



ROOM DOCUMENT 14

A Comparative Study of Evaluation Policies and Practices in Development Agencies

The attached note contains the executive summary of a Comparative Study of Evaluation Policies and Strategies in Development Agencies. The note is submitted by l'Agence française de développement for information at the 7th meeting of the DAC Network on Development Evaluation, 20 – 21 February 2008.
The full report will also be made available at the meeting.

A Comparative Study of Evaluation Policies and Practices in Development Agencies

Final Report to the French Development Agency

Marta Foresti

with Christina Archer, Tammie O'Neil and Richard Longhurst

October 2007

Overseas Development Institute
111 Westminster Bridge Road
London SE1 7JD
UK

Tel: +44 (0)20 7922 0300 Fax: +44 (0)20 7922 0399
www.odis.org.uk

Executive Summary

This report aims to understand experiences and challenges in implementing development evaluation reform and to assess a variety of solutions, looking at different development institutions. It hopes to lead to improved knowledge of evaluation systems in different development agencies and to provide some pointers to support AFD and other development agencies in their evaluation reform processes.

An overview of the evaluation units in the institutions under study looks at: accountability; information and knowledge; partnerships; capitalisation of experience; and feedback. The report compares policies and practices across institutions, looking at different choices and their consequences, as well as strengths and weaknesses. The analysis also identifies the main internal and/or external factors involved for each institutions. The report reviews management arrangements and structural position of evaluation units; their main features, processes and tools; and practices involved in commissioning, managing and supporting evaluation processes. It does not aim to identify good practice or to set standards for the AFD to follow.

The report is based on full case studies (including primary data collection) of five evaluation units, at: DFID; Sida; the World Bank; KfW; and the AfDB. Staff for interviews mainly came from the evaluation unit. It also covers four desk-based studies (relying mainly on documentary analysis), at: the EU; the IMF; Danida; and Oxfam. In addition, a small number of key informant interviews were carried out with leading development evaluation experts.

Evaluation policy, strategy and mandate

Evaluation policies, mandates or similar organisational documents refer to the overall purposes and objectives of the evaluation function for the particular agency. Most agencies are either reviewing these or waiting for them to be approved. If there is no evaluation policy or mandate, evaluation units refer to other organisational or procedural documents. Most such documents refer to both accountability and learning as the main objectives of evaluations, although there are some differences regarding to what degree each should be fulfilled. Although accountability and learning should be mutually reinforcing, some degree of confusion persists among staff, related to the (perceived) tension and trade offs between the two.

Agencies without an agreed policy sometimes experience confusion regarding the evaluation unit's mandate and/or roles and responsibilities.

Independence, integration and isolation of evaluation

With the exception of the multilateral banks, units have all undergone significant organisational changes in recent years. In most agencies, evaluation units are not subordinate to the management, operational or policy departments of the agency; the head does not report to management. Some report to a governing body; some are headed by staff recruited from outside and who cannot be employed by the agency at the end of the mandate; and some sit within the line management structure of the agency. Sida has set up an external and separate organisation (SADEV) to look at the totality of Swedish aid. Sitting outside management and policy structures allows for greater independence, one of the cornerstones of the evaluation function in aid and development agencies, but in some agencies' experience risks isolation from operational units.

Behavioural independence is more prominent in some agencies than in others. In a number of agencies some degree of tension exists between the independence of evaluation and its integration with other functions. Meanwhile, evaluation units normally rely on consultants to provide an external and objective assessment in their evaluations; those who rely on regular commissions can be under considerable pressure not to be too critical. Concerns were expressed that independence could jeopardise integration and 'voice' within the agency.

In the lending agencies, the unit budget is approved by the Board. However, in most bilateral agencies, the unit 'bids' for its annual budget, as with other departments. Increasingly, units are submitting workplans.

Types of evaluation and other activities

Overall, most agencies are focusing increasingly on policy, strategic and country evaluations. Some are still involved in project evaluation, though to a lesser degree than in the past. This can imply that evaluation units now have fewer opportunities to engage directly with field offices and operational departments.

Most evaluation units carry out on average between five and ten country programme evaluations every year; these aim to review and guide overall approach and strategy. Timing of these is not always consistent with country strategy and planning processes, however. The shift towards strategic, policy and thematic evaluations can present some challenges, particularly when findings are perceived as 'political' or it is not clear who should lead. However, it has generally been accepted that this shift is necessary in order to influence policy development and organisational choices. For this to happen, decisions on which policies or themes to evaluate should be taken in consultation with relevant staff in operational and policy departments.

Joint evaluations have also become more frequent. In line with the principle of harmonisation enshrined in the Paris Declaration on Aid Effectiveness, the increasing emphasis on joint evaluation is seen as inevitable and, in principle, a much needed effort on the part of the donors. Actual experiences have been mixed.

The main focus for all agencies is on 'independent evaluations'; internal reviews, self-evaluations, monitoring systems and, increasingly, project evaluation and completion reports are increasingly the responsibility of management, operational departments and country programmes. Quality of the latter is often a matter for evaluation units to deal with, sometimes through quality control systems and ratings. Most agencies rely on the OECD/DAC Evaluation Principles as a foundation for quality rating systems. Overall, there is more focus now on rating systems, management follow-up, quality control and knowledge management, to ensure that evaluations contribute to a 'learning culture' within the organisation and to some concrete proposals to influence practices and behaviours. There is more work to be done here, though.

In some agencies, the unit is responsible for producing some form of global report or synthesis. In addition, many agencies have mechanisms and processes to monitor management follow-up and response; these are not always formal or structured.

Evaluation methodologies

Agencies tend not to adopt a uniform methodology, although almost all use the DAC criteria as a common basis. Efforts are being made to ensure more unified and consistent approaches, to make evaluations more comparable and transparent.

Usually, individual consultants propose a methodology which is then discussed and agreed in consultation with the units. Quality can be an issue here too. Impact evaluations are an area of increasing interest. Some agencies are already carrying these out, or at least highlight impact as a key feature.

Roles, responsibilities and capacity of staff

Specific roles and responsibilities of staff vary considerably; they can include: staff involvement in evaluations; consultations with other departments on specific evaluations; tendering processes; and evaluation management, training and capacity building. Seldom independent evaluations are managed and carried out by unit staff. Increasingly, some form of direct staff involvement in evaluations is considered an opportunity to foster individual and institutional learning, although most agencies rely on external consultants to carry out all evaluations.

Evaluation units deal with the tendering process, including development of ToRs and recruitment and management of consultants. In most cases, ToRs are adapted to specific requirements, although there can be a lack of flexibility on this written into tendering procedures. Level of consultation with operational and policy departments varies across agencies although consultations are more intense during implementation, reporting and dissemination. Difficulties may arise in this process because units do not have a high level of access to country programmes.

Several units look after evaluation capacity building, both within and outside the agency, of in-country researchers/consultants, partner governments or internal staff.

Internal staff capacity is considered critical: the number of staff in evaluation unit is often inadequate, leaving staff members with little time to engage in the content of studies while overburdened with administration. In addition, staff do not always have previous evaluation experience. Attracting high-calibre staff and limiting staff turnover is a challenge, although many agencies have tried to tackle this and raise staff profile and credibility, and there is recognition of the need to bring in new skills, such as in learning and knowledge management.

Communication, dissemination and use

This is an area of growing importance for all agencies, in part motivated by a general awareness of current limitations. Only a few agencies have an explicit disclosure or communication policy/procedures, although almost all agencies publish evaluation reports. In some cases, management response is also published. Products range from reports, to synthesis papers, policy briefs, newsletters and e-bulletins, as well as annual reports. Other initiatives include dissemination seminars, both internal and external, special events, conferences, or press releases/conferences.

It is important to recognise that the pressure to communicate effectively with very different audiences can result in dilution of the more challenging findings or oversimplification of the more technical or rigorous aspects of evaluation. This could undermine the impartiality or objectivity of the evaluation process.

Internal feedback and learning often remain very weak. Some efforts are being made to improve management response and follow-up, as well as information management systems and databases. However, this does not reach the wider internal audience (made up of those who could still benefit from sharing lessons learnt). There are some systematic efforts to improve this situation, but these remain few.

Concluding remarks

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.

However, there is considerable diversification, in terms of both internal arrangements and different roles and responsibilities fulfilled. This is partly explained by the fact that units are increasingly expected to fulfil a variety of different roles and to engage in a wide range of activities. Differences lie in: roles and responsibilities beyond traditional evaluation; references audiences; types of evaluation; and tools and products.

Against this background, the evaluation function within aid and development agencies emerges as a function 'in search of identity'. In some cases, the search is quite advanced; in others, this is a work in progress. Such searches are exemplified in the frequent changes agencies have undergone in recent years.

Overall, there is an apparent disconnect between the rhetoric and the rhetoric on the strategic and growing importance of development evaluation, and evaluation practice in development agencies. More investment is needed, both financial and (more importantly) in terms of human resources and capacity. This disconnect could be further investigated and assessed through an evaluation initiative, aimed at assessing the 'goodness of fit' between the current arrangements and the objectives of evaluation units.

This 'institutional evaluation gap' calls for greater attention to institutional approaches to evaluation, arrangements and capacity: Investing in the quality of evaluation should be a priority for most agencies. In addition, greater attention needs to be paid to the nature, origin and potential of the demand side of the evaluation equation: in recent years, there has been an increased demand for transparency and accountability of aid and development interventions. This has the potential to inject some new energy into the practice and use of development evaluation which, in time, should be reflected in the performance of evaluation units in all development agencies.