

DFID and DCED organized a small workshop in London, with support from ITAD and others, to look at two topics:

- the interface between monitoring and evaluation in private sector development; and
- current state of the art in use of theory-based evaluation techniques in PSD.

Discussion focused on bringing together qualitative and quantitative techniques under a theory-based umbrella; it also considered the implications of the new trend in appointing evaluation units attached to individual programmes throughout implementation.

Participants included DFID staff and consultants supplying monitoring and/or evaluation services to DFID; they heard presentations from a number of programmes during the day, and discussed issues arising. These short notes summarise the discussion, without attributing comments to individuals or programmes; they could form the basis for a second, similar event in 6 months' time.

One over-arching conclusion was that it is very important to be clear on the objectives of the evaluation, because this drives the design and methods used:

- Internal or external learning or accountability; internal or external validity
- Who will use the findings, so what level is the evaluation pitched at? Intervention / programme / market / sector?

The objectives should also be realistic.

The monitoring / evaluation interface

"Optimum independence is not maximum independence. Accurate and fair evaluations combine intellectual detachment with empathy and understanding"¹. How independent can and should evaluation units attached to programmes be?

Being in at the start means that the evaluation units do not have to establish their credibility in mid-stream, catching up with the implementation team. Early dialogue, and good chemistry, between implementers and evaluators are important; the relationship is one of 'critical friend'. This relationship could be damaged if evaluators are given responsibility for conducting annual reviews, and/or for appointing DCED auditors. Accountability hasn't been successful in driving better measurement. What creates incentives for better measurement is the drive to be effective.

The evaluation unit in one such programme made institutional learning more important than accountability. An early Theory of Change workshop was found to be very useful, as was signing a Principles of Work agreement, setting out respective responsibilities (see Annex). It also helps if the evaluators have input to the design of the log-frame and if the evaluation function is in place till the end of implementation. The following questions emerged (among others):

- If a programme is adapting its strategy as it goes along, when should it do a baseline? As soon as intervention starts to 'take off'?
- In one challenge fund, the evaluation unit had to sign confidentiality statements. Others feel that data gathered in evaluations should be made public (and there are a number of relevant [principles](#), legislation and other [policies](#) requiring *inter alia* that all evaluation reports are published). Is more guidance needed? Or should evaluation units just negotiate with

¹ www.wipo.int/export/sites/www/about-wipo/en/oversight/iaod/evaluation/pdf/iadci_dfid.pdf

implementers on a case-by-case basis?

- Should a DCED Audit be governed differently to an evaluation? In particular, how does it work if the Evaluation Unit commissions the DCED Audit? Would a DCED audit accept elements of measurement conducted by the evaluation unit instead of the programme, and audit it accordingly? ²

Several questions emerged around the theme of who should gather which data: evaluator or implementer? There is a spectrum of possibility, from evaluators collecting all their own data, through to validation of monitoring data. Some implementers are concerned that evaluators could disrupt their relationships with companies, asking inappropriate questions, raising expectations, being too numerous and burdensome, etc. In those cases, evaluators could focus on the counter-factual? Are evaluators better placed and skilled to gather and analyse qualitative data? Evaluators and implementers could sort this out through negotiation, but DFID need to decide in advance, for tendering and contracting purposes. There's a spectrum of involvement from the evaluator, from full independent data collection, to quality assurance of data, to capacity building/support of programmes in collecting data

Theory-based evaluations of PSD

Evaluations should be theory-driven rather than methods-driven; evaluation data and findings could be put to greater use in strengthening market systems. One evaluation unit expects to focus on keystone nodes - points in the results chain that are really crucial to success in changing market systems, and which occur across several interventions or sectors. In another programme, evaluators are using realist evaluation - focusing in some depth on the theory of change to understand it fully and unpack the assumptions on which it is based. The following questions emerged from the discussions:

- Programmes should produce good theories of change at the end of implementation, when it is clear how things worked. That might inform consideration of which assumptions regularly hold true across multiple programmes. But what makes a theory of change good? Articulating the complexity in the right level of detail is as much an art as a science, especially in PSD.
- How rigorous are theory-based approaches? What would it take to prove them? Could they be aggregated? [5 principles for degrees of evidence](#) have been proposed in the US. The risk is that these approaches aren't sufficiently well understood for people to trust the conclusions. And how do we keep the jargon, especially around names of new methods, to the minimum.
- What is special about Theory Based Evaluation in PSD? Is it different for catalyzing capital flows, returnable capital, trade etc.?
- Staff turnover is a problem at all levels; what institutional arrangements minimize the impact of this? Should we understand better the cost of evaluations - depending for example on the information being sought, the complexity and geographical spread of the programme, orderly documentation (or not) and other parameters that will affect the actual cost in practice.

Next steps

A second meeting, in 6 months' time, could be focused around one or two of these questions; it could also feature content that could not be included in the first meeting. Meanwhile, the group can be a good sounding board for draft concept notes and other material on the topic.

² Questions probably for DCED in particular

Annex: Example Principles governing the relationship between DFID, [Programme Implementer] and [Evaluation Implementer] for the Evaluation Title

Date

1. Purpose of the Document

1.1. This document outlines the principles which will guide the working relationship between the signatory parties. The document is drafted in the spirit of collegial professionalism and mutual respect for each of the respective parties' roles and responsibilities.

2. Purpose of the Collaborative Relationship

2.1. The signatory parties share a common objective of a successful [evaluation title].

3. Signatory Parties

3.1. The signatory parties to this document are

3.1.1. DFID – the funding entity

3.1.2. [Programme name] – an independent entity made up of consortium partners which are contracted by DFID to implement the programme

3.1.3. [Evaluation Contractor name] – a company contracted by DFID to conduct an independent evaluation of the programme.

4. Guiding Principles and Shared Commitment

4.1. The three signatory parties agree to:

4.1.1. Independence: Recognise the independence of the programme and the independence of the evaluation. The independent evaluation is led by people with no other involvement in the project. This is important to prevent bias, either real or perceived.

4.1.2. Collaboration: We will work in close collaboration, while maintaining the independence of the evaluation.

4.1.3. Communication: Communicate proactively to ensure effective joint planning and on-going survey operations.

4.1.4. Cooperation: Deal with any methodological and operational issues as they arise, in a timely and cooperative manner, recognising the efficiencies of dealing with matters at the lowest appropriate level.

4.1.5. Confidentiality: We recognise that in the course of the project the three signatories may share material and information that might be regarded as sensitive or confidential. If we are in any doubt about the status of such information we will seek clarity from our fellow signatory before making use of this information.

4.1.6. Resolving Disputes: We will adhere to agreed procedures for resolving disputes, where an agreement cannot immediately be reached.

4.1.7. Transparency: Programme management and evaluation reports will need to be quality assured to ensure they meet the required standard, and then reports and all associated datasets will be made available to all signatories.

4.1.8. Methodology: The Evaluation must assess the impact of the programme, as outlined in the evaluation Terms of Reference, and also assess why and how things work, including investigating the theory and assumptions behind the interventions, and checking for unintended effects. To do this the Evaluation should adhere to the

internationally-agreed evaluation criteria³: relevance, effectiveness, efficiency, impact, and sustainability.

4.1.9. Data sharing: We recognise the interdependence of our roles and to this end agree to share data relevant to the programme, its successes and its weaknesses, openly with one another. We recognise that this shared information and data should be treated sensitively and with respect for our partners and their respective organisations.

4.1.10. Ethical Consideration: The Evaluation must adhere to the DFID Ethics Principles for Evaluation and Research.

5. Areas for collaboration between Evaluator and Programme Implementer

5.1. The programme Implementer and the Evaluator agree to work cooperatively. This will include but not be limited to coordination of timings for activities relevant to the Evaluation, coordination of efforts to consult stakeholders, communities and local government officials, and sharing of logistical and field support where feasible.

5.2. For quantitative data, the content of surveys and other data collection instruments will be developed cooperatively between the programme Implementer and Evaluator. The baseline and subsequent quantitative surveys will be developed by the programme Implementer in the first instance with collaborative input from the Evaluator.

5.3. For qualitative data both the programme Implementer and the Evaluator will share and request input from the other in terms of interview design, indicator descriptions, interview protocols, and other information which is relevant to the execution of this task.

5.4. The Evaluator should determine the design of the evaluation, in consultation with the programme Implementer.

5.5. In addition the programme Implementer and the Evaluator agree to:

5.5.1. utilise where appropriate programme's internal monitoring systems for the collection of data

5.5.2. respect the core evaluation questions agreed upon by DFID as part of the independent evaluation

5.5.3. respect the independence of the evaluation, which may in some cases, require the Evaluator to produce its own instruments or collect data independently.

5.5.4. respect the need for timely data collection and analysis to inform interventions and progress

5.5.5. respect the potential for research fatigue among communities and respondents

5.5.6. be cognisant of the implications of proposed changes on field and data management systems, including the mechanics of survey design, manuals and training materials, data entry templates, cleaning script and the code for generation of indicators and other relevant data points

6. Dispute Resolution

6.1. The principles outlined below are drafted in the interests of efficiently resolving disputes between the parties and to ensure that the programme and the evaluation are not impacted in a negative way. In keeping with the overall guiding Principles of Work as

³ OECD Development Assistance Committee (DAC)

outlined above, the parties agree that decision making should operate under the principles of group consensus ensuring that the principles of robust evaluation are maintained.

- 6.2. If there is disagreement between the Evaluator and the programme Implementer, the issue should be escalated from the team members to the team leaders, and programme or project directors as needed. These names and roles are reflected in Table 1 below. For example, if an issue cannot be resolved between A and L, it will be escalated to B and M, and should the issue still not be resolved then it will be escalated to C and N. DFID should be notified and updated if issues need to be escalated.
- 6.3. As a last resort, and only when all other avenues have been explored shall the disagreement be escalated to DFID to intervene and resolve the disagreement. DFID may seek external advice on the resolution of the disagreement and undertakes that the issue will be resolved in a timely manner, with a view to minimising the impact on the programme, its implementation and its evaluation.
- 6.4. DFID shall be kept informed by the programme or project manager of any disputes that are escalated.

Table 1: Names and Roles of individuals responsible for dispute resolution to be completed

Organisation	Responsibility	Supervisor	Next supervisor
DFID	A	B	C
Implementer	L	M	N
Evaluator	X	Y	Z

7. Data ownership

- 7.1. All new data collected for the project will be jointly owned by DFID, the programme Implementer and the Evaluator. Ultimate ownership of the data and all intellectual property related to the evaluation will be DFID's under the terms of contract.

8. Deliverables

- 8.1. DFID must be satisfied that all deliverables meet an acceptable quality and standard against the terms of reference for the Evaluation, before submitting for peer review.
- 8.2. After DFID's internal quality assurance process is complete, if the Evaluator publishes any information relating to the programme, full acknowledgement of the contributions of the programme team and others as appropriate will be made.

9. External and Public Communications and Publications

- 9.1. No formal publications, from the Evaluator, the programme Implementer or DFID, are to be made public before the internal DFID quality assurance process is complete.
- 9.2. DFID, the programme Implementer and the Evaluator must all be fully aware and notified early of any planned publication and associated communications for public consumption.

The parties, through their duly authorised representatives, agree to the above named principles governing the relationship between the Department for International Development (DFID), the programme Implementer and the independent Evaluator for the [Evaluation Title].