

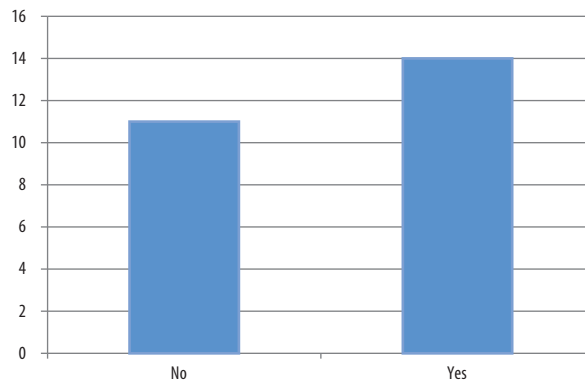
Better Aid

# Evaluation in Development Agencies



This is an extract from the “Evaluation in Development Agencies” 2010 publication, from the OECD DAC *Better Aid* series.

Figure 2.12. Publication of management responses to evaluations  
Are management responses and follow-up actions made public?



### Co-ordinating and sharing evaluation plans

As a step towards further systematising sharing of evaluation ideas and increasing collaboration in development evaluation, an online evaluation plan-sharing initiative was launched in 2008. The Secretariat collects evaluation plans and makes these available to all members on a website. The goal is that evaluation units will take into account the evaluation work of others and, in consequence, inform their own plans, and adjust when feasible to reduce overlap or maximise opportunities for collaboration. Eventually this platform should contribute to more joint work, increasingly transparent planning approaches and a more rational distribution of labour in the field of evaluation.

Twenty-nine evaluation units currently have their plans on the evaluation inventory.<sup>10</sup> All responding members stated that they share or plan to start sharing their evaluation plans with the Secretariat. A large majority (28 out of 37 respondents) said they consult other's evaluation plans to look for possible areas of collaboration when setting the evaluation programme or planning an evaluation (see Figure 2.13 for responses). Several specific groups were cited as stimulating these consultations: the Nordic Plus donor group, the Evaluation Cooperation Group (ECG) for multilateral agencies and most frequently the DAC Evaluation Network. Several members indicated that there is a specific small group of "like-minded" donors or donors with shared languages or working with the same partner countries whom they contact or collaborate with on a regular basis. It is clear that the majority of co-ordination takes place among these small subgroups.

Those units that did not consult the evaluation programmes or plans of other evaluation departments for possibilities to collaborate gave several reasons for this. Many commented that the primary goal of evaluation is to respond to the needs identified within their own agency, others that they lacked staff time to do so. Three members stated that finding useful partnerships was too difficult due to their specific type of aid modalities (for example, loans or private sector support), work in a limited number of countries or language barriers.

It is not yet clear from available data what impact the sharing of plans has had. However, comments from several members indicate that even though others' plans are consulted this does not always result in concrete joint evaluation work being identified or undertaken.

Nearly all responding members (36) refer regularly to evaluation reports completed by other donors, as shown in Figure 2.14. Referring to others' evaluations of similar

development interventions is viewed as a potential first step towards reducing redundancy in evaluation and increasing lesson learning. However, comments in response to this question and descriptions of the evaluation planning process indicate that evaluations are generally consulted as a point of reference or background research for further work, and are not (yet) used instead of undertaking new evaluations. It is hoped that a general improvement in evaluation quality and standardisation through measures such as the agreed DAC Quality Standards for Development Evaluation will eventually make it easier to learn from others' work rather than repeating all or part of it. Evaluations undertaken primarily with an accountability focus for an individual agency tend to have agency-specific lessons that are harder to "borrow" or generalise, limiting the scope for joint work.

Despite progress in sharing evaluation plans and a high level of collaboration among members, it remains to be seen whether member agency practice has actually become more joint in terms of setting evaluation plans. Responses to the member questionnaire regarding the planning process demonstrate that the majority of evaluation topics are selected based on agency or domestic priorities identified through primarily internal planning processes. Planning processes focus on building ownership input from agency staff and management. Consultation of other donors' plans tends to be informal and on an *ad hoc* basis. Contact with other agencies is undertaken when actively looking for specific opportunities to carry out joint evaluation (for example, when the agency already has a topic in mind), with the majority of the planning and topic selection process turning on the priorities identified within the development agency. Progress on conducting joint evaluations is covered in the following section.

Figure 2.13. Sharing evaluation plans

Does the unit consult other donor evaluation plans to look for possible areas of collaboration?

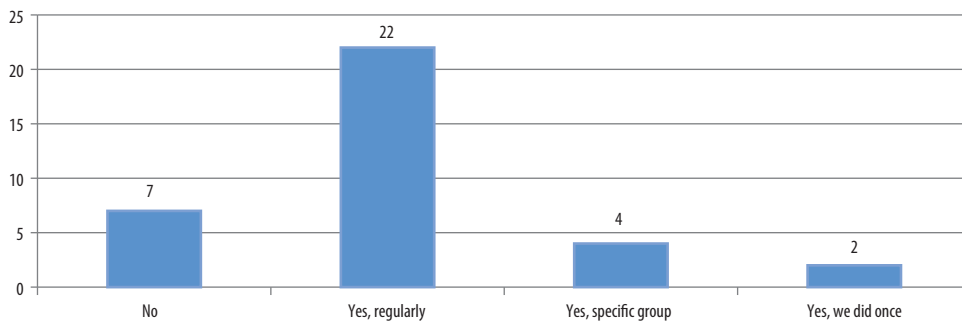
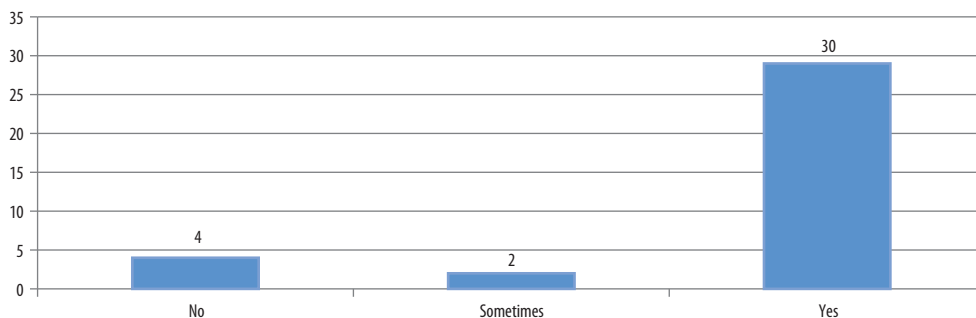


Figure 2.14. Consulting relevant evaluation reports

Does the unit regularly consult relevant evaluation reports by others before planning a new evaluation?



## Joint evaluation

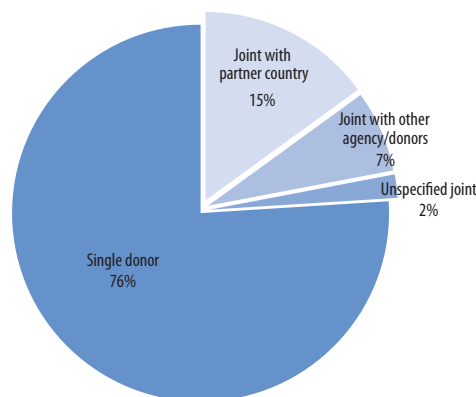
Wide variation exists between network members on joint evaluation, with some conducting nearly all of their evaluations jointly and others not currently engaged in joint work at all. Overall, single donor evaluations represent 76% of the 696 evaluations carried out by member departments on average per year, while joint work makes up 24% (or 159 evaluations per year). An analysis of the average proportion of joint work per member also finds about 96 joint evaluations involving partner countries are completed per year, representing approximately 15% of evaluations. Joint work with other donors/agencies represents around 7% or 49 evaluations (see Figure 2.15). If Japan's JICA (reporting a very large number of reports and an above average proportion of joint work) is excluded the figures shift slightly to: 73% single donor, 9% with other agencies, 4% with partners and 3% unspecified joint.

Some members did not make the distinction between types of joint evaluations (with other agencies or with partner countries) when reporting and these evaluations were included under "unspecified joint". There also may be some overlap in these two categories as joint evaluations often involve both multiple agencies/donors and partner countries; a separate category was not included for this "joint-joint" type of evaluation though it would be useful to make this distinction in future studies. The definition of "joint evaluation" used for self-reporting varied, with certain members using a quite broad definition and others restricting their reporting to "fully joint" processes with shared ownership. Other data sources were not available to triangulate these self-reported figures, though input from members indicates that there may be some over-reporting of joint evaluation.

While comparable data do not exist for earlier periods, this certainly represents a substantial increase from ten or fifteen years ago when the concept of joint evaluation was just emerging, judging by the number of joint evaluations discussed at earlier meetings of the network or published on member websites and DERE. Figure 2.17 gives an overview of the proportion of joint work for each individual member.

As illustrated in Figures 2.16.1 and 2.16.2, there is large variation between bilateral donors and the group of multilateral banks. Multilateral institutions conduct just 9% of all evaluations jointly (Figure 2.16.2), compared to 29% for bilateral donors (Figure 2.16.1). This disparity is explained in part by the large number of joint evaluations reported by Japan JICA, which skews the average for joint evaluations upwards for the group of bilateral donors.

Figure 2.15. Evaluations completed by all members of the DAC Network on Development Evaluation (average percentage of joint and single donor evaluations per year)



Please note: For Figure 2.15, a total of 696 evaluations were reported. Data not available for Italy and US DFA.

Figure 2.16. Joint evaluation: Comparison of bilateral donors and multilateral banks

Figure 2.16.1. Evaluations by bilateral members

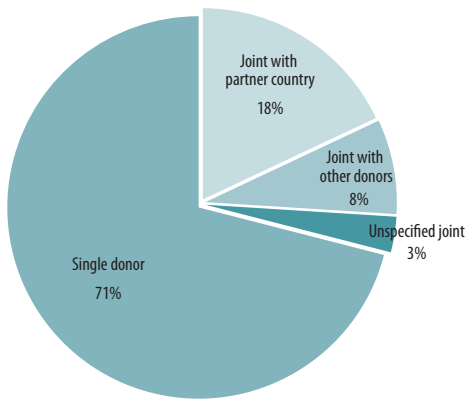
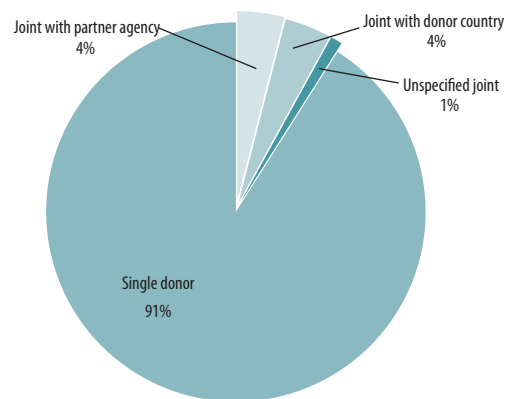


Figure 2.16.2. Evaluations by multilateral members



The study on joint evaluation undertaken by SADEV for the network in 2008 highlighted many of the challenges to practically implementing joint work in member agencies. Further reflection is needed to determine the ideal level of joint work. How well current levels of joint evaluation reflect overall “jointness” in development programmes and policies is also not clear.

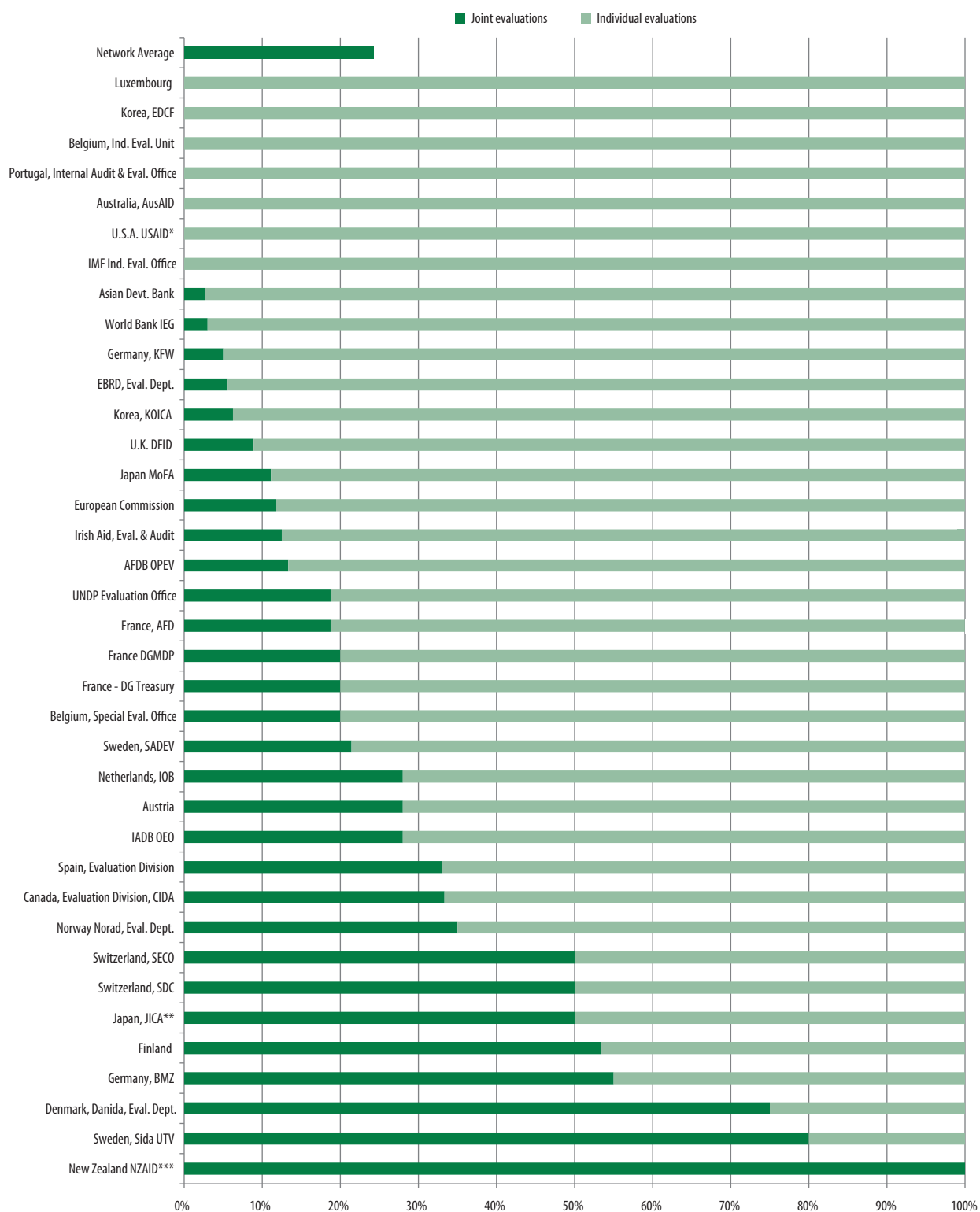
In short, progress is being made towards more joint work, primarily through collaboration among smaller subgroups within the network. While evaluators play a role, further progress may take time until development modalities themselves become more joint, making joint evaluation a logical consequence.

### Partner country involvement in evaluation

A key objective of the network’s efforts on joint evaluation is to improve mutual accountability and ownership of evaluation processes, to make development evaluation more responsive to partner country needs and interests, and to strengthen the evaluation capacities of developing country partners. The role of partner countries has been given importance in policy statements and network documents since the 1970s and for nearly twenty years DAC members have committed to the principle of stakeholder involvement in development evaluation (DAC Principles for Evaluating Development Co-operation, 1991). Yet progress in this area is mixed.

The section above outlined the overall involvement of or both other agencies and partner countries in joint evaluations, as reported by members. To understand these figures better, involvement of partner countries at specific key stages of the evaluation process was covered in detail in the member questionnaire. Members reported the frequency (never, sometimes, regularly/for certain types of evaluations, or always) of “partner country stakeholder involvement” at different points of the evaluation. The goal was to clarify the actual level of engagement, and distinguish between a “more joint” partnership-based approach and the simple consultation of partner stakeholders (for example, during data collection) – both of which involve partners but which connote very different approaches in terms of ownership and mutual accountability. The term “involvement”

Figure 2.17. Proportion of joint evaluations



Please note: Figure 2.17 reflects reported average proportion of joint evaluations per year. Data not available for Italy and US DFA. \*USAID: projected number of evaluations for 2010. \*\* JICA: Reported evaluations include large number of programme-level evaluations, many of which involve partner countries. \*\*\* Applies only to the two to three strategic evaluations completed by NZAID each year. The majority of programme-led evaluations undertaken by NZAID are not joint (see New Zealand profile).

was not explicitly defined in the questionnaire to allow for the fact that the appropriate type of participation varies from evaluation to evaluation. This, however, led to some discrepancy in reporting. Input from individual members seems to indicate that actual levels of engagement with important stakeholders may be lower than reported (as noted above on joint evaluation).

Partner country stakeholders, are defined as, “agencies, organisations, groups or individuals who have a direct or indirect interest in the development intervention or its evaluation” (OECD DAC, 2002). While consultants contracted to evaluate development programmes are not generally considered stakeholders in this sense, the involvement of consultants from partner countries was included in this question because it is one of the primary ways members say they incorporate country perspectives into the evaluations they commission. Hiring consultants from partner countries, however, does not in any way preclude the need to engage other stakeholders as they play distinct roles in the evaluation process, such as in programming evaluations, defining the scope of and approach to evaluations, as well as making use of the evaluation’s findings, conclusions and recommendations. The findings are summarised in Figure 2.19.

The results show that partner country stakeholders are most frequently involved during data collection and field visits. Individuals from partner countries (or institutions based in developing countries) are often hired to conduct evaluations. About half of respondents reported the partners are regularly involved in designing evaluations (including drafting terms of reference). Partner country stakeholders are consulted or involved at the planning stage by only a very small minority of members, with over seventy-percent involving partners only occasionally (26%) or not at all (53%). Eight respondents (20%) reported that this was done regularly or for certain types of evaluations (frequently for country level evaluations). Only one member reported always involving partners in some way in deciding its evaluation plans, though how exactly partners are involved was not specified. Participation as members of a reference or steering group varies widely, with twelve members reporting regular involvement, fourteen reporting occasional involvement and nine never involving partner stakeholders. It is also interesting to note the relative lack of involvement of partner country stakeholders at the distribution, follow-up action stages.

Most respondents said that evaluations respond somewhat to the needs of both donors and partners (72%), as shown in Figure 2.18. Several stated that this varies widely from evaluation to evaluation with some evaluations responding primarily to partner needs and others focused on contributing to donor accountability or learning. Others stated that their goal is to respond to domestic needs and that addressing partner concerns is not in their mandate.

**Figure 2.18. Relevance of evaluations for partner countries**  
How well do evaluations address issues of concern to both donors and partner countries?

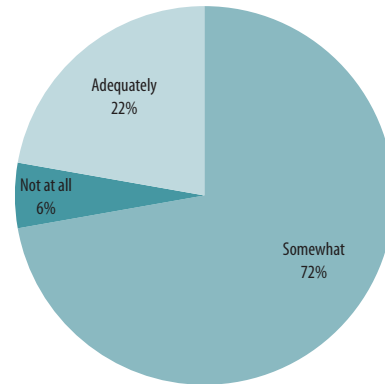




Figure 2.19. Partner involvement in evaluation  
Are partner country stakeholders involved...



## Evaluation capacity development in partner countries

In addition to helping improve the capacity of its own members, the DAC Network on Development Evaluation has the mandate to “promote and support evaluation capacity development in partner countries.” The network and its members work toward this goal by developing international evaluation standards and guidance, implementing targeted capacity building interventions, and through the involvement of partner country stakeholders in member evaluation work.

There is consensus among members that improving partner capacity is important – particularly in the context of commitments to improve mutual accountability and ownership, and efforts to do more joint work with country partners.

Despite this consensus, there are differences among members in terms of the roles evaluation departments play and the degree of resourcing and sophistication of their capacity work. A 2006 study by Japan for the network found that 22 members (of the 26 responding agencies) were currently conducting evaluation capacity development work.<sup>11</sup>

About half of member units do not have the mandate to deal with capacity development, either because it is covered by another department or because it is not a priority area for the development agency. Responses to the 2009 questionnaire show that the other 50% of members have “evaluation capacity development in partner countries” in their evaluation policies (see Figure 2.1). There is a range of coverage, with some policies simply mentioning “the importance of capacity development” and others providing a strong mandate to evaluation departments to actively contribute to capacity building – often with a dual mandate of supporting capacity both within their own agency and with outside development partners. For example, Danida’s mandate includes “contributing to the development of evaluation capacity in partner countries through bilateral and multi-lateral co-operation and contributing to the development of evaluation capacity in NGOs and the Danish resource base.”

Japan provides another interesting example. The Evaluation Department of JICA, in collaboration with partner country Ministries of Planning (or equivalent organisation), jointly plans and supervises the ECD process. JICA agreed to support ECD in Vietnam, Philippines, and Indonesia by signing memoranda of understanding for co-operation in evaluation. Through these supports, JICA aims to help partner countries to establish management methods of a project cycle in which the lessons learned and recommendations from the evaluations would be utilised in future development projects. Furthermore, Evaluation Department conducts annual ODA Loan evaluation seminars since 2001. The seminar targets government officers (in charge of development projects in planning agencies in partner countries), and provides capacity development training in evaluation system and techniques.

The questionnaire asked about capacity development strategies, finding that 59% of member departments had a strategy while 41% did not – either because they do not undertake capacity work or because they are involved in ad hoc capacity development work. Compared to the information cited above on inclusion of capacity development in evaluation policies, this finding would seem to indicate that there are at least two or three members that do not have a specific mandate to do capacity work, but are nonetheless involved in some way. Furthermore, comments from individual respondents highlighted that most capacity development work involves providing funding for a few evaluation training programmes. There seems to be relatively little strategic engagement on capacity

development, even among those members that have a mandate to do so. Further guidance on how donor evaluation units can best support partner capacity may be needed.

The level of implementation varies widely, with some members holding individual trainings for development staff in one or two countries, while others actively involve partners in joint work as part of an overarching capacity development strategy (see section on Partner Country Involvement). Comments show that many members support international capacity programmes, such as the International Program for Development Evaluation Training (IPDET), the African Evaluation Association (Afrea), or other regional/national associations.

## Conclusions

This chapter has described the current development assistance evaluation policies of the members of the DAC Network on Development Evaluation based on a member questionnaire conducted in 2009 and a literature review, including recent DAC Peer Reviews. In addition to providing an overview of the general state of affairs in development evaluation, the section has highlighted some emerging trends, notably improvements in independence of the evaluation function, the diversification of actors involved in evaluation and increased co-ordination between evaluation departments.

Network members together produce over six hundred evaluations per year, with the average evaluation unit completing 19 evaluations each year. The network-wide average budget for central/main evaluation units is USD 4.7 million. For DAC countries, central aid evaluation units have an average budget of USD 3.3 million, representing the equivalent of 0.1% of the development co-operation budget the unit is charged with evaluating. For the multilateral institutions, the average evaluation budget is USD 10 million, representing about 1.4% of the overall administrative budgets for these institutions. Challenges remain in obtaining comparable budget figures.

There is a shared concern regarding resources, particularly human resources, and it is widely acknowledged that joint work and impact evaluation in particular suffer from resource constraints in evaluation departments. However, resources are clearly not the only barrier to joint work, as some of the most well-financed units participate least in joint work and some very small units (both in terms of staff and budget) are deeply engaged in collaborative evaluations.

Staff levels and technical evaluation skills (of both staff and consultants) are areas of concern for just over half of reporting members. Slightly more than half of reporting members consider skills for evaluating cross-cutting themes as less than fully adequate. Network-wide, gender balance of evaluation staff is near parity (45% male and 55% female). Progress has also been made on mainstreaming gender equality into evaluation. Most evaluation units are now mandated to examine issues of gender equality as part of their regular evaluation work (often as a topic for thematic studies, an additional evaluation criterion or a special theme in all evaluation reports). However, women are under-represented in management posts.