



Department  
for International  
Development



# **Rapid Review of Embedding Evaluation in UK Department for International Development**

**Executive Summary**

**February 2014**

## **Purpose of the rapid review**

Since 2009/10, there has been a drive within the Department for International Development (DFID) to strengthen the evidence base upon which policy and programme decisions are made. Evaluation plays a central role in this and DFID has introduced a step change to embed evaluation more firmly within its programmes. The primary purpose of this rapid review is to inform DFID and the international development evaluation community of the progress made and the challenges and opportunities encountered in embedding evaluation across the organisation.

## **The embedding vision and the embedding process**

The direct impetus for embedding evaluation came with the change of UK government in 2010. The incoming Secretary of State (SoS) for International Development took steps to set up an independent body, the Independent Commission for Aid Impact (ICAI), with a strategic aim to “provide independent scrutiny of UK aid spending, to promote the delivery of value for money for British taxpayers and to maximise the impact of aid”. The role of the existing Evaluation Department (EvD) within DFID was redefined, as was the role of evaluation within the organisation.

A vision was set out for DFID to:

- become world class in using evidence to drive value for money and impact and influence other donors to do the same
- drive programme design through rigorous evaluation of what works, allowing DFID to test, innovate and scale up
- take measured risks using high quality evidence of impact on poverty
- help partners to generate and use evidence (DFID 2010b).

Embedding evaluation would be achieved by a systematic change process. Key elements of this were:

- more staff in operational and policy divisions able to undertake specialist work on commissioning and using evaluations
- much clearer standards on evidence when new programmes were commissioned
- a rapid increase in work on rigorous Impact Evaluations (IEs) within programmes
- development of a new professional cadre of accredited evaluation specialists from across the organisation.

In practice, the embedding evaluation approach has fundamentally changed the way evaluation is managed and used across the organisation. Evaluation is embedded in Business Cases (BCs) for all new initiatives, and programme teams and embedded Evaluation Advisers (EAs) have become the front line for commissioning evaluations. EvD’s location, role and function shifted to guide and support evaluation activity across DFID.

## **Capacity to evaluate**

There has been a strong drive to recruit, accredit and train staff in evaluation in DFID since 2011. There have been 25 advisers working in a solely or shared evaluation role, a further 12 advisers in roles with an evaluation component, 150 staff accredited in evaluation and 700 people receiving basic training. This is a major achievement in a relatively short period of time, and illustrates the commitment from senior management and staff across the organisation to build evaluation into DFID's policy and practice.

While the scaling up of capacity has been rapid, the depth of this capacity is less than required. The number of embedded advisory posts created is significantly fewer than envisaged at the outset, with eight of 25 advisers working 50% or less on evaluation. The distribution of these posts has also been uneven, with some Country Offices (COs) and operational departments having little or no evaluation advisory support. In several cases, EvD staff are effectively filling the evaluation roles of other departments. In some cases there has been a deliberate strategy to utilise support from other accredited staff members in a CO instead of having dedicated evaluation support. There are ten offices where one EA has to support between five and 20 evaluations using only 30-50% of their time.

The establishment of an Evaluation Cadre as a resource and community of practice for those working on or supporting evaluation has been viewed positively. The accreditation process has generated huge interest. However, concerns have been raised as to whether the cadre can support the professionalisation of such a large number to a level where they can competently commission and manage evaluations, or whether, at this stage in the embedding process, the focus should be on a smaller core group. Of those accredited to date, 81% are at the foundation or competent level, neither of which qualifies them to manage a substantial evaluation without specialist support. The link between accreditation to a cadre and recruitment into evaluation posts has not been consistently maintained, leading to concerns that some decentralised staff may not be receiving the appropriate information and support required to be effective in their roles.

DFID's focus on supporting external partners has been quite progressive, particularly in IEs. However, it has not sought to build capacity through the evaluations it is financing. Most contracting is with northern companies, which have few or no requirements to work with partners in the global south. This is a lost opportunity.

## **Effect on quantity, coverage and type**

The embedding evaluation approach has contributed to a significant, but uneven, increase in the quantity of evaluations commissioned by DFID. These have increased from around 12 per year, prior to 2011, to an estimated 40 completed evaluations in 2013/14. Health, poverty and education areas are well covered by evaluation, while wealth creation and governance are less so. This is a concern due to the increased focus which DFID has on these latter areas and the current shortage of evidence to support policy and programme decisions.

In terms of planning, the Africa divisions intend to evaluate a larger proportion of their programmes, both by number and value, than other divisions. The Western Asia Department (WAD), Middle East and North Africa Department (MENAD) and to a lesser degree Asia, Caribbean & Overseas Territories (AsCOT) do not appear to have embraced the need for evaluation to the same extent as the Africa and policy divisions. Given DFID's increasing focus on fragile and conflict affected states there is a need to review the extent to which programmes in these areas are currently being evaluated.

The focus of evaluation has changed to become almost exclusively programme oriented. There are very few thematic or country level evaluations planned whereas previously these types of evaluations accounted for the majority of DFID's evaluation portfolio. This presents a challenge to DFID as it seeks to synthesise the learning from individual projects and programmes into broader lessons for policy and programme planning and design.

### **Effect on quality**

Evidence from independent Quality Assurance (QA) between 2012 and 2013 indicates variation in the level of quality by region and stage in the evaluation cycle. Pass rates (green and amber ratings on evaluation QA) indicate that Asia has the lowest rate at 54%, Africa 75%, MENAD 78%, and global programmes 77%. This variation points to potential gaps in technical support and oversight, which are also highlighted by differences in the coverage of country and regional EAs. Quality issues cut across most parts of the evaluation cycle, although recently the lack of attention paid to managing evaluation implementation to ensure quality has been highlighted as arguably the most critical challenge. Clarifications and improvements in guidance are also required and are being addressed.

### **Effect on demand and use**

The embedding process has placed evaluation firmly within the programme cycle and thus increased the actual and potential demand. This has taken place in a context of wider organisational focus on using evidence, demonstrating results and increasing accountability for resources used. However, the institutionalisation process does appear to be drawing attention away from external audiences – stakeholders who can influence design and quality, and who, ultimately, use the evaluation findings.

There is a greater appreciation of evaluation in DFID and evaluation findings are already being used to some degree in decision making. However, the quality of evaluation management responses is variable and a clearer process is needed to ensure that these are timely and used consistently. Furthermore, there is not yet an agreed approach to disseminating evaluation findings and promoting evaluation use.

There is scope for DFID to adopt a much more systematic and creative approach to promoting evaluation use in the context of the wider evidence agenda. There are institutional routes for this, for example through the work of the Research and

Evidence Division (RED) on evidence uptake, resource allocation rounds, policy processes and advisory cadres. Although some of these channels are being utilised, they have not yet been systematically identified and pursued.

In line with DFID's transparency agenda, all evaluations are now published, although visibility is limited as there is no platform on the DFID website where they can be easily accessed. There are some good examples of sharing and uptake of evaluation findings beyond immediate stakeholders, but DFID is not yet positioned to make best use of the emerging body of evaluation evidence either internally or externally.

### **Effect on value for money**

The embedding evaluation approach has been accompanied by a significant increase in the number of evaluations which has, in turn, led to an increase in the total amount spent on evaluation. However, the average total cost per evaluation has changed little since 2010.

Externally procured evaluation costs appear to be in line with those of other donors. However, forecasts of future spending on evaluation indicate a likely increase in the median amount that DFID pays directly for evaluations. For non-impact evaluations the median budget is £200,000 and for IEs the median budget is £500,000. This represents a significant under-estimation of evaluation costs.

Evaluation accounts for a median of 1.9% of programme value, which is in line with expectations. The amount DFID spends on IEs is higher at 2.6% of programme value but this is consistent with the figures of other donors such as the Millennium Challenge Corporation and the World Bank.

### **Fitness for purpose**

The extent to which the structures and processes are appropriate to achieve the aims of embedding evaluation within the organisation was reviewed at two levels: i) COs and other spending units, and ii) EvD.

In COs and UK based operational and policy departments, the decision to evaluate is being made predominantly within programme teams during the preparation of BCs. This has led to an imbalance across portfolios and a lack of a broader strategic focus, though the focus of evaluation within teams is of benefit to programme performance and future design.

The constraints to fitness for purpose are primarily capacity and the need to ensure the relevance and quality of evaluations. Evaluation staffing is uneven and where EAs exist they are almost always split posts with results or statistics responsibilities alongside evaluation. As the organisation focuses increasingly on improving programme management, following the recent 'End-to-End' Review of Programme Management, it is likely that individuals in positions involving evaluation and results will face increased challenges in adequately carrying out their evaluation work. EAs will have to pay substantial attention to the utility of evaluations commissioned. In

addition, a stronger recognition of the relationship between evaluation and results monitoring is required by DFID centrally. The mutuality of the dual roles at country level will need to be better understood and supported, particularly with the increased focus on programme management. This, in turn, might necessitate more central support to strategic, complex and impact evaluations, including jointly managed initiatives, to better ensure quality and utility at both the country and corporate levels.

Since being moved to RED, EvD has served two primary functions: establishing and maintaining the policy framework and systems for supporting decentralised evaluation, and providing a one-on-one technical support service to those designing and managing complex, impact and strategic evaluations. Both of these functions are appropriate in the context of a strongly decentralised evaluation function. To better support decentralised evaluation attention will need to be paid to EvD's responsiveness to demands for support, the quality of this support, and ultimately the quality of evaluation products. The Evidence Survey conducted in 2013 found that despite a positive view of evaluation, 45% of those interviewed felt that the evaluation advice and support on offer was only partially sufficient or insufficient for their needs. While the primary responsibility rests with decentralised EAs, the support from EvD is still too thinly spread. EvD needs to reconsider priorities and find ways to ensure the decentralised evaluation system is effective and efficient in delivery.

### **Implications and options**

Although it was not in the scope of the Review to make formal recommendations, a number of suggestions arising from the findings are proposed for further consideration.

#### ***Evaluation coverage***

The findings show that gaps have developed in terms of coverage, geographically, thematically and in areas such as aid modalities. These gaps are an almost inevitable result of the current situation, which includes decentralised decision making without strong coordination mechanisms. This is an issue that needs to be addressed, but without undermining decentralised ownership. It could possibly be tackled through an overall DFID evaluation strategy and through much stronger involvement of decentralised senior management in decisions to evaluate.

#### ***Support to operational units***

DFID has made efforts to build the capacity of decentralised units to commission, manage and use evaluations. However, the QA of evaluation Terms of Reference (ToR), inception reports and final evaluation reports has identified gaps in current practice. In particular, there is a need to focus efforts on ensuring evaluations are improved by:

- understanding why and when to commission evaluation
- enhancing the contexts of evaluations and engaging stakeholders in an appropriate and timely manner

- selecting and implementing appropriate evaluation approaches while ensuring reliability of data and validity of analysis
- reporting and presenting information in a useful and timely manner.

These issues point to a critical need for DFID to ensure that operational units have the capacity to manage evaluations well, including engaging stakeholders during evaluations and undertaking QA of evaluation activities.

### ***Making better use of evaluation***

There is a sense that enthusiasm for evaluation has waned since the embedding initiative was started in 2010. This is in part due to staff changes, particularly at senior levels, and also because of a reduced visibility of evaluation results. To justify the amount of resources currently being spent on evaluation, it will be important for DFID's corporate centre, whether EvD or other research, evidence or policy teams, to communicate evaluation findings more broadly and at higher levels. Interest can only be maintained and increased if the use of evaluation is demonstrated.

### ***Strengthening evaluation management***

There has been considerable enthusiasm shown by programme managers for conducting IEs, which now comprise 28% of planned evaluations. EAs have expressed some concern that there may be a lack of understanding of what is required to conduct a good IE, in terms of time, resources and the technical capacity to manage such an evaluation properly. It will be important for COs to develop a more intensive engagement with EvD and research teams when they undertake these types of evaluation, and to ensure that central teams have the resources to provide this more intensive support.

For this to happen, EvD will need to make certain decisions. The current proportion of total DFID evaluations being supported is too great to provide the level of support required. To address this capacity challenge the commissioning of evaluations across the organisation needs to be managed more strategically, staffing in EvD needs to be increased, support responsibilities need to be shared with other parts of RED, or a much greater investment in staffing needs to be made at the decentralised level.

### ***Enhancing evaluation governance and strategy***

There is a need to strengthen the governance arrangements that determine DFID's central oversight of evaluation across the organisation. The current Evaluation and Evidence Strategy Group (EESG) has fulfilled its initial purpose and now requires a new mandate, more senior membership and a clearer reporting line to the Executive Management Committee (EMC) and/or the Investment Committee (IC). The mandate needs to reflect the main purposes of evaluation in DFID – accountability for results, organisational learning and global evidence building – through generating new knowledge on development effectiveness. To do so, the mandate would include

commissioning the development of a DFID evaluation strategy. This would specify priority sectors, geographies and thematic evaluation areas. The new mandate would also give the group responsibility for advising Directors General (DG) and directors of any changes required in their divisions for effective implementation of the strategy. The membership would include directors of geographic divisions and policy divisions to ensure organisational commitment and strategy implementation.

The development of DFID's evaluation strategy will address two key findings in this report. Firstly, regional and thematic coverage is variable and there is minimal involvement of relevant policy or evaluation teams in decisions to evaluate. Secondly, there is need to reinvigorate decentralised engagement with evaluation. A more strategic approach to evaluation has the potential to ensure optimal use of evaluation resources. It can thereby fulfil the overall purpose and function of evaluation as an essential tool for increasing the effectiveness and impact of development assistance.