ALNAP 20th Biannual meeting – Rome – 6th Dec 2006

REPORT ON HALF-DAY WORKSHOP ON JOINT EVALUATIONS

1. INTRODUCTION

One of the stated aims of the Tsunami Evaluation Coalition (TEC) was 'to test the TEC approach as a possible model for future joint collaborative evaluation'. During the course of its work the TEC has facilitated three After Action Reviews and has recently conducted an email survey of its wider membership. Reviewing the experience (successes and learnings) of the TEC as a joint evaluation exercise has been the focus of each of these activities. In addition, ALNAP hosted a joint ALNAP/ ECB/ TEC/ DEC roundtable on joint evaluations in London in June 2006. The reports from all of these events, and including the email survey, can be found on the TEC's website: www.tsunami-evaluation.org

The purpose of this half-day workshop at the ALNAP Biannual meeting in Rome was to build upon the experience and learnings that have emerged and been captured to date, in order to set the agenda for joint evaluations of humanitarian assistance in the future. The workshop began with three presentations[1]:

(i) 'History of joint evaluations' since the mid 1990s by Niels Dabelstein of the Ministry of Foreign Affairs, Denmark
(ii) 'The ECB experience' of joint evaluations by Jock Baker of CARE US
(iii) 'TEC lessons learned' by Rachel Houghton of the TEC Secretariat

Discussion was then open to all workshop participants (approximately 35), and focused on the following six questions:

(1) What are the top 3 or 4 reasons for doing joint evaluations?
(2) In what context should joint evaluations be considered (ie, for what type of emergencies, sectoral/ thematic issues)?
(3) When should they be carried out? (eg, real-time, during/ after an emergency)?
(4) How is it possible to get all groups on board? (eg, NGOs, donors, recipient governments, field and headquarters)
(5) What is an appropriate management structure?
(6) What role should / could ALNAP play in taking forward joint evaluations within the sector?

(NB: There was little discussion on the first question as this was mainly answered in the three presentations.)

2. DEFINING AND CATEGORISING JOINT EVALUATIONS

The workshop began by making a distinction between joint evaluations which involve two or more agencies and 'system-wide' evaluations which involve a broader range of organisations from across the humanitarian sector. It was agreed that the discussion would focus on the latter. However, it was also noted that agencies typically involved in system-wide evaluations are those which are represented in ALNAP meetings. There are other agencies and actors within the humanitarian sector

---

[1] Some of the key points from each of these presentations are captured in the following text. See www.tsunami-evaluation.org for the powerpoint version of the first and third presentations.

that do not normally participate in system-wide evaluations (for example, local NGOs and host governments).

Niels Dabelstein provided a useful categorisation of joint evaluations in his presentation, as follows:

- *Classic joint evaluations:* Participation is open to all stakeholder agencies. All partners participate and contribute actively and on equal terms. Examples include: the Rwanda Evaluation, the CDF evaluation, the General Budget Support (GBS) Evaluation and the Tsunami evaluation.
- *Qualified joint evaluations:* Participation is open only to those who qualify, through membership of a certain grouping (eg, DAC, EU, Nordics) or through active participation in the activity (eg, jointly-implemented programmes) that is being evaluated.
- *Framework Evaluations:* Participating organisations agree on a common evaluation framework. Responsibility for implementation of individual evaluations is then devolved to different partners resulting in individual case evaluations and a synthesis report. An example is the joint IDP evaluation.

There was broad consensus within the group that system-wide evaluations are not a conspicuous feature of the humanitarian landscape, especially when compared to the development field. At the same time initiatives such as the TEC suggest that the number of system-wide evaluations undertaken is likely to increase in the future – hence the value of this workshop in doing some forward-thinking and setting the agenda.

3. *WHY DO JOINT EVALUATIONS, OR NOT?*

Views were expressed by the presenters and by the workshop participants both in favour of, and against, doing joint evaluations. These are captured below:

Reasons for doing joint evaluations:

- <u>Broader scope</u>: answers questions that cannot be addressed by one actor alone, such as coordination and coherence; also enables sensitive issues to be addressed.
- <u>Objectivity and legitimacy</u>: increased weight of the evaluation if it is undertaken with partners.
- <u>Advocacy tool:</u> opportunity to influence at the highest level, and can contribute to ongoing reform initiatives
- <u>Rigour</u>: joint evaluations generally demand a higher water mark of rigour than single agency evaluations.
- <u>Attribution</u>: it is usually easier to capture attribution in a joint evaluation.
- <u>Efficiency</u>: rationalisation, harmonisation and reduced transaction costs for all partners (except the lead agency).
- <u>Participation and alignment between agencies:</u> there is an opportunity for peer review and peer learning.
- <u>Evaluation capacity</u>: it is a way of developing evaluation capacity within the sector.
- <u>Beneficiary voice</u>: opportunities for doing large beneficiary surveys are usually greater in joint evaluations than single agency ones.
- <u>Social capital:</u> builds social capital amongst the agencies involved.

Reasons against doing joint evaluations:

- <u>Complexity</u>: can hamper joint evaluations of humanitarian assistance.
- <u>Lack of agreed common standards</u>: this can get in the way of joint evaluation within the humanitarian sector.
- <u>Time</u>: it takes much longer to plan and execute joint vs single agency evaluations.

- Management: it requires a complicated management structure to work.
- Transaction costs: these are usually higher for participating agencies (especially for the lead agency) compared with single agency evaluations, although this needs to be balanced against the value that participating agencies gain from the exercise.
- Focus of recommendations: need to guard against inadequately targeted recommendations which can reduce their impact.
- Detail required for a single agency: a joint evaluation may not provide the detail required to fulfil accountability requirements on the part of a single agency.

The overall consensus, however, seemed to be in favour of joint evaluations, albeit with some important caveats – see below.

Workshop participants were reminded that it is the *process* of the joint evaluation that is often as important as the final reports, not least in terms of the learning that takes place during the evaluation. Participants also recognised that joint evaluations will not necessarily replace single agency evaluations required for accountability purposes. The key is to ensure that joint evaluations do not duplicate what is being done by individual agencies, but instead build on that effort and stay focused at the system-wide level. There is evidence that some agencies did forego single agency evaluations when the TEC was launched, although many others continued to mount single agency efforts.

4. WHEN AND WHERE TO DO JOINT EVALUATIONS?

Current status

Niels Dabelstein reported that approximately 50 percent of evaluations carried out by the Nordic and Irish governments are now joint evaluations, but also noted that many more of these are in the development rather than humanitarian sector. Jock Baker reported that CARE, Oxfam, IRC, CRS, Mercy Corps and Save the Children are all participating in the ECB, which is exploring and experimenting with joint evaluations among NGOs. Rachel Houghton described the recent TEC experience as a 'hybrid joint evaluation' involving more than 40 agencies. Interestingly, in the recent survey of participating agencies, the majority cited 'learning' (for improved performance) as their principal reason for involvement in the TEC. 'Accountability' ranked fifth as a reason for involvement, partly explaining the plethora of single agency evaluations that took place alongside the TEC. But it should also be noted that learning from the response was the principal stated objective of the TEC.

In the future

One participant proposed that joint evaluations should take place 'in all contexts and at all times'. But this view was modified in the ensuing discussion, with broad agreement that joint evaluations should frequently be considered but will not always be appropriate. For example, governments (especially donor and host governments) do not always share the same standards; ensuring participation of the host government is very difficult in the context of failed and 'difficult' states (see below); and an absence of joint policy formulation, joint programme planning and joint implementation can get in the way of effective joint evaluation. On the flip side of this last point, when the UN has pooled funding joint evaluations may become easier.

Key to deciding when and where to do joint evaluations is to consider the intended audience. System-wide evaluations are most likely to be place-based enquiries, with a particular focus on policy. Two views emerged: that the level of the evaluation should correspond with the level of humanitarian intervention; but it was also suggested that a lot could be learned from joint evaluations of certain small-scale emergencies.

Joint evaluations tend to be retrospective, especially if they are policy-focused. But if greater attention is paid to utilisation then there is an argument to carry out policy-oriented evaluations in ‘real-time’, ‘when people are listening’.

5. HOW TO BRING IN ALL RELEVANT GROUPS

Ownership of joint evaluations was seen as critical if they are to become more utilisation-focused, and this includes ownership by field-staff. It was suggested that a stakeholder map be drawn up early in the planning process for a joint evaluation: this will help to identify which actors need to be substantively involved and which can play a more peripheral role – for example, just being kept informed. This was described as a sliding scale of involvement, recognising that some agencies may be ‘silent partners’ by providing funding or receiving information and reports. Yet when a joint evaluation is launched there tends to be a point at which a critical mass gathers behind it, encouraging others to follow because they do not want to be left out. Above all, there needs to be clarity upfront about how agencies can participate and what they can expect to get out of it.

Much of the discussion focused on how to engage NGOs more in joint evaluations in the future. It was pointed out that the transaction costs of joint evaluations compared with single agency evaluations are often too high for NGO staff, who have limited time available for meetings. Also, Geneva is not their focal point (where the first TEC meeting was held), unlike donors and UN agencies who have tended to dominate system-wide evaluations so far. Greater engagement of NGOs in system-wide evaluations would undoubtedly mean more focus at the field level. Indeed, to achieve greater local ownership one participant urged the involvement of an informed local person / people, not least to act as a kind of ‘liaison officer’ if the joint evaluation is to continue to be headquarters-driven.

The MSF representative made clear his agency’s fundamental objection to joint evaluations: there are not yet agreed principles or unified standards against which international humanitarian action can be judged, although this was regarded as a very idealistic stance by some other participants.

In terms of involving the recipient / host government, this may depend upon whether it is a natural disaster like the tsunami or a conflict-related emergency. In the case of the latter, engaging the host government is likely to be much more difficult, especially if it is party to the conflict. At a minimum it is important to ensure that the affected country’s government is kept fully informed and has the opportunity to participate and comment if it wants to. Unfortunately this was missed in the TEC as the resources were not in place to make it happen.

One participant likened the process of joint evaluation to a social movement. It is hard to legislate for, it requires real leadership and interest as well as a certain minimal level of resources to mobilise, and it can be hampered by a fuzzy ideology.

6. MANAGING JOINT EVALUATIONS

The presenters reflected on the experience of managing joint evaluations to date. Niels Dabelstein differentiated between a one-tier management system, whereby there is a management committee that delegates contracting and day-to-day management to one agency, and a two-tier system, whereby there is a steering committee *and* a management committee. An example of the former was the five-donor joint evaluation of assistance to Afghanistan. The multi-donor Rwanda evaluation is an example of the latter. Rachel Houghton identified lessons learned from managing the TEC – see Box 1. Jock Baker highlighted the difficulty faced by the ECB in getting an agency to take the lead in NGO-joint evaluations.

Box 1: 10 Key Lessons from the TEC

Start with clear groundrules and parameters: This is important for efficiency, transparency and accountability. For example, information concerning:

- Aims and objectives (provide a clear rationale for the process)
- Criteria for membership
- The extent and limitations of authority of member agencies (have a ToR for participating agencies, and include organograms of roles and responsibilities)
- Dispute mechanisms

If such groundrules and parameters are in place, any occasional discomfort over roles can better be avoided and disputes more readily resolved.

Communicate these clearly with member agencies and other interested parties: This will require creating a communications strategy from the beginning of the joint evaluation (JE) process, and setting up a central website as soon as possible.

Provide information management and research support: Decenteralised and distributed evaluations (ie, among a number of different agencies) require a central research service as well as active coordination. Create a secretariat through the employment of a coordinator and an administrator right away. Not doing so for the TEC created a delay in the process.

Locate the secretariat in a sector-wide network: Location of the TEC secretariat independent of member organisations and in a sector-wide network was viewed as key to the success of the TEC.

Management structure: Learning from the TEC process suggests that the overall management structure should include a smaller core group with a wider 'steering committee', more along the lines of the Rwanda evaluation. Having steering committees and working groups for three of the five thematic evaluations was viewed as excellent practice, and contributed much to those evaluations. In the case of the other two evaluations, the employment of special advisors was extremely beneficial.

Representivity in management and the wider group: While those involved in JE's are generally self-selecting, when undertaking a multi-agency JE it is necessary to undertake **face-to-face advocacy** to ensure broad (and relevant) representation. For example, with INGOs, local NGOs, country-office staff and affected-country governments.

Activity in the affected region: Learning from the TEC process has shown the need to either locate a secretariat function in the affected region or at least have some of the management meetings there. This will increase utilisation – and therefore impact – of the study (see below).

Ensure a utilisation focus: If learning from JEs is to be maximised, the evaluation must be utilisation focused. For example, it will be important to work early on with the country offices of member agencies. The creation of a utilisation-focused communications strategy in a workshop with TEC member agencies was considered good practice.

Budget fully: Seed money as well as contingency money needs to be in place before the next JE. Importantly, a utilisation-focused JE will require a budget for follow-up work once the evaluation phase is complete. Agencies involved in management of the JE must understand the responsibility to either provide funding or be involved in fundraising for the process from the beginning. Once most TEC management agencies publicly provided funding, other donors followed suit.

Cross-fertilisation between teams: The provision of opportunities for cross-fertilisation between the evaluation teams, both before, during and after the evaluation, was considered good practice.

In the discussion it was pointed out that many joint evaluations tend to be front-loaded, for example the TEC, which means there is often a gap and loss of momentum when the reports are finished. Yet this is when there needs to be real engagement to address the issues and recommendations that have been identified. If evaluations are more genuinely utilisation-focused this may be avoided in the future. However, donors need to understand that this component of the evaluation needs planning and funding upfront.

Some general learnings to guide management of joint evaluations emerged:

- (i) management structures should be kept simple and light
- (ii) as the experience of doing joint evaluations grows, so will the trust between agencies thus making management easier
- (iii) it is critical to have a core group of 4 or 5 agencies involved at an early stage to move it forward
- (iv) there should be a clear delineation of roles and responsibilities
- (v) joint evaluations require full-time coordination, administration and research support
- (vi) a good communications strategy is critical to keep agencies involved, and can help to build trust in the process amongst agencies that are less-engaged and/ or more sceptical
- (vii) there must be adequate funding, including a contingency budget (if, indeed, the dissemination and follow-up is not funded upfront).

7. A FUTURE ROLE FOR ALNAP

There was broad consensus that ALNAP does have a future role to play in joint evaluations in the humanitarian sector, as follows:

1. Acting as a repository of knowledge and good practice on joint evaluations, with the suggestion that it could update the DAC guidance booklet on system-wide joint evaluations.
2. Holding seed money that is required to ‘kick-start’ a joint evaluation, possibly hosting the secretariat as it has with the TEC.
3. Some participants felt that at least one joint evaluation per year should be on ALNAP’s agenda.

It was also suggested that ALNAP be pro-active in terms of helping agencies prepare for joint evaluations and understand what they entail, for example, through running workshops and providing a guidance booklet.[2]

7. FINAL REFLECTIONS FROM THE FACILITATOR

There is a real sense of the ‘coming of age’ of joint evaluations, that they are widely recognised as having much to contribute and will be part of the humanitarian landscape in the coming years. Discussion seems to be less focused on ‘if’ there should be joint evaluations in the future, but instead on ‘how’ they can best be managed and run. Although some of the debate still focuses on the pros and cons of joint versus singe agency evaluations, this sometimes seems to miss the point that both can be fundamentally different exercises. The value of a system-wide joint evaluation is in addressing high-level policy issues and challenging operational practice across the sector. Single agency exercises are more likely to focus on internal learning and specific accountability requirements. Rather than debating whether one should be promoted versus the other, a more constructive approach would be to explore how they can best complement each other. The most pressing challenges for the next system-wide joint evaluation in the humanitarian sector appear to be:

---

[2] Some participants informed the group that there is a 1.5 day module dedicated to joint evaluations as part of a training course on international evaluations held in Canada. This could be a resource for ALNAP to draw upon.

(i) how to involve the government of the affected country (assuming this is both desirable and feasible);
(ii) how to engage the NGO sector more, both INGOs and especially local organizations;
(iii) how to make the joint evaluation exercise more field-oriented, in terms of its accessibility to field-based staff and organisations and the utilisation of its findings.

Reference documents and background materials for the workshop

- Recent TEC survey results (since updated; www.tsunami-evaluation.org)
- The three AARs undertaken with the TEC CMG and evaluation team members (www.tsunami-evaluation.org)
- Proceedings from the joint evaluation workshop convened by ALNAP in June 2006
- DAC guidance on joint evaluation
- DAC paper: Working paper: Joint evaluations: Recent experiences, lessons learned and options for the future
- Other related papers at: http://www.oecd.org/searchResult/0,2665,en_2649_201185_1_1_1_1_1,00.html
- ODI Working Paper 215 'How the Sphere Project Came into Being': A Case Study of Policy-Making in the Humanitarian Aid Sector and the Relative Influence of Research Margie Buchanan-Smith https://files.odi.org.uk/wdrive/ALNAP/TEC 2005-06/rh_desktop/Data/Evaluation-Research to practice/20030731-How Sphere came into being-Buchanan-Smith.pdf

Margie Buchanan-Smith
9th January 2007