order to achieve its aims the OECD has a number of specialised committees. One of these is the Development Assistance Committee (DAC), whose members have agreed to secure an expansion of aggregate volume of resources made available to developing countries and to improve their effectiveness. To this end, members periodically review together both the amount and the nature of their contributions to aid programmes, bilateral and multilateral, and consult each other on all other relevant aspects of their development assistance policies.

The members of the Development Assistance Committee are: Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, the United Kingdom, the United States and the European Union.

Please cite this publication as:

Foreword

This summary of the work of the DAC on evaluation by Niels Dabelstein is a little gem. My first involvement with the Group goes back to 1998, when he was the chair, and I remember being impressed by his encouragement of our operational discussions in a friendly atmosphere.

Evaluation is about accountability and lessons learnt, and we are reminded here of issues which have been constant themes in the work of the group, including contentious ones. What strikes me is that the influence of the Group has been very impressive: its five criteria have become accepted as the key structure for evaluation, it has contributed enormously through the development of standards and guidelines, and it has constantly been at the forefront of efforts to adapt evaluation to new challenges such as evaluation in fragile situations, and the evaluation of budget support. It has helped to encourage the development of evaluation capacity in other areas, not only in multilateral organisations and in partner countries but also across cultures by enabling the translation of normative work into many languages.

This success is all the more remarkable in view of the limited resources available, and has been achieved thanks to the effort of former chairs and vice-chairs, but especially thanks to the Secretariat whose energy and focus has maintained the highest standards. We are fortunate to have Hans Lundgren and his team who can take credit for this.

In times of austerity it is usual for societies to become introverted and for development to slip down the political agenda. Today it is therefore particularly important that decisions are evidence-based, and that evaluation can quickly and clearly offer conclusions on what works and what doesn’t. Not only do we therefore need to ensure that evaluations are of good quality so that they are credible, we need also to be able to communicate our messages in ways which relate both to the general public and to policy-makers. The use of Facebook and Twitter to communicate is now part of the Group’s daily life, but we need to learn how to express often complex conclusions in clear, everyday language which doesn’t leave the reader asking what it means in simple terms.

Just as important is the need for us to do our jobs as efficiently as possible, which means being able to share evaluation conclusions more easily and to work together on a regular basis. With the help and support of several members we have recently been looking at how we can increase our use of systematic reviews which aim to get the big picture from existing evaluation material using credible, robust methods.

Such honesty in addressing the problems we all face in our work, and in giving a lead in providing the tools to take us forward, is sorely needed. This is why the DAC will remain in the future the focal point of evaluation in the development field and will play a key role in ensuring the increased effectiveness of our work.

Martyn Pennington
Chair of the DAC Network on Development Evaluation
Acknowledgements

Writing this account of 30 years’ work of the DAC Evaluation Network has been a pleasant journey down memory lane. Having participated in 38 of the 51 meetings of the Evaluation Network as member, vice-chair and chair has been a great privilege, a challenge, a fantastic learning experience and last but not least lots of good company, several of which have led to genuine friendships. The DAC Evaluation Network has been a very important part of my professional and social life for 25 years and I take this opportunity to thank the many colleagues I have worked with for shorter or longer time.

This recount is primarily based on the 51 records of the meetings and the documents tabled at meetings as well as papers and reports produced by the Network, all of which has been meticulously archived by the OECD. I have chosen not to make scientific reference to each and every document from which I have gleaned, rephrased or outright copied sections or phrases. This would have made the story unreadable. The few other sources used are acknowledged in the text and a list of sources is provided in an annex.

As is evident from the text I have not written this alone. Thanks are due to the many colleagues who have contributed with opinions and anecdotes presented in boxes: Klaus Winkel, my former boss and mentor, Rob van den Berg, Eva Lithman, Hans Lundgren, Haven North, Michael Quinn Patton, Peter Walker, Nick York and Bernard Wood.

When writing history choices have to be made of what to highlight and what to let pass by. It has not been easy to make those choices, but they are mine and I take full responsibility for any omissions and apologise for any possible offences to colleagues who feel their interest has not been reflected adequately.

Special thanks go to Megan Grace Kennedy-Chouane, Ted Kliest and Hans Lundgren for commenting on consecutive drafts, jogging my memory and providing useful suggestions and detail.

Last but not least I want to thank Hanne, my life’s companion, for bearing with my being away in Paris – and so many other places – so often, and for living with my living in the past the last couple of months.

Niels Dabelstein

May 2013
Table of Contents
Foreword ............................................................................................................................................... 3
Acknowledgements .............................................................................................................................. 4
1. How it came to be .............................................................................................................................. 6
2. Members and modalities .................................................................................................................... 8
3. Generating knowledge ..................................................................................................................... 11
4. Joint evaluations .............................................................................................................................. 17
5. Reaching out – Collaboration with partners ...................................................................................... 27
6. Advocating and sharing best practice in evaluation – The normative role ........................................... 32
7. Sharing and communicating evaluation results .................................................................................. 40
8. Some concluding remarks ................................................................................................................. 41
Annex A ............................................................................................................................................... 42
References and Further Reading ........................................................................................................... 43

Acronyms

AfrEA: African Evaluation Association
CLEAR: Regional Centres for Learning on Evaluation and Results
DAC: Development Assistance Committee
Danida: Development co-operation programme of Denmark
DEReC: DAC Evaluation Resource Centre (website)
ECD: evaluation capacity development
ECG: Evaluation Co-operation Group (of multilateral organisations)
EvalNet: DAC Network on Development Evaluation
IDEAS: International Development Evaluation
IPDET: International Program for Development Evaluation Training
OECD: Organisation for Economic Co-operation and Development
UN: United Nations
UNEG: United Nations Evaluation Group
1. How it came to be

Evaluation has been a major theme in the DAC’s work from the start. In the first Chairman’s Report in 1962, DAC Chair James Riddleberger said that, “Member Governments would be well advised to devote more attention to critical evaluations of the results of their aid operations and exchange experience in this field.” The Chairman concluded rightly that the issue “of the efficacy of foreign aid will clearly be one of the major pre-occupations of the Committee in its future work” (DAC, 1962:41). He added that DAC should consider ways and means of assisting in such evaluations.

The Nordics, Netherlands and Germany prioritised early evaluation efforts. There were seminars in Berlin (1966), the Netherlands (1970 and 1973) and Denmark (1975) on the problems of aid evaluation and the DAC produced a paper “Evaluating Development Assistance” (available in OECD archives) in 1972 which described problems of method and suggested ways in which evaluation might be approached for use as an effective management tool.

There was a renewed emphasis on evaluating the effectiveness of aid in the early 1980s following directives from the November 1979 DAC High Level Meeting which led to the creation of a Group of Correspondents on Aid Evaluation tasked to, i) report on existing findings on aid effectiveness; ii) consider feedback into policy-making; iii) report on ways in which evaluation could be supportive of public information; and iv) consider how to support evaluation work in developing countries. The priority was on producing a report on aid effectiveness to be tabled at the 1982 High level Meeting.

Before the Correspondents submitted the report several issues were debated that would prove important in shaping the work of the group overtime. In particular, the question of whether evaluation findings could be used to strengthen public support for aid was an early example of how there was sometimes a mismatch between what the DAC wanted from the Evaluation Group and what the Group itself felt was possible or desirable. While acknowledging that its function was to service the needs of the DAC, from the very beginning the Group began asserting its obligation to refine the questions posed by the DAC to make them more amenable to meaningful study. DAC, as a political body tended to have unrealistic expectations of what evaluation can deliver. In this case, for example, they were perhaps too optimistic that evidence about the results of aid would convince donor publics that aid was a good idea. In contrast, the Evaluation Group understood that evaluations might have mixed or ambiguous results, which could be unconvincing or even damage public perceptions of the usefulness of aid. As professional evaluators, the Group chose to be independent in its advice and has continued to be so since. This has contributed to the credibility of the Group’s work, but also, sometimes, slowed its response to DAC requests.

The evaluation report 1982 high-level meeting was sobering reading. It frankly acknowledged that the existing body of evaluations could not provide substantial evidence of the effectiveness of aid. The Group of Correspondents used the opportunity to advocate for further work on improving evaluation and the use of evaluation by making a “resolute departure” from the static retrospective approach taken so far (does aid promote development?) towards a more forward looking approach which emphasised what could be done to improve the impact of aid on development.
The 1982 DAC High-Level Meeting agreed to establish the DAC Expert Group on Aid Evaluation with the mandate to:

- Strengthen exchange of information, experience and co-operation between national and, as appropriate multilateral evaluation units, with a view to:
  - improving the evaluation activities of individual members,
  - encouraging standardization of methodological and conceptual frameworks, and
  - laying the basis for improved donor co-ordination in the planning of major evaluation studies.

- Contribute to improved aid effectiveness by attempting to distil from evaluation studies operational lessons for project and programme planning and design, for consideration by the DAC.

- Examining the feasibility of launching a set of joint or co-ordinated studies, undertaken by interested members, of aid effectiveness in a major sector, such as agriculture, or possibly in particular countries.

- Seeking ways to promote and support developing countries' evaluation capabilities.
2. Members and modalities

Members of the Expert Group were the heads of evaluation in the DAC members’ development agencies with the World Bank and the International Monetary Fund as observers. Already at the first meeting in March 1983 it was agreed to invite the OECD Development Centre, the UNDP and the three regional development banks as observers. The membership has not changed since, except when DAC membership expanded, but the Group has occasionally debated whether to expand to include other development partners. The question of membership has been raised in the context of growing efforts to support a collaborative, partner-led approach to development co-operation. However, the group has always decided to maintain its more limited membership, while finding ways to collaborate with other partners. Many members felt that an important strength of the group is the “peer” nature of its members. Working with colleagues who face common challenges and play similar roles in their respective countries, contributes to a strong sense of camaraderie among members and supports learning. Instead of expanding membership, non-OECD member countries and organisations have been invited on an ad hoc basis when items on the agenda warranted. The Group has also organised many seminars, workshops and conferences with wide participation.

The group’s mandate has remained remarkably stable over the years despite changes of name: In 1998, following a DAC review of all subsidiary bodies, the Expert Group was renamed the Working Party on Aid Evaluation, with the same mandate, and after yet another review in 2003 it was renamed the DAC Network on Development Evaluation.¹

In 2012 the DAC again reviewed its subsidiary bodies and revised the mandate. It now calls on the network to:

- Advocate and share knowledge on best practice in evaluation.
- Synthesise evaluation evidence on key development topics and share findings and lessons
- Work with developing country partners to increase collaboration and promote and support evaluation capacity development.

This mandate, in fewer words and less specificity, contains the same fundamental elements as the 1982 mandate. This continuity reflects the on-going relevance of the group’s work. It also demonstrates the need

¹ The terms Evaluation Group or just Group will be used in the rest of this paper.
for continuous learning and improvements of the evaluation function and its products to adapt to changes in development co-operation.

The continuity does not indicate that nothing has changed in the past 30 years – much has – both in terms of the specific work programme of the group and in its ways of working. The group has contributed substantially to improve evaluation norms, standards and methods; it has contributed substantially to evidence based knowledge about development through synthesising evaluations, coordinated evaluations and joint evaluations; and it has provided substantial support to evaluation capacity development, both to its members and, more importantly, to partner countries. The following chapters will tell about these endeavours.

But first, a little about how the group works:

The first meeting of the group which I attended was held in January 1988 in a high ceilinged, wood panelled salon in the elegant Château de la Muette. Fifty men in dark suits and ten women in sombre dresses formally re-elected the chair before getting on with a long agenda. This overtly formal atmosphere was deceptive; it hid what was actually a collegial group of professionals. The group was (and is) composed of evaluation experts who share a dedication to their work and who have developed a modus operandi that combines the formal requirements of an international organisation with the relaxed, friendly atmosphere of a community of professionals. Like with all DAC bodies the meetings operate on the basis of consensus and proceed only when there is no objection.

In accordance with OECD rules, the chair is elected at the first meeting in a calendar year. Over time this requirement has been adapted to mean that a chair generally serve as long as he or she is willing – and not opposed. A few chairs have sat for four years or more, others for less than a year. In 1990 the group decided that a larger bureau was needed to manage an increasing workload and a vice chair was elected to support the chair. The bureau now consists of the chair, two vice-chairs and the Secretariat to co-ordinate the work.

The group meets every eight to nine months in Paris, with smaller meetings taking place between meetings as needed. From the very beginning much of the work has been undertaken by smaller groups of members, in a flexible way. Themes are tabled either in response to DAC requests or by group members with an interest in a particular issue. Members have then taken responsibility for leading and financing work, including studies that have informed the work of the group, and also been shared with the DAC.
The DAC Evaluation Group is supported by the OECD Secretariat in several ways. Traditional ‘secretariat functions’ such as support to the chair, organising meetings, agendas, records, publications, web site management and communication have always been provided with high quality. But the Secretariat has also been deeply engaged in substantive work. It has pushed agendas forward and is the institutional memory that prevents reinventing the wheel as members come and go. The dominant model in the 1990s was member-driven work, often carried out by consultants, guided and financed by members backed up the Secretariat. Today many of the work streams directly involve or are led by the Secretariat, and sometimes also initiated by it. Task teams remain however an important way of working with members volunteering to take work forward.

The secretariat support has been not only excellent but also remarkably stable in that only two ‘managers’ have been responsible for supporting the Group: John White 1982 - 1989 and Hans Lundgren from 1990 until now. Since 2007, Megan Kennedy-Chouane has contributed in important ways to the group’s work.

The continuity has been a great asset for the group and not least for the chairs who have been able to draw on the long experience of the Secretariat.
3. Generating knowledge

The major reason for the creation of the DAC Evaluation Group was the need for systematic information about the effectiveness of aid and the need for lessons to improve development assistance. The group developed several approaches to fill this role, including synthesising existing evaluations, carrying out joint evaluations, and assessments of multilateral organisations.

3.1 Synthesizing evaluations

“The contribution to improved aid effectiveness by attempting to distil from evaluation studies operational lessons for project and programme planning and design, for consideration by the DAC.”
- 1982 mandate

The first report on “Operational Lessons of Evaluation” presented to the DAC in 1984 was drafted by the Secretariat under the guidance of a working group comprising the United Kingdom (chair), the European Commission, the Netherlands and the United States. The paper was structured around a number of issues identified by the group through members’ inputs, which were in most cases syntheses of their evaluations. The report pointed out that DAC members’ aid often repeats (well known) errors: inadequate project preparation, inappropriate technology, tying of aid, etc. The report led the DAC to initiate work to develop the DAC Principles for Project Appraisal published in 1988. These principles were an important milestone in the development of the DAC’s overall good practice principles for development co-operation.

Although the report produced useful findings, the approach of synthesizing messages from member evaluations had its limitations. It proved difficult to synthesise such broad issues based on project evaluations not designed to answer higher-level questions. Thus the Group explored other avenues to generate knowledge, such as focusing on specific sectors or themes where there was a larger body of evaluation evidence. The Netherlands undertook to synthesise evaluations of health projects and programmes. The report, based on 77 health sector evaluations, raised a number of substantive questions which were deemed so important that the report was submitted in draft form to the DAC meeting on primary health care in September 1988. The report was subsequently finalised and published in 1989. This process provides an example of a difficulty the group has continuously grappled with: timing. The lead time to produce credible evaluation-based reports is often long, which can make it difficult to respond in a timely manner to the changing interests and expectations of the DAC.

2 The DAC produced several other ‘principles’ or ‘good practices’ for technical assistance, programme assistance, women in development environmental assessment, procurement and (un)tying of aid and evaluation. They were published together in “Development Assistance Manual, DAC Principles for Effective Aid” in 1992
A different approach to gathering more systematic evidence was suggested by Denmark in 1985: To incorporate a standard set of questions relating specific topics to all evaluations conducted in 1986 (later expanded to include 1987). The topics chosen were: Sustainability, Impact on Women and Impact on the Environment. Despite initial support several members did not follow through and deliver. Twelve countries delivered on sustainability, eight on Women in Development and seven on Environment. The varied response rate is indicative of another recurring difficulty: even when activities are unanimously agreed, members participate unevenly, based on their own interests and work agendas back home. The reports were submitted to the DAC in 1988 and subsequently published as the first in the series “Selected Issues in Aid Evaluation.” In particular the report on sustainability was influential: its seven “sustainability factors” became household terms in development co-operation and strongly influenced the DAC Principles for Project Appraisal. The report on Environment, based on a small sample of evaluations, did not attract much attention. The Women in Development report was finalised in co-operation with the DAC Expert Group on Women in Development and is seen as the start of a close relationship between the two groups (see box on right).

A third approach to synthesising is represented by the work on Participatory Development and Good Governance. The 1993 High Level Meeting of the DAC approved the creation of an ad hoc Working Party on Participatory Development and Good Governance. The mandate of this Working Party was to help bring “PD/GG” into the mainstream of development co-operation through a three-year program of activities. In parallel the Evaluation Group formed a Participatory Development and Good Governance Steering Committee to develop a review in the following five main areas of activity: a) legal systems; b) public sector management; c) decentralisation; e) human rights; and f) participation. Each of these themes corresponded to an important area of activity of the Participatory Development and Good Governance Working Party. Thus, the experience and analysis reflected in each theme paper produced by the evaluation contributed to the forward looking work of the PD/GG Working Party. A synthesis of the papers “Evaluation of Programs Promoting Participatory Development and Good Governance (PD/GG)” was published in 1997.
Collaboration between DAC networks

One of the most innovative and valuable activities of the subsequent work of the DAC in cementing the role of Participatory Development and Good Governance as an integral element in development co-operation was the collaboration between these two subsidiary bodies: the Working Party on PDGG and the Expert Group on Aid Evaluation. This example of the importance of evaluation to the formulation and implementation of sound policies and practices (which is also an example of the benefits of collaboration between subsidiary bodies of the DAC) demonstrate the value that can be added by the DAC when its Members are willing to devote additional time and effort to joint activities that reflect their shared values and interests”.

James H. Michel, USA,
Chair, Development Assistance Committee 1995-1999

Following this study, UNDP and Germany conducted a joint evaluation in 1999 of UNDP-supported programmes in the area of decentralisation and local governance. As the result of the insights gained from this evaluation, the Group decided to launch a broader project covering the activities of many donors. This resulted in the report “Lessons Learned on Donor Support to Decentralisation and Local Governance” in 2004. The report identified a number of areas where positive results have been achieved, particularly in the field of fiscal decentralisation and financial management, as well as in relation to the strengthening of civil society at local level. The report also pointed to a number of issues in need of further evaluation and research. One of these was the sustainability of donors’ local governance support initiatives, which could not be evaluated at the time, as interventions take long time to bear fruit.

The donors’ focus on environment in the ‘90s and a little prompting from the DAC Working Party on Development Co-operation and Environment inspired the Evaluation Group, led by the Netherlands, to commission a synthesis of evaluations of institutional capacity development in environment in 1997. The report, based on analysis of 70 evaluations and of relevant literature as well as a survey and interviews, was presented to the DAC Working Party on Development Co-operation and Environment in 1999 and published in 2000.

These syntheses were planned and executed over relatively long periods of time, sometimes several years. Only a few times was the Evaluation Group able to respond rapidly to the needs of the DAC as in the case of reconstruction in Afghanistan. At the November 2001 meeting the group mandated the outgoing chair to prepare a paper on key evaluation lessons relevant for reconstruction in Afghanistan. The note: “Aid Responses to Afghanistan: Lessons from Previous Evaluations” was prepared during the first week of December, at which time meetings of Afghan factions with international actors were beginning to address post-conflict reconstruction issues. The paper was discussed at the DAC Senior Level Meeting in December 2001 and was a key background paper, which usefully informed discussions at a DAC Experts’ Meeting on “Afghanistan Reconstruction and Recovery: Seeing Round the Corner” in May 2002. A similar note, “Iraq and Rehabilitation, Lessons from Previous Evaluations” was prepared by the chair and vice-chair in May 2003 for the High Level Meeting and subsequent “DAC Reflection on Reconstruction in Iraq,” in July the same year.
The large synthesis exercises have to a large extent been replaced by joint evaluations, which aim to focus evaluation efforts in a particular policy field (such as budget support or peacebuilding) to build up the knowledge base over time, though some synthesis work continues. In response to a request from the DAC Governance Network and the Working Party on Aid Effectiveness in 2011, the Evaluation Network synthesised findings in the field of public governance, drawing in particular on the findings of several recent joint evaluations in this field, to produce both a full length summary paper and a short policy brief. The paper was shared in advance of the Busan High Level Forum on Aid Effectiveness and provided evidence to inform the work of groups deliberating how best to support country systems and good governance. Findings from major joint work on budget support (described below) was also synthesised in a working paper and in a shorter Evaluation Insights, both of which were widely distributed to the DAC, its subsidiary bodies and other interested groups. Findings on conflict prevention and peacebuilding were likewise summarised and widely disseminated.

In early 2009 the Secretariat launched a new initiative to pull findings from evaluation reports available in the group’s online database and make them more broadly available. The “DEReC Front Page” highlights evaluation reports that are relevant to major current events and policy issues, with the goal of encouraging wider use of evaluation results to improve development policy and practice. New editions of the Front Page are featured on the DEReC homepage periodically, for example looking at findings on gender equality on International Women’s Day or featuring reports on aid for trade during high-level policy discussions on aid for trade. However, while the Front Page significantly increased traffic on the group’s website and helped disseminate specific reports to a broader audience, the exercise again faced the challenge of distilling “key messages” of broader significance from reports that were not necessarily designed to answer such questions. It proved difficult to identify lessons that were at once credible and had external validity, while also providing some new insight. This experience again highlighted the importance of building up high quality evaluative material to increase the overall base for such syntheses, in addition to synthesizing existing material (as was done, for instance, for the Afghanistan and health sector reports mentioned earlier). Such experiences also provided further impetus for the work on developing quality standards and good practice guidance for evaluation (as described below).

In 2011 a new series of policy briefs, called the Evaluation Insights were launched, highlighting emerging evaluation evidence to inform international policy debates. First editions covered budget support, the Haiti earthquake response, rural water and sanitation, food security, gender mainstreaming and governance. New briefs are developed on the initiative of members, either on topics that have been identified as top political priorities or when it is felt that a topic is “ripe” for drawing broader conclusions because there is a wealth of evaluation activity going on in that field. For example, food security was identified as a key theme for OECD development work in 2011/12, and the Netherlands later commissioned a systematic review on this topic, which was used to produce an Insights note. Most of the Insights draw on new meta-evaluations or on systematic reviews, synthesizing key findings across a body of evaluation literature. The briefs are very widely distributed, including at policy meetings, internal and external events and through the Evaluation Network’s social media networks, and there was excellent uptake in relevant development communities. The Insights website received over 6,000 visits in the first six months (after release of three Insights). The Insights are an example of how the group adds value to the work of individual member’s evaluation departments by leveraging its unique position together with Secretariat support to reach a wider audience and use evaluation lessons to influence policy discussions.
3.2 Country evaluations

While early development evaluations focused mainly on assessing individual projects or programmes, evaluation of ‘non project aid’, including sector support, country strategies, food aid and budget support, has a long history in the group. In 1985, the United States produced a paper on methodology for evaluating ‘non-project aid’; the paper noted that “the main difficulty that had been encountered was in the demonstration of causal relationships. In the case of policy-linked lending, this difficulty became acute.” Work continued on developing a methodology for evaluating different types of aid. A near final paper was presented in 1987; however, some members, notably France and Germany, felt that the group was venturing into policy issues beyond its mandate. As a result non-project aid fell off the group’s agenda for several years, though concern around the methodological challenges of assessing such flows remained. Many of these issues resurfaced when the group began work to review experiences and share lessons around approaches to country programme evaluation and on evaluating budget support activities (as described in the section on Joint Evaluation below).

The widespread adoption of country assistance strategies during the 1990s challenged evaluators to assess results at the country level. Development agencies experimented with a range of different approaches to carry out country programme evaluations. In order to review these experiences and to share lessons about strengths and weaknesses inherent in different approaches to country programme evaluation, the group organised two workshops in Vienna in 1994 and 1999. The second workshop resulted in a report, “Evaluating Country Programmes.” This publication provides a unique overview of approaches and methods for country programme evaluation which were presented and discussed at the workshop.
In 2003, the DAC chair called for a more comprehensive perspective on understanding the impact of development co-operation initiatives at the country level. In response, Denmark, the Netherlands, and Sweden presented a proposal for a joint pilot evaluation of the impact of the total combined ODA at country level. The group worked on this ambitious proposal for more than two years to develop a suitable approach. However, in the end it proved too difficult to reach consensus. In the end the initiative was overtaken by the Paris Declaration Evaluation (described below), which provided a broad country-level perspective on the effectiveness of development activities.
4. Joint evaluations

“Examining the feasibility of launching a set of joint or coordinated studies, undertaken by interested members, of aid effectiveness in a major sector, such as agriculture, or possibly in particular countries;” – 1982 Mandate

Joint evaluations, that is, evaluations involving multiple donors or both donor and recipient countries, are encouraged as a way of improving co-ordination and broadening coverage and analytical scope. The group discussed this issue at almost every meeting in the ‘80es. During this period joint evaluations were mostly of multi-bi projects (projects financed by bilateral donors and executed by multilaterals) or projects co-financed with multilateral banks. Overtime, the discussion of how to encourage and support joint evaluation shifted from a focus on multi-bi evaluations, to multiple (bilateral donor) evaluations and finally to an increasing focus on country-led evaluation. This shift was encouraged by the concurrent move towards joint assistance strategies, sector-wide programmes and other “joint” approaches to development assistance.

In response to a DAC High Level Meeting request the Group debated the issue in 1989 (12th meeting) and concluded that there was no appetite for developing a formal programme of joint evaluations. Thus joint evaluations continued to be ad hoc and carried out by small groups of interested countries. However, the group has continuously promoted joint working and the networking of its members has facilitated the launch of numerous joint evaluations. In addition to several joint evaluations by smaller groups, frequently the Nordics, the 1990’s saw a number of flagship joint evaluations, including: the evaluation of the World Food Programme led by Canada, the Netherlands and Norway (1994), The International Response to Conflict and Genocide: Lessons from the Rwanda Experience (1996) initiated by Denmark and the evaluation of European Union Food Aid (1997) led by the Netherlands.

In 1998 a report “Donor Experience with Joint Evaluations -- A Typology and Lessons Learned” demonstrated “that joint evaluations have had positive impacts on donors and recipients as they support the development of a common evaluation culture, increase awareness on different approaches and methods; contribute to the spread of good evaluation practices; and help in strengthening partnerships.” Despite the continued lack of appetite a formalised programme for joint evaluations a growing number of joint evaluations have been undertaken since 1998., including: Joint Multi-Donor Evaluation of European Union Aid Managed by the Commission 1995-1999, led by France; Joint evaluation of the road sub-sector programme Ghana (2000) initiated by Denmark; The Comprehensive Development Framework evaluation: Toward Country-led Development (2003) led by the World Bank; Basic Education: Local Solutions to Global Challenges (2003) initiated and led by the Netherlands; of the Triple C Concept in EU Development Co-operation Policy (2007) initiated by the EU Development Group and led by Sweden and the Netherlands.
In addition to evaluations led by individual countries, occasionally when an issue was considered of special importance joint evaluations were proposed to be “DAC evaluations”. One such attempt was the evaluation of the international response to the genocide in Rwanda. The evaluation was suggested at the meeting in October 1994. A number of members expressed support for the proposal and a willingness to contribute, but two countries (Japan and France) indicated some hesitation arguing that it was too early to evaluate. As consensus was not reached it was agreed that an informal working group of interested members would meet and discuss next steps, the overall scope and approach to the evaluation. Eventually all members of the group, except the World Bank, participated in the 38 member steering committee which also included UN agencies and NGOs. This evaluation had a significant impact on the humanitarian sector as shown in the box below.

Meeting of the steering group managing a joint evaluation of the co-operation of the European commission, Belgium, Denmark, France and Luxembourg with Niger (2009).
Rwanda joint evaluation as a catalyst

The mid 1990’s saw a period of rapid and urgent change across humanitarian agencies. As the Cold War morphed into something new, intra-state conflicts proliferated and the amount of donor state funding flowing via agencies into crisis zones rocketed. Humanitarian aid was no longer a side show; it was there in the glare of the TV lights alongside military intervention, peace conferences and the fall of regimes.

Aid agencies felt the internal pressure of rapid change, the growing research around effectiveness in aid and the heat of the spotlight. Talk and action was already underway to create more universal standards across the endeavour, codes of conduct, operational standards, shared handbooks and annual “state of the business” reports were all in development, but the Rwanda Joint Evaluation catalysed the process. It allowed a community of donors, agencies and academics, already working closely on the evaluation of one of the biggest tragedies and aid operations of the generation, to take that sense of joint commitment forward and apply it to their own projects. The evaluation threw up the need for better accountability, more effective human resource management, evidence based standards for aid, and a clear sense of shared, principled, values.

Thus the Rwanda Evaluation catalysed a host of nascent quality assurance initiatives: The Code of Conduct for Relief Workers, the Sphere project on Minimum Standards, the Active Learning Network on Accountability and Performance (ALNAP), People in Aid and the Humanitarian Accountability Project.

Sometimes history is a matter of happenstance, there comes a moment when political will, funding, knowledge and organizational space all come together to make change happen. In many ways the Rwanda Evaluation provided this.

Dr. Peter Walker, Irwin H. Rosenberg
Professor of Nutrition and Human Security
Tufts University

Several other joint evaluations of humanitarian assistance were also undertaken, such as Humanitarian and Reconstruction in Afghanistan (2005); the Evaluation of Assistance to Internally Displaced Persons (2005) and the 2006 Indian Ocean Tsunami Evaluation Coalition involving DAC members, UN agencies and NGOs and managed by the Active Learning Network for Accountability and Performance in Humanitarian Action (ALNAP).

The influx of a high number of large and small humanitarian aid providers after the Haiti Earthquake in 2010 was a concern to many observers and practitioners of humanitarian assistance. On the evaluation side, it was feared that the fragmented response would be mirrored by an evaluation tsunami – an influx of uncoordinated evaluation activities by different actors. This led to a joint initiative where the DAC Network on Development Evaluation worked with the United Nations Evaluation Group (UNEG) and the Active Learning Network for Accountability and Performance in Humanitarian Action (ALNAP) to produce a shared Context Analysis and Evaluation Framework and create an online Haiti Learning and Accountability Portal. One year after the earthquake, a light synthesis of lessons from the Haiti earthquake response was published on the Evaluation Network’s website to share emerging insights. Individual members also supported the creation of an Evaluation and Learning Support Office in Port-au-Prince, which aimed to encourage collaboration and sharing of information.
4.1 Joint evaluations of budget support

As budget support became an increasingly important form of assistance, assessing non-project aid came back on the agenda. The United Kingdom hosted a workshop in 2003 on budget support and a huge joint evaluation of general budget support was embarked upon. The exercise involved both developing a methodology and carrying out seven country studies. The United Kingdom led the steering committee which included 24 donors and 7 developing countries. The evaluation was completed in 2006. It could not evaluate the poverty effects of general budget support, but it did conclude that “most of the effects of ‘poverty reduction general budget support’ inputs so far are likely to have been on access to services, rather than income poverty and empowerment of the poor” (Joint Evaluation, 2004). One of the challenges was that general budget support had only been practiced for a very short time in some of the partner countries being studied. What were evaluated were the enabling factors of general budget support (better public financial management). Despite these limitations a follow up study concluded that “many organisations had made significant use of the evaluation in policy and operational work.”

The methodological challenges of addressing poverty were taken up in the subsequent evaluation, led by the European Commission supported by other countries. Three pilot country evaluations were completed in 2012: in Zambia, Mali and Tunisia. The overall conclusion (OECD 2012b) was that:

In the three countries, especially where the evaluation looked at a longer time period, the evaluations show that budget support has supported economic growth and social achievements, which have brought about a significant reduction in poverty, especially in urban areas, and have also alleviated exclusion from basic services in rural areas. However, low effect on the rural poor was particularly evident in Zambia, where rural income poverty has not been addressed adequately by the government. In Tunisia, despite a constant decrease of the total number of poor over the last fifteen years, there has similarly been deterioration in the poverty situation in the rural areas.

The Director General of Budget, Mali describes the importance of budget support for improving budget execution, opening dialogue and improving development results, during a seminar on budget support evaluation hosted by Belgium in March 2012. (Photo: Kennedy-Chouane)
4.2 Joint Evaluation of the Implementation of the Paris Declaration

Another major joint undertaking was the Evaluation of the Paris Declaration, probably the largest evaluation ever undertaken. The evaluation was in a sense ‘commissioned’ by the Paris Declaration itself which, unusual for an international declaration, committed the partners to “...explore independent cross-country monitoring and evaluation processes .... to provide a more comprehensive understanding of how increased effectiveness contributes to meeting development objectives” (Paris Declaration, paragraph 11).

Within a few months of the signing of the Paris Declaration in March 2005 plans were under way for its evaluation. The Secretariat presented a discussion note outlining the challenges, and some potential ways forward. A number of members indicated support for the development of a common evaluation framework, to be applied within different country contexts. In line with the principles of ownership and harmonisation, it was underlined that the evaluation should engage partner countries at the earliest stage and to the greatest possible extent. The Secretariat was asked to further explore and define possible ways forward.

At the March 2006 meeting a draft “Options Paper on Evaluation Network Follow-up to the Paris Declaration” was presented. Many members saw this work as highly important. Some members, however, expressed caution that the group might be starting work on the evaluation too early in the overall Paris Declaration process. A task force led by the United Kingdom and Denmark carried the work forward in close collaboration with the Secretariat and in consultation with the Joint Venture on Monitoring the Paris Declaration (JVM), a sub-group under the Working Party on Aid Effectiveness, which included developing
countries. During the following months the Secretariat, the Chair and Denmark consulted partner countries on the way forward and garnered support for the evaluation.

By November 2006 there was broad agreement on a two-step approach. The first step would be to evaluate the early implementation of the Paris Declaration. This work would be delivered to the 3rd High Level Forum in 2008. The second step would be a more summative evaluation of the Declaration’s contribution to development results. The final reports would be delivered to the 4th High Level Forum in 2011. The DAC Guidance for Managing Joint Evaluations, which was published in 2006, would be a key reference document for managing the process.

In March 2007 a Reference Group comprising the participating donors, multilateral agencies and partner countries met and agreed on the evaluation framework, and the evaluation started in July 2007. The first report was delivered in time for the High Level Forum in Accra in September 2008 and the second in time for the Busan High Level Forum in December 2011. In all, 22 partner countries and 18 donors/organisations participated in the evaluation. The Paris Declaration Evaluation was important, in terms of its findings, which influenced the debate around aid effectiveness substantially. The evaluation process itself was also important applying the DAC Evaluation Quality Standards and Guidance on Joint Evaluation and engaging partner countries at par in preparation and execution of the evaluation in full respect of the Paris Declaration principles of ownership and mutual accountability. The evaluation received the American Evaluation Association 2012 Outstanding Evaluation Award. Those interested in the evaluation set-up and the process is referred to a special 2013 issue of the Canadian Journal of Programme Evaluation.
4.3 Challenges and progress on joint evaluation

Joint evaluations come about when and where interests converge. The DAC Evaluation Network, with the active support of the Secretariat, has continuously promoted joint evaluations and developed several approaches to facilitate joint work: exchange of evaluation plans, consolidation of evaluation plans and analyses of trends, matrixes of donors’ engagement in countries and sectors, informal ‘market place’ events and a web-based platform for identifying opportunities. The issue has been on the agenda in one way or the other in just about every meeting. All this has greatly facilitated joint work and in some cases directly generated joint evaluations.

Specific challenges to participation in joint evaluations faced by member agencies were identified in a report, “Challenges, opportunities and approaches for increasing joint donor programming of evaluations” (SADEV, 2008) which was discussed at the meeting in November 2008. An agreement was reached at that meeting on seven high priority recommendations which were aimed at i) facilitating Group member participation in joint evaluations and ii) increasing partner country participation in and ownership of joint evaluation processes. The 2010 review of member systems also looked at how exactly partner countries were involved in “joint” evaluation work and identified weaknesses at the beginning and end of the
The DAC Network on Development Evaluation

evaluation process: most partner stakeholders are consulted only after an evaluation has already been launched (often during field visits) and they are only rarely involved in designing evaluations, selecting topics or following up on the findings. Thus ownership remains low, despite an increasing number of “joint” initiatives. The findings of these studies prompted the group to develop a monitoring matrix, which was used for several years to track progress among members in meeting their commitments to involving partner countries in evaluation work.

At subsequent meetings joint evaluations featured prominently on the agendas, both in terms of updates on on-going evaluations, but also as debates on how to further joint evaluations and, in particular, how to engage partner countries more.

In February 2012 France hosted a workshop on lessons learned from joint evaluations with focus on engaging partners in joint work. The group agreed to serve as a hub to maintain the partner country network coming out of the Paris Declaration Evaluation and to identify opportunities for further joint work and link this to evaluation capacity development.

More recently, there has been a bit of a backlash to joint evaluations because of the high transaction costs often associated with co-ordinating multiple partners and the increased accountability pressure on member evaluation units to assess the contribution of their country’s own development assistance. Yet, the scope for different forms of collaboration continues to expand, with members increasingly sharing evaluation plans and finding ways to borrow or build on what has been done by others – for example by sharing context analyses or partnering up for field visits, even while carrying out separate evaluations. Some recent examples of silent partnerships in joint evaluation, such as an initiative by the KfW Entwicklungsbank, Agence Française de Développement and the European Investment Bank to agree on a shared evaluation approach that can be used by all three, are also encouraging as these encourage an efficient division of labour while meeting different partners’ needs. At the same time, the importance of partner countries taking a leadership role in development evaluation was given new emphasis in the 2011 Busan declaration on Effective Partnership for Development Co-operation. Some members clearly see the future of joint evaluation rests in partner countries, where the focus will be on supporting country’s to assess their own development policies and programmes, in collaboration with external partners.

Joint Evaluations at Country or Programme Level

One of the unique features of the network is to enable direct interaction between its members. In 2003 we took the initiative to create a "market place" for joint evaluations, where members could present a proposal and get an immediate response from other members. This led to several joint evaluations that tackled country level or program level issues; not as ambitious as the global or regional initiatives, but often very practical and helpful to the countries concerned, reducing the evaluation burden, and useful for the donors as it led to evaluations with a broader perspective than just their own efforts. Creating this enabling environment for collaboration is one of the great benefits of a network like EvalNet.

Rob D. van den Berg
DAC Evaluation Network chair, 2001-2004
4.4 Evaluation of multilateral organisations and agencies

In 1988 interest began to focus on multilateral organisations. The smaller donors were particularly interested to look at both effectiveness of the organisations and at their internal evaluation capacity. They were driven by the need to understand how well multilaterals performed to inform their own decision making. Only the United States had to a certain degree systematized assessment of the effectiveness of the multilaterals and the United Kingdom had a ‘litmus test’ for assessing the multilateral agencies’ evaluation systems.

Norway produced an approach paper in 1990 that discussed a number of issues related to both aspects. The idea of joint evaluations received general support while the idea of looking at the quality of their evaluations was less enthusiastically received.

Over the next years several UN agencies were evaluated by smaller groups of donors: UNICEF (1992), UNFPA (1993), the Global Environment Fund (1994), WFP (1994 and 2005), UN Research Institute for Social Development (1997), UN Capital Development Fund (1999), IFAD (2003), UNFPA and IPPF (2004), and of the International Trade Centre (2006). While valuable and useful, these evaluations were time and resource intensive exercises which led to a degree of fatigue amongst donors and the multilaterals themselves.

As a result Denmark in 2004 reintroduced the idea of assessing the multilateral agencies’ own evaluation capacities with the hope of identifying reliable evaluation systems that could be used by all stakeholders. They suggested a peer review approach. An approach was developed by a joint DAC Evaluation Network and United Nations Evaluation Group (UNEG) taskforce and UNDP volunteered to be “guinea pig” to test out the method. UNDP’s Evaluation Office was peer reviewed by a panel composed of evaluators from Denmark, International Development Evaluation Association (IDEAS), IFAD, the Netherlands and the United Kingdom in 2005. The pilot review was used by UNDP to develop a new Evaluation Policy and by the joint task force to further refine the methodology. UNICEF was peer reviewed in 2006, still as a pilot testing the methodology, by a panel composed of evaluators of the African Development Bank, Canada, IDEAS, Ireland, Norway, and UNIDO and the methodology further refined.

Concurrently, in 2005, UNEG initiated work towards establishing a “quality stamp” on the evaluation function in member organisations in the hope that quality evaluation functions would reduce the need for individual donor assessments.

These two work streams came together when in 2006 the DAC Evaluation Network and the UN Evaluation Group (UNEG) formalized the joint task force “to establish an internationally recognized form of peer review of the evaluation function in UN agencies.” The task force aimed to promote peer reviews of interested UN organisations. The peer review would focus on the evaluation function within the organisation; take its central evaluation unit as a starting point and working towards understanding how evaluation operates at the various levels, in order to review the quality of the function in light of the objectives of the organisation and international standards. Reviews look at existing capacity and need for capacity strengthening, use of relevant and appropriate evaluation methodologies, and the quality of the evaluations undertaken, in terms of preparation, implementation and reporting. Another key focus is on looking at collaboration in evaluation with local partners and stakeholders, as well as harmonisation and co-ordination with other external partners.
Subsequently several agencies were peer reviewed: the World Food Programme (2007), OIOS and GEF in 2009; UNIDO in 2010 and UNEP, UN Habitat and FAO in 2012. The reviews have been very useful to the organisations and agencies as a means to improve, to differing degrees the policies, independence and quality of evaluation. They have, however, not reduced to any significant degree the donors’ demands for independent, evidence-based assessment of the humanitarian and development effectiveness of the programs of multilateral organisations – partly because of the political perception that individual assessments are needed. Still, it could be argued that these assessments have shifted in focus with a higher recognition of the evaluative evidence that is emerging in the organisations themselves. This broad assessment is confirmed by a lessons learned study of the peer reviews being conducted in 2013.

Other instruments have been developed to supplement the peer reviews and fill the information gap:

In 2011 the group endorsed a new approach, pioneered by Canada, for assessing the development effectiveness of multilateral organisations. The assessment relies on the content of published evaluation reports produced by the multilateral organisations supplemented with a review of documents and consultation with headquarters staff. The method uses a common set of assessment criteria derived from the DAC evaluation criteria. It was pilot tested during 2010 using evaluation material from the Asian Development Bank and the World Health Organisation. The first reviews using the approved methodology were completed in 2011/2012 of UNDP and WFP followed in 2013 of the African Development Bank and UNICEF. It is envisaged that members of the Group applying ‘division of labour among themselves’ will be involved in rolling out a series of such reviews implemented in a coordinated way with the Peer Reviews and to complement the reviews by MOPAN

Ultimately, such development effectiveness reviews should be replaced by regular, evidence-based, field-tested reporting on development effectiveness provided by multilateral organisations themselves.

3 All of the reports, and further information on assessing multilateral effectiveness are available on the OECD DAC website: www.oecd.org/dac/evaluation/evaluatingmultilateraleffectiveness.htm
5. Reaching out – Collaboration with partners

“Seeking ways to promote and support developing countries' evaluation capabilities.” - 1982 Mandate

Developing countries – the recipients of aid – clearly have an important role to play in assessing the effectiveness and results of development co-operation. The partner countries are perhaps more directly concerned with using evaluation to understand and improve results. The importance of involving partner countries in evaluation was acknowledged early on in the DAC and by the time the Paris Declaration officially called for mutual accountability, the notion of collaboration with partners was already well established in the development evaluation community. In 1985 Sweden took the lead in establishing collaboration between the group and partner countries. The impetus was a Swedish review of support to building evaluation capabilities – primarily through support to audit offices – which showed disappointing results. Several members were hesitant at the idea of a workshop with partner countries “until it had clearly been established that there was a real demand to be met.” Others noted the difference in perspective between authorities in developing countries and the managers in donor agencies. Sweden and Norway together with the Secretariat continued exploring the possibility of organising a seminar with partners, recognising the importance of better collaboration. Preparations took time and resistance in the group from those who were primarily concerned with donors’ own accountability needs had to be overcome.

The 1987 seminar opened the door to a common understanding of the role of development evaluation although there were some differences of view about the interest (demand) and capacities of the developing countries between developing countries and donors. It was clear that there were wide variations in the extent to which evaluation had been developed in different countries and regions, requiring flexible responses from donors in each case. Regardless of these variations in evaluation capability, however, all developing-country participants expressed clear interest in the possibility of expanding their activities in this field. A theme that ran right through the seminar was that evaluators in aid agencies could and should start to think in terms of moving away from running their "own" evaluation programmes, and should do more to integrate their activities with the evaluation processes in the developing countries.
A seminar with participants from developing countries: A moment of controversy

At the 6th meeting in 1986 it was proposed to have a joint seminar with representatives from evaluation offices from developing countries. The aim was to become better acquainted with those in developing countries who have a shared interest in the evaluation of development programs, learn more from each other and promote the role of evaluation.

When the proposal was put forward there was general agreement with one exception. One delegate from a donor country was strongly opposed, on principle, to inviting the developing countries to participate in a DAC seminar even for a onetime DAC meeting. The DAC and its meetings were exclusively for donors. The delegate from Norway and others generously offered to cover the travel costs for representatives from developing countries: a potential sticking point. But still no consensus; thus, no seminar.

I suggested that we adjourn for coffee while we sorted it out. The opposing delegate agreed not to intercede and allow the consensus to stand.

The joint seminar was held in the DAC’s chateau de la Muette, 11 – 12 March 1987. Approximately 75 delegates came from evaluation offices in countries in Africa, Asia, Latin America, Europe and North America. The seminar was chaired jointly by the head of evaluation in India and me.

W. Haven North, USA
Chair 1985-1988

The success of this seminar led the group to embark on a series of regional seminars, the first of which was held in 1990 in Abidjan, co-sponsored with the African Development Bank. The seminar identified a number of ways in which donors could support evaluation capacity building in Africa – mainly by providing technical assistance. The seminar report “Evaluation in Africa,” however did not go in-depth on the real problem: the lack of accountability of partner governments, “Since evaluation can serve as a source of information to the public on the use of public funds, evaluation can be seen as a tool of democracy. The principle of public access to information is, however, not accepted in all countries.”

A regional seminar on “Performance Evaluation in Asia and the Pacific” was organised jointly by the DAC Evaluation Group and the Asian Development Bank and held in Kuala Lumpur in May 1992. Although more focused on technical aspects of performance evaluation, it also touched on the lack of political demand for evaluation and institutional capacity building. An action plan for strengthening performance evaluation in the region was prepared during the seminar. A follow up survey a year later received limited response, and little activity could be traced carrying the agenda forward. However, the survey did show that half of the respondents had started networking and sharing experiences.

In November 1993 the Inter-American Development Bank and the DAC Evaluation Network organised a similar seminar for Latin America and the Caribbean in Quito, Ecuador.
After this meeting Denmark, in 1994, proposed to gather experiences from donors’ support to evaluation capacity building. Given the limited amount of evaluations on the subject, a survey was conducted, and the issue was subject to a special session at a seminar in Canberra in 1996 hosted by Australia. The report “Evaluation Capacity Building - Donor Support and Experiences” was published in 1996, and its recommendations were:

- Continue to promote a common methodology and terminology shared by donor and host countries.
- Encourage and facilitate members in coordinating sector and country evaluations and in sharing them with host-country evaluation institutions.
- Continue supporting regional networking to promote the exchange of expertise and to share evaluation information among recipient and donor countries.
- Support regional and south-south seminars and training.
- Facilitate the exchange of information on donor evaluation capacity building support activities and promote the assessment of these activities.

These recommendations, although not revolutionary, have provided a sound basis for subsequent work of the group, as will be shown below.

A second Regional Seminar and Workshop on “Monitoring and Evaluation Capacity Development in Africa” was held in Abidjan in November 1998. It reported a more optimistic view than in 1990, concluding that “… the trend towards more responsive, accountable and efficient government is lasting and will influence future development strategies.” The workshop also underlined that not all countries were at the same level of institutional development and that each country followed its own specific path on capacity development according to its administrative culture.

This was the last of this type of seminar. It is difficult to ascertain the direct outcomes of the seminars, but the topic remained important and some effects are evident: several DAC Members, as well as the multilateral development banks and the UNDP, actively pursued the dialogue with the participants established at the regional seminars. In some cases this was done through joint evaluations or provision of technical assistance or both, and through support to a long list of activities supported by individual members or multilateral initiatives:

Regional seminars/conferences in Spain, South Africa and China have been organised by the banks and UN agencies. Japan organises annual workshops on Overseas Development Assistance Evaluation focusing on Asia and the Pacific. Several donors and UN agencies have supported the African Evaluation Association (AFREA) founded in 1999 and the International Development Evaluation Association (IDEAS) which was launched in Beijing in 2002. Individual donors provide financial support and bursaries for the conferences organised by the associations as well as for participation in the conferences by the American Evaluation Association and the European Evaluation Society.

In 2003, a workshop “Partners in development evaluation – learning and accountability” was hosted by France. The workshop involved partner countries as well as many civil society organisations. Discussions and a number of room documents looked at partner capacity and involvement in joint work and the role of donors in supporting partner capacity. The resulting workshop report pulled together this knowledge.
At a workshop on joint evaluation held in Nairobi in April 2005 in collaboration with developing country partners, the issue of capacity was raised in the context of enabling developing country stakeholders to take on more active roles in joint evaluations. One of the key recommendations from the workshop was that “developing country governments should be supported to build their institutional capacity for initiating and leading joint evaluations and all partners need to look at innovative ways of providing funding for aid recipients to build their evaluation capacity.”

A 2006 fact-finding study led by Japan found that extensive evaluation capacity development work continues. The study included 26 agencies (21 bilateral and 5 multilateral) and a total of 88 separate evaluation capacity development interventions were reported including training and scholarships (37), workshops (31), technical support to projects/programmes (18), financial support (18), joint evaluations (22), dialogue at policy levels (10) and other types (8). The diversity of interventions in this area is characteristic of both the multi-dimensional nature of capacity development work, and of the lack of a clear definition of what exactly constitutes capacity development (which leads to variation in reporting). The fact-finding study came up with many suggestions for effective capacity development measures which have helped inform member efforts.

At the Third International Roundtable on Managing for Development Results, held in Hanoi in February 2007, capacity issues were a key dimension in the discussions, underlining the importance of renewed and focused attention to evaluation capacity.

A very prominent evaluation capacity building initiative is the International Program for Development Evaluation Training (IPDET) founded in 2000 by the World Bank and Carleton University in Ottawa, Canada.
with support from several DAC members who provide funding for stipends from developing countries. Many staff of the member evaluation offices teach at IPDET and DAC evaluation normative standards and guidance work are also used there.

In 2008 the Independent Evaluation Group (IEG) of the World Bank proposed to support centres to provide in-region capacity development and technical assistance services, and global learning to strengthen practical knowledge-sharing on monitoring and evaluation across regions. The program would extend the expertise and experience gained through IPDET and its offspring SHIPDET in Shanghai. The proposal was supported by several donor countries and the development banks. By 2011 Regional Centres for Learning on Evaluation and Results (CLEAR) were established in Anglophone Africa, East Asia, Francophone Africa, Latin America and South Asia.

AfREA, CLEAR, IPDET, IDEAS and the many national evaluation societies have all contributed to the emergence of an evaluation profession in developing countries. Many of these groups have been actively supported by DAC Evaluation Network members and the group has also collaborated with these networks on various occasions.

In 2009 the Secretariat presented an overview of the Group’s past work on capacity development (summarised in the section above) and a proposal to focus on the capacity implications of members’ own evaluation policy and practice. There was general recognition of the need to continue to collaborate with other actors on capacity development and to focus more on the challenge of addressing the demand side of evaluation. A concrete result of the discussions was that the Secretariat produced a brief and succinct “Tips for capacity friendly evaluation in development agencies” (OECD 2010) which provides useful advice to evaluators on how best to support evaluation capacity development through their regular evaluation work. Finland now leads the network’s task team on supporting evaluation capacity development. The team works to encourage collaboration, share experiences and develop a clear understanding of how DAC members can best support partner countries in strengthening capacities and building an enabling environment for evaluation.

The DAC Evaluation Group has come a long way in terms of openness and outreach to other communities. It has reached out to partners, civil society associations and to new and emerging donors. Particular attention must be given to managing the process of collaboration to ensure it adds value – involving many players is more cumbersome in many ways but can achieve better results in the long run.
6. Advocating and sharing best practice in evaluation – The normative role

“Strengthening exchange of information, experience and co-operation between national and, as appropriate multilateral evaluation units, with a view to:

- improving the evaluation activities of individual members;
- encouraging standardization of methodological and conceptual frameworks”

- 1982 mandate

The early discussions around how to improve the evaluation practices of the members of the group focused very much on sharing and discussing methods such as the use of logical framework, the need to see projects in a broader macro-economic framework, the difficulties of evaluating non-project aid, the use of social cost-benefit analysis, and the role of controls (and with-and without / before-and-after analyses). All these were issues that would stay with the group for the next three decades.

At the meetings throughout the 80’s considerable time was devoted to sharing experiences at every meeting. These exchanges covered methodological issues as well as systemic and political issues and were deemed very useful by the members. The tradition has continued in that at every meeting time is set aside for exchange of information. Today this element of the formal agenda is mainly about major institutional or political changes while methodological issues are dealt with in work streams and in the informal workshops and meetings around the formal meetings.

6.1 Methods and procedures in aid evaluation

The United Kingdom suggested a systematic approach already at the second meeting, suggesting the group produce a handbook of evaluation methods and procedures. The DAC Secretariat pointed out that the group’s mandate did not provide for the production of a major report focused on methodological and procedural aspects of evaluation. In respect of this and the sensitivities of some members (who were hesitant to produce normative work) it was underlined that it should be “a collection of information on existing evaluation methods and procedures.” This is borne out by the title of the final “Methods and Procedures in Aid Evaluation – A Compendium of Donor Practice and Experience” (1985). The compendium was the first in a series of publications on principles, norms, standards and guidance documents’. The group decided not to produce specific guidelines as this was considered to be inflexible. Instead the focus was on developing agreed upon normative guidance that could inform member practice, improve the quality of evaluations and create a basis for collaboration.
6.2 The DAC Principles Evaluation of Development Assistance

The DAC Principles for Project Appraisal which were approved by the DAC in 1988, and to which, as mentioned above, the Evaluation Group’s first report on evaluation findings in 1984 and the report on sustainability in 1988 had contributed significantly, constituted a challenge to the group: it did not have a set of principles for its own work. In 1989 the Group decided, despite some reservations about the utility of principled standards, to work on a set of Principles for Evaluation and tasked the DAC Secretariat with producing an outline. The outline was discussed in June 1989 and various members committed to produce drafts of sections of the “principles.” The first draft was discussed in February 1990 when it was agreed to produce a concise set of principles rather than a lengthy all-encompassing documents. Several drafting meetings were held in conjunction with the regular meetings in Paris and the first regional seminar in Abidjan. The final version (DAC Principles on Evaluation of Development Assistance was agreed to in March 1991 and subsequently approved by the DAC.

The DAC Principles for Evaluation of Development Assistance is probably the most important single product by the DAC Evaluation Group. The basic principles of independence, impartiality, credibility, transparency, utility, timeliness, and inclusiveness were not new, but it was important to establish them as principles. A very important element of the principles was the definition of evaluation:

“An evaluation is an assessment, as systematic and objective as possible, of an on-going or completed project, programme or policy, its design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, developmental efficiency, effectiveness, impact and sustainability. An evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision-making process of both recipients and donors.”

From that definition five words or concepts have had a profound impact on development evaluation: Relevance, Efficiency, Effectiveness, Impact, and Sustainability. These have become the widely accepted criteria upon which every development evaluation will base its assessments—or make excuses for why not.

A comprehensive review of the DAC Principles on Evaluation of Development Assistance published in 1998, concluded that the key principles remained useful and valid and did not need revision. However, it suggested complementing them with guidance for evaluators on good/best practices in certain fields, notably in areas such as feedback and communication of results; how to promote an evaluation culture within Ministries and agencies; country programme evaluations, humanitarian aid, joint evaluations and partnerships in evaluations. This review, together with the members’ needs for more direction in specific areas, helped shape the group’s agenda.

The five evaluation criteria from the ...OECD/DAC have been a strong foundation for international development evaluation since 1991. They have been the most prominent and widely adopted criteria used for aid evaluation by most bilateral and multilateral donor agencies, as well as international nongovernmental organizations (INGOs) of development evaluation.

T. Chianca, 2008
6.3 Guidance in specific areas

The first guidance developed by the group was the Guidance for Evaluating Humanitarian Assistance in Complex Emergencies published in 1999. It was an offspring of the Rwanda Evaluation and focused on how to design and manage evaluations of humanitarian assistance.

The second was “Effective Practices in Conducting a Joint Multi-Donor Evaluation” in 2000. It was the result of a comprehensive study led by the United States. As experience grew with the increasing number of joint evaluations, the Evaluation Group decided to conduct a follow up study. This work was led by Denmark and resulted in the report “Joint Evaluations: Recent experiences, Lessons Learned and Options for the Future” in 2005. The report was very large and comprehensive and the Group requested the Secretariat produce a condensed version. The resulting “Guidance for Managing Joint Evaluations” was published in 2006.

In 2005 the DAC Network on Development Evaluation and DAC Network on Conflict, Peace and Development Co-operation, on the initiative of Utstein Group, began working together to improve methodology for evaluating conflict prevention and peace building interventions. A Steering Group comprising Norway, the United Kingdom, Sweden and the Secretariat led the comprehensive work that resulted in draft guidance produced in 2008 for a trial application period. The guidance was tested in a series of evaluations in South Sudan, the Democratic Republic of Congo, Afghanistan, Sri Lanka, Haiti and elsewhere. The lessons from implementing these evaluations were brought together in “Evaluating Peace Building Activities in Settings of Conflict and Fragility – Improving Learning for Results” (OECD 2012). This work embodied several key areas of the group’s activities: normative work (in developing the guidance), joint evaluation (in testing the guidance in various conflict settings) and collaboration with DAC policy groups (in this case, INCAF).

6.4 Results based management

The role of results management systems in development co-operation agencies constituted an important area for the group, primarily because results systems constitute an important input to rigorous evaluation and the two are closely linked in practice. In order to respond to the need for an overview of the rapid evolution of RBM, the Group initiated a study of performance management systems in 1999.

The ensuing report “Results Based Management in Development Co-operation Agencies” based on a limited sample, was very comprehensive and was one of the first reviews of results management in development. It contributed to a number of agencies’ considerations on how to take this agenda forward internally, as well as contributing to various international events such as the 2002 Multilateral Development Banks’ round table on better measuring, monitoring and managing for results. The issue stayed on the agenda, e.g. the glossary is both containing RBM and evaluation language. In 2012, monitoring and managing for development results was again at the top of the DAC’s political agenda and the group again discussed experiences with results based management systems, based on corporate evaluations done by some members.

---

4 Later called the International Network on Conflict and Fragility (INCAF)
5 The Utstein Group is a group of key likeminded development ministers from Germany, the Netherlands, Norway, Sweden and the UK. Canada has since joined the group.
6.5 The Glossary of Key Terms in Evaluation and Results Based Management

In 1999 it was suggested that a glossary of terms used in evaluation and results be produced. As new concepts such as results based management emerged it had become necessary to clarify language and build an agreed common basis for collaboration between institutions. A Task Force, chaired by the World Bank, led the overall project, in collaboration with the Secretariat. France took the lead on the French version, whilst the Inter-American Development Bank produced the Spanish translation. The French version of the glossary was developed by an independent working group. The “Glossary” was published in 2002. The process was not easy as is evident from the foreword: “The process has been guided by the highest considerations of clarity and conciseness and a spirit of collaboration and compromise in terms of the willingness of major development agencies and banks not to impose their specific vocabulary on others.” Another example of the difficulties is that there was more than 20 definitions of ‘impact evaluation’ identified in member documents and academic literature at the outset of the process.

The Glossary has now been translated into 16 languages including Arabic, Kiswahili, Turkish and Swedish. It is probably the most popular and used product of the group. The glossary illustrates the added-value of the group: by creating a platform for sharing experiences and building consensus the group can create unique products that could not be produced by an individual member.

6.6 The DAC Quality Standards for Development Evaluation

When the Evaluation Group discussed the draft Review on Gender and Evaluation in 2003 the Group took note of a clear criticism of the varied, often poor, quality of the evaluations assessed. Australia offered to do a more systematic assessment of existing evaluations. However, the discussion revealed that several members were hesitant to subject their evaluations to a transparent quality assessment. Instead it was agreed to focus on developing a set of agreed norms and standards, which could be used for quality checks and would help improve the quality of evaluations overtime. The work was taken forward by Australia and Denmark and involved a series of workshops both internally and with partners. One such workshop was held in Australia in June 2005 with partner countries from the region participating. The draft was also discussed at two sessions at the IDEAS’ 1st biannual conference in New Delhi 2005. The draft standards were approved for a three-year test phase in 2006. They were revised in several versions by the Secretariat in 2009 based on a survey of the use of the standards during 2006-2008, a 2009 workshop held in Auckland with partner countries, and comments submitted by the members of the group. The standards were approved by the group in January 2010 and endorsed by the DAC on 1 February 2010. The standards are now available seven languages.

In June 2010 the Secretariat pulled together the Principles and the Quality Standards and other key documents into the, “Summary of Key Norms and Standards”, for the first time making available all of the key elements of the DAC evaluation framework in a single document, in an electronic format. This brief compilation is a useful international reference point for anyone interested in development evaluation.
6.7 Assessment Tool for DAC Peer Reviews and Evaluation Managers

In 2004 the Secretariat suggested that the Evaluation Group work more systematically to strengthen members’ evaluation systems. The background was a study of eight years Peer Reviews’ findings on evaluation which indicated that many members’ evaluation functions had difficulties meeting both old and new challenges, notably the call for more rigorous evaluations of impact and evaluations of new (complex) aid modalities such as budget support, sector and sector budget support. The study also revealed a need for more attention to the way future Peer Reviews could usefully address the complexity of these new demands placed on evaluation units. The Secretariat and France, with input from members, produced ‘An Assessment Tool for DAC Peer Reviews and Evaluation Managers’ which, as the title indicates would also be useful for self-assessment.

This was in 2007 followed by “A Comparative Study of Evaluation Policies and Practices in Development Agencies” (of 10 agencies) which concluded that “The picture that emerges here is a rather composite one. Evaluation units share a number of common features and challenges, and are on a similar journey (although at different stages), from a relatively straightforward model of project evaluation aimed at internal management and accountability, towards a more complex model of policy, country and partnership-led evaluation, which requires new skills, roles and organisational arrangements.”

Supporting the DAC Peer Reviews

Peer reviews of OECD Development Assistance Committee (DAC) members are based on the dual principles of peer pressure and peer learning. The reviews are the only international processes to regularly examine key bilateral development co-operation systems. Following a request by the DAC, a Client Survey of the OECD DAC Peer Reviews was conducted by the Working Party on Aid Evaluation and its findings were reported to the DAC Senior Level Meeting in December 2002. The client survey found that peer reviews were clearly of substantive value to members and many members reported that the peer reviews have had high or some impact on their aid policies and programmes. However, despite their overall positive assessment of the importance and impact of peer reviews, DAC members sought improvements in three main areas: experience-sharing and collective learning, methodology of peer reviews, and thematic coverage of peer reviews. As a result, it was agreed to have regular methodology meetings to deal with broad review issues and further develop the methodology. The first DAC Peer Review Methodology meeting took place in January 2003 and has since been an important feature for dialogue on how to continually improve the peer review process.

In view of the increasing demands on evaluation as a key component of the accountability, management and learning structures of development agencies and the need to identify examples of good practice, the Secretariat took the lead on conducting a study of members’ evaluation systems and resources. Published in 2010 the study “Evaluation in Development Agencies” describes the role and management of evaluation in development agencies and multilateral banks, based on questionnaires, findings from peer reviews by the DAC, interviews and a literature review. The study includes information about the specific institutional settings, resources, policies and practices of each of the DAC Evaluation Network’s 32 members. The study identifies major trends and current challenges in development evaluation, covering: human and financial resources, institutional set-ups and policies, independence of the evaluation function, reporting and use of evaluation findings, management response systems, donor co-ordination, joint evaluation, and the involvement of partner countries in evaluation work. The review was structured around the assessment tool and the core evaluation principles.
6.8 Impact Evaluation

"When Will We Ever Learn? Improving Lives through Impact Evaluation," published by the Center for Global Development in 2006, called on developing countries, aid agencies, foundations and NGOs to close the evaluation gap by adopting good practices in terms of independently evaluating the impact of their own programmes and by collectively creating and financing an independent entity to promote and finance rigorous impact evaluations of interest to the entire international community. Even before the publication the initiative sparked a lively, if not heated, debate on evaluation methodologies as indicated in the box below.

The impact evaluation debate

The “Evaluation Gap initiative” started by the Centre for Global Development (CGD) in 2004 triggered a heated debate. Bill Savedoff from the CGD presented a discussion paper at the second meeting of the Network on 9-10 November 2004. Savedoff advocated for using experimental approaches with randomized control trials as the Golden Standard of development evaluation in order to find out “what works” in development.

The minutes from the meeting merely indicate some of the issues of the intense debate that accompanied the CGD initiative. Many members of the network were deeply challenged, not to say professionally offended by the notion that their evaluations would not be rigorous enough. Nevertheless the CGD paper and the subsequent creation of the International Initiative on Impact Evaluation (3IE) did stimulate the search for and development of approaches and methods to evaluate the effectiveness of development interventions.

Although not all evaluation departments were in favour of randomized approaches the CGD initiative gained support from policy and research departments. Impact evaluations have since become a staple of the evaluation programs of many development evaluation departments including the newly formed independent evaluation institutions.

Eva Lithman, Sweden, Chair (2004-2006)

Following a number of informal consultations the Network of Networks on Impact Evaluation (NoNIE) was created in 2005 as a forum to continue discussions on methodology and define a common platform for addressing issues of rigour in assessing impact. NoNIE brought together the DAC Evaluation Group, the United Nations Evaluation Group (UNEG) and the Evaluation Cooperation Group of the multilateral development banks (ECG), as well as the International Organisation for Cooperation in Evaluation (IOCE). NoNIE published the very useful guidance: ‘Impact Evaluations and Development: NoNIE Guidance on Impact Evaluation’ in 2009. The same year UNICEF, 3ie and NoNIE organized a conference on ‘Perspectives in Impact Evaluation’ in Cairo that was well attended by people from many different disciplines and evaluators from partner countries, especially from Africa. NoNIE is supported by many members of the evaluation group.
6.9 The utility of the normative work

The normative work has proven much appreciated by the membership and beyond. The United Nations Evaluation Group and the Evaluation Cooperation Group of the multilateral development banks have built upon and adapted the normative work of the DAC Evaluation Group to suit their particular needs.

The thorough consultative processes on norms and standards – two years for the Principles, three years for the standards and for the glossary – may seem very long. However, in the end, such collaboration and the negotiation processes created ownership resulting in the products being used extensively in the evaluation community. Not least the Principles have demonstrated their value throughout their long existence and continued use. In 2013 the Secretariat developed a new publication, 12 lessons on evaluating development activities – part of a series of 12 lessons from the OECD DAC. The document draws on examples from the group and from DAC peer reviews, and provides a new resource for development agencies with tips and real-life examples on settings up a strategic framework for evaluation, delivering quality evaluations and supporting a learning culture.

"Over recent years, OECD-DAC has exercised a leadership role in this professionalizing [of development evaluation]. These principles, criteria and standards have been broadly adopted by donors and now serve as the generally recognized benchmarks for development evaluation.

- L. Armitage (2011)
The Secretariat’s 2010 review of members found that core normative pieces, including the criteria and the DAC Quality Standards on Development Evaluation are widely used by DAC members and other development actors. For instance, nearly all DAC member country’s evaluation policies specifically refer to the DAC Evaluation Principles and the importance of meeting DAC evaluation standards (OECD 2010).

6.10 Supporting new members and engaging beyond the DAC membership

The DAC evaluation group’s normative framework has also provided a useful basis for supporting both new and old members in professionalizing their evaluation systems. The normative work provides a useful entry point for engaging with other countries beyond the DAC. Turkey’s international development agency, for example, collaborated with the group to produce a Turkish version of the glossary. Many countries have found such translations to be a useful way to begin discussing evaluation concepts and to start bringing their system in-line with international good practice. Working with the United Arab Emirates and the evaluation departments of the Islamic Development Bank and the African Development Bank, a trilingual Arabic/English/French version of the DAC quality standards was produced. This collaboration also created an opportunity for members of the network and the Secretariat to visit Abu Dhabi and Jeddah and support capacity development efforts in the region.

Network Chair Nick York (left) and HE Dr. Ahmad Mohamed Ali, President of the Islamic Development Bank Group during a workshop to support evaluation capacities in the region and launch the trilingual (Arabic/English/French) version of the Quality Standards for Development Evaluation in Jeddah, Saudi Arabia. (Photo: Kennedy-Chouane, January 2012).
7. Sharing and communicating evaluation results

“Strengthening exchange of information, experience and co-operation between national and, as appropriate multilateral evaluation units.” – 1982 mandate

One issue came to the forefront early – that of use of evaluation results and how to feedback evaluation results to policy makers and programme designers. Evaluations did not always show positive results of aid, which meant they were not necessarily popular with policy makers. Several members had seen negative findings repressed or evaluation findings used selectively to suit political interests. The question was: how to get the message across? In 1987 the Group conducted a survey of donors’ experiences and discussed the issue at a seminar hosted by Canada in September the same year. This was followed by another survey. Canada pulled the information together and “A Review” of Donors’ Systems for Feedback from Aid Evaluation” was published in 1990. The report was another example of documenting and sharing experiences, and apart from the use made by individual agencies it informed the DAC Principles for Evaluation of Development Assistance. After the review of the Principles in 1998, which included a survey of evaluation users, it was concluded that feedback and communication practices needed to be improved. To take this work forward Japan hosted a workshop in Tokyo in September 2000 to share experiences and draw lessons from current best practices. A Background Synthesis Report and the workshop report were compiled and published in 2000 as “Evaluation Feedback for Effective Learning and Accountability.” This report illustrates many different approaches to feed-back which are still useful.

The exchange of information between members in the early years was, seen with today’s eyes, rather primitive: In addition to the Exchange of Information session at every meeting, evaluation plans and evaluation reports were exchanged by mail. This was not effective as many members did not share information. In 1986 Canada presented a study on ‘automated retrieval systems’. About a quarter of the members had some kind of ‘automated system’ for storing and retrieving evaluations. Discussing the report, members were divided roughly between those who had already ventured into automated systems and found them not very useful, or who feared an indigestible flood of material, and those who saw automated systems as a way of reducing the flow of information to essentials. Time has proven them all wrong. Canada continued exploring ways of collecting and sharing evaluation reports electronically and in 1988 the “DAC Evaluation Reports Inventory System” was launched. Initially the ‘Inventory’ was restricted to members only (as several agencies did not publish their evaluation reports) and reports were submitted on diskettes and the inventory updated once a year and distributed by diskette. By 1994, 29 agencies contributed to the inventory and it contained more than 1,500 evaluation abstracts and 945 planned evaluations. In 1998 the Inventory migrated to the Internet and most of its content became available to the public. In 2005 the inventory, which by then contained more than 1300 full reports was replaced by DEReC: the DAC Evaluation Resource Centre online. In 2000 the group established its own (restricted) website on OECD DAC Website. And in 2005 a public website was launched. Since then, in tune with the rapid development of the Internet, the Secretariat has kept the website lively by adding features and links, updating documents and reports as well as using Facebook, Twitter and YouTube to convey messages to a wider public.
8. Some concluding remarks

Looking back over the past 30 years it is evident that the Evaluation Group has fulfilled its mandate. It has

- strengthened exchange of information, experience and co-operation between national and multilateral evaluation entities as well as regional and global evaluation societies;
- contributed to improving the evaluation activities of individual members;
- established principles, norms and quality standards for evaluation that have been models for others;
- improved donor co-ordination in the planning of major evaluation studies;
- contributed to improved aid effectiveness by synthesising evaluation studies providing lessons for policy making and operational lessons for project and programme planning and design;
- launched many joint or co-ordinated studies of aid effectiveness in major sectors and
- actively promoted and supported developing countries’ evaluation capabilities.

For an intergovernmental group like the DAC Evaluation Network, however, a clear mandate and support at the political level is needed to help establish the group’s formal status on the international scene and enable individual members to participate fully. The Network’s existence is not taken for granted – in fact its mandate is subject to a sunset date, which has meant that the group needs to continuously prove its relevance to a shifting development agenda. That it has done so is witnessed by the fact that its mandate has been renewed by the DAC every 2 years.

As aid modalities have shifted, the evaluation focus and the work of the Evaluation Group have also shifted, moving from project evaluations towards assessing sector or country programmes and aid modalities such as general budget support and humanitarian assistance. In methods, the emphasis has moved more towards assessing outcome and impact. In outreach the Network has been increasingly open to collaboration with partner countries as well as other networks and evaluation societies.

The aid and development landscape is changing rapidly: The developing countries taking control of their own development, the new aid providers, intensified engagement with the private sector and civil society, the increasing mixing development co-operation with trade interests, shrinking or stagnant development budgets in traditional donor countries and demands for value for money pose new challenges for the Evaluation Network.

The dynamic and inclusive nature of the way the Evaluation Group works has been key to its achievements. This dynamism does not come out of no-where. It comes from the engagement and enthusiasm of the members, small and large. I am confident that the Evaluation Network will continue to meet the challenges in the same collaborative and constructive way as in the past thirty years.
## Annex A.

### Bureau Members of the DAC Evaluation Group

<table>
<thead>
<tr>
<th>Chairs</th>
<th>Country</th>
<th>Dates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Network on Development Evaluation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mr Martyn Pennington</td>
<td>European Commission</td>
<td>Feb. 2012 - present</td>
</tr>
<tr>
<td>Mr Finbar O’Brien</td>
<td>Ireland</td>
<td>Nov. 2006 – Nov. 2008</td>
</tr>
<tr>
<td><strong>Working Party on Aid Evaluation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mr Rob D. van den Berg</td>
<td>Netherlands</td>
<td>May 2002 - July 2004</td>
</tr>
<tr>
<td>Mr Niels Dabelstein</td>
<td>Denmark</td>
<td>Nov. 1997 – May 2002</td>
</tr>
<tr>
<td><strong>Expert Group on Evaluation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mr Jean Quesnel</td>
<td>Canada</td>
<td>March 1991 – Oct. 1993</td>
</tr>
<tr>
<td>Mr Klaus Winkel</td>
<td>Denmark</td>
<td>June 1988 – March 1991</td>
</tr>
<tr>
<td>Mr Haven North</td>
<td>United States</td>
<td>June 1985 – June 1988</td>
</tr>
<tr>
<td>Mr Basil E. Cracknell</td>
<td>United Kingdom</td>
<td>Dec. 1983 – June 1985</td>
</tr>
<tr>
<td>Mr Allan Mau</td>
<td>Canada</td>
<td>March 1983 – Dec. 1983</td>
</tr>
<tr>
<td><strong>Vice-Chairs of the DAC Evaluation Group</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mr William Carlos</td>
<td>Ireland</td>
<td>2011 - present</td>
</tr>
<tr>
<td>Mr Henri Jorritsma</td>
<td>Netherlands</td>
<td>2011 - present</td>
</tr>
<tr>
<td>Mr Benoit Chervalier</td>
<td>France</td>
<td>2008 – 2011</td>
</tr>
<tr>
<td>Ms Penny Hawkins</td>
<td>New Zealand</td>
<td>2008 - 2011</td>
</tr>
<tr>
<td>Mr Nick York</td>
<td>United Kingdom</td>
<td>2008 - 2008</td>
</tr>
<tr>
<td>Ms Belen Sanz Luque</td>
<td>Spain</td>
<td>2006 – 2008</td>
</tr>
<tr>
<td>Mr Dominique de Crombrugghe de Loorinche</td>
<td>Belgium</td>
<td>2006 - 2011</td>
</tr>
<tr>
<td>Mr Finbar O’Brien</td>
<td>Ireland</td>
<td>2005 - 2006</td>
</tr>
<tr>
<td>Ms Satoko Miwa</td>
<td>Japan</td>
<td>2005 – 2006</td>
</tr>
<tr>
<td>Mr Niels Dabelstein</td>
<td>Denmark</td>
<td>2004 – 2005</td>
</tr>
<tr>
<td>Mr Daniel Kamelgarn</td>
<td>France</td>
<td>2004 – 2005</td>
</tr>
<tr>
<td>Mr Colin Kirk</td>
<td>United Kingdom</td>
<td>2002 - 2004</td>
</tr>
<tr>
<td>Ms Anne-Marie Fallenius</td>
<td>Sweden</td>
<td>1997 – 2001</td>
</tr>
<tr>
<td>Mr Niels Dabelstein</td>
<td>Denmark</td>
<td>1996 – 1997</td>
</tr>
<tr>
<td>Mr John Morris</td>
<td>United Kingdom</td>
<td>1994 – 1995</td>
</tr>
<tr>
<td>Ms Hedy I. von Metzsch</td>
<td>Netherlands</td>
<td>1991 - 1993</td>
</tr>
<tr>
<td>Mr. Jean Quesnel</td>
<td>Canada</td>
<td>1990 – 1991</td>
</tr>
</tbody>
</table>
References and Further Reading*


* Further information on the DAC Evaluation Group, current work and publications are available online:

www.oecd.org/dac/evaluation
The DAC Network on Development Evaluation – 30 years of strengthening learning in development

On the occasion of the group’s 30th anniversary, this document takes a look back at the story of the DAC Network on Development Evaluation and its members. It provides an overview of the group’s history, highlighting major accomplishments. It examines how the group has worked to promote collaboration and good practice in evaluation and to provide credible evidence to inform learning and decision-making.