

A ripple in development?

Reflections on the LRRD2
Joint Evaluation Process

Final Report

Margie Buchanan-Smith

Author: Margie Buchanan-Smith.

The views and interpretations expressed in this report are the author's and do not necessarily reflect those of the organisations commissioning this study.

Copyright: The author and the organisations commissioning the evaluation.

Registration No.: 2008–001230

Date of final report: September 2009

Cover: Niklas Hägglund

Published by Sida 2009

Printed by Edita, 2009

Article number: SIDA52512en

This publication can be downloaded from www.sida.se/publications

Table of contents

1. Introduction.....	3
2. Why LRRD2 became a joint evaluation	5
3. The management and planning of LRRD2	6
4. Involvement of partner government representatives.....	9
5. Implementation	13
6. Dissemination and utilisation.....	15
7. Conclusions – a milestone in joint evaluations?	15
References.....	17
Annex 1 List of people interviewed	18
Annex 2 Timeline for LRRD2 evaluation	19
Acronyms.....	20

Reflections on the LRRD2 Joint Evaluation Process

1. Introduction

The quest to improve performance, learning and accountability in the international humanitarian system has spawned a growing interest in the practice of joint evaluations. (A joint evaluation can be defined as an evaluation carried out by two or more agencies, evaluating the work of two or more agencies)¹. But by definition these are complex and challenging exercises, balancing the interests and needs of a range of often very different stakeholders. In 2007, ALNAP commissioned a meta-evaluation of joint evaluations of humanitarian action (Beck and Buchanan-Smith, 2008). Not only did this review and compare the quality of joint evaluations with single-agency evaluations, showing that the former are generally of superior quality, the meta-evaluation was also an opportunity to take stock of the current practice of conducting joint evaluations within the humanitarian sector. It included a particular focus on the second ever system-wide humanitarian evaluation, the work of the Tsunami Evaluation Coalition (TEC), set up to evaluate international disaster support in the wake of the devastating earthquake and tsunami of 26th December 2004². See Box 1. The meta-evaluation concluded that the trend towards joint evaluations in the humanitarian sector is likely to gather pace, but that there are some important learnings emerging from the growing body of experience of joint evaluations, for example in terms of the time it takes to do a joint evaluation, which has consistently been under-estimated, the importance of negotiating and not rushing the formulation of the terms of reference (TOR), and in ensuring in-country buy-in which is critical to the success of many joint evaluations.

The TEC's work was organised according to five themes, one of which was 'linking relief, rehabilitation and development' (LRRD), and the evaluations were published in 2006³. In 2007, the Swedish International Development Cooperation Agency (Sida) took the lead in launching a second phase evaluation of the links between relief, rehabilitation and development (LRRD2). Some three years after the tsunami the role of the international community had shifted. Instead of playing a central role in the response in some countries, they were more involved in supporting national and local recovery efforts. This raises different sets of LRRD issues and shifts the focus more to national and local government agencies.

Channel Research, the consultancy company that was awarded the contract for the LRRD2 evaluation, spotted the opportunity to capture the learnings from this most recent joint evaluation process, with the objectives of advancing learning on the successful implementation of joint evaluations and building on the experience of the earlier TEC studies. Thus, a supplementary piece was added to the main LRRD2 evaluation, an exercise reflecting on the *process* of the evaluation. One of the major criticisms of the TEC was its failure to engage regional actors, especially the governments of the countries affected by the tsunami. (This is a widespread problem in joint evaluations.

¹ This was the definition adopted by the ALNAP meta-evaluation of joint evaluations published in 2008 (Beck and Buchanan-Smith, 2008).

² The first ever system-wide evaluation in the humanitarian sector was the multi-agency Rwanda evaluation in 1996 (Borton et al, 1996).

³ See <http://www.tsunami-evaluation.org/>. The other themes were: coordination, needs assessments, funding and impact on local and national capacities.

See DAC, 2005b). In response, LRRD2 specifically set out to engage representatives of the three countries covered in this evaluation – Indonesia, Sri Lanka and the Maldives. This reflective exercise explores how well this worked and captures lessons that have been learned.

Box 1 Key findings from the ALNAP meta-evaluation’s review of the TEC process

Overall the TEC evaluations were of good quality, for example in reviewing needs and livelihoods assessments and in addressing issues of gender equality, but were weaker on advocacy issues and in presenting recommendations. Other relevant findings for this review include:

- there was a lack of overall conceptual framework to guide the evaluation
- roles and responsibilities could have been more clearly defined (for example between the Core Management Group, evaluation team leader and team members)
- the TEC faced constraints in tying into national capacity
- there were lost opportunities to consult with beneficiaries
- impact and policy received inadequate attention
- recommendations were of mixed quality and follow-up had been poorly planned at the outset

Source: Beck and Buchanan-Smith, 2008

Methodology

Although it would have been ideal to carry out the reflection as a facilitated workshop exercise, this was not possible due to the diverse geographical locations of the main stakeholders. Instead, one-to-one interviews were carried out with all members of the evaluation’s Management Group (MG), selected members of the Joint Steering Committee (JSC), and two members of the evaluation team, guided by a checklist of questions. See Annex 1 for a record of those interviewed. The author also participated in a workshop in Colombo in December 2008 when the preliminary evaluation findings were presented (an opportunity to meet with the Sri Lankan government representatives involved in the evaluation), and has drawn upon the findings of two ‘After-Action Reviews’ carried out with the evaluation team (in November and December 2008 respectively).

This paper presents the findings of these reflections and relates them to the wider debate on conducting joint evaluations. It starts by asking the question of whether it was appropriate for LRRD to be a joint evaluation – section 2. Section 3 reflects on what worked in the management and planning of LRRD2, as well as the challenges and learnings. This is elaborated in section 4 which explores the nature and effectiveness of the engagement of partner country governments in this joint evaluation. Section 5 captures some of the learnings from the implementation phase of LRRD2 while section 6 reflects on dissemination and utilisation. The concluding section pulls together the main learnings from the whole evaluation process and their implications for joint evaluations in the future.

2. Why LRRD2 became a joint evaluation

All interviewees agreed that it was entirely appropriate for LRRD2 to be a joint evaluation because of the topic it was addressing. Linkages between types of programming also means linkages between agencies and actors, and thus raises systemic issues. The opportunity to evaluate beyond the somewhat artificial boundaries of one donor's portfolio was welcomed by the evaluation team.

However, LRRD raises different types of issues for the international community than it does for national governments and has been mainly the preoccupation of international agencies in the last decade (with the possible exception of the disaster risk reduction agenda that national governments have been encouraged to take up). Thus, it is not surprising that the initiative for LRRD2 came from the international community, and that their perspectives and concerns have dominated the evaluation despite the involvement of national government representatives from the region (see section 4 below). The focus on international aid was clear early on:

'Although it is the aid effort that shall be evaluated, the evaluation should consider the aid response in the perspective of the wider trends that have driven recovery, including household, community and government initiatives and the wider economic and market related context'. (italics added)
(Sida, 2007:1)

Indeed, some members of the evaluation team felt that the aid-centric nature of LRRD2 was a missed opportunity to learn more about government policies and response⁴.

The four themes selected for LRRD2 built on the structure of the LRRD1 report (Christoplos, 2006):

- (i) the return of the state and civil society
- (ii) poverty, livelihoods and economic recovery
- (iii) rebuilding the social fabric and community development
- (iv) reduction of risks from natural hazards and conflict

A fifth theme was added – capacity-building – in response to feedback from JSC members (see below).

Using aid as the starting point for looking at linkages was of less interest to governments in the region. The Government of Sri Lanka representatives in LRRD2 confirmed that they were more concerned with questions of impact than with 'linkage' questions⁵. The Government of Indonesia representatives would have liked more focus on national government policy, for example on rehabilitation and housing (although this is not necessarily incompatible with linkage issues)⁶. Indeed, one interviewee made a convincing argument that if the role of government in responding to the tsunami and linking the different phases of response had been the primary concern in LRRD2, then a different set of countries should have been selected to capture learning and good practice: the tsunami response was much more government-driven in Thailand and India than in Sri Lanka, Indonesia or the Maldives.

⁴ Sida's evaluation manager would have liked to have included a stronger focus on government response, but felt uncertain about the mandate of such a joint evaluation.

⁵ However, it is also worth noting that TRIAMS (the Tsunami Recovery Impact Assessment and Monitoring System) was specifically designed to focus on national outputs and impacts.

⁶ Danida's recent publication on partner country involvement in joint evaluations came up with similar findings. Partner country governments were more interested in impact and closer alignment with national policy development (Jensen and Eriksson, 2009).

Not surprisingly, one of the challenges faced by the LRRD2 evaluation was a waning of interest in the tsunami response some three years after the event. This is evident from the half-hearted engagement by some members of the international community; for many their priorities had shifted to more recent humanitarian crises. And it partly explains the limited engagement of national government representatives. For the latter, the tsunami response had been eclipsed by recent and bigger political events. In the Maldives there were elections and a change in president during the LRRD2 evaluation period; in Sri Lanka, the war in the north had entered a new phase of intensity; in Indonesia, government representatives said they were as, or more interested in issues of post-conflict recovery than in issues of post-tsunami recovery. In the words of one interviewee, natural disasters like the tsunami can become 'ancient history' very quickly in terms of national government priorities. 'Selling' the concept of an evaluation on LRRD had not been easy for Sida at the outset.

Despite these challenges, most interviewees felt that the LRRD2 evaluation succeeded in addressing systemic issues if the focus of the 'system' is taken to be the international aid system. The objective of LRRD2 was more to contribute to learning, in keeping with the TEC, than it was to do with accountability although reporting on results was still judged to be important. The opportunity to do this system-wide on linkage issues that are common to almost all international agencies was welcomed. This was captured by one interviewee in the following words: 'the wide umbrella of the joint evaluation meant that wider lessons were identified with a deeper understanding of the problem'.

Key learning

Evaluations that take place some time after the crisis will almost always face the problem of waning interest and of other issues taking priority. This needs to be taken into account in planning the evaluation, and especially the timing, for example tying it into policy formulation processes and identifying other windows of opportunity to influence decision-making.

3. The management and planning of LRRD2

The organisation of LRRD2

As with LRRD1, Sida's Evaluation Department took the lead in initiating and managing LRRD2. The evaluation was funded by four donors who each have a strong track record in supporting and/ or leading joint evaluations. These were the Ministry of Foreign Affairs (MFA) in Norway, the MFA in the Netherlands, Danida and Sida. Other agencies who were actively involved, in addition to governments in the region, include: NORAD, CIDA, OCHA, UNICEF, IFRC, and CARE International.

Sida proposed the commonly used two-tier management structure for LRRD2. Thus, there was a four member Management Group (MG), and a larger Joint Steering Committee (JSC) representing 17 governments/ agencies at the outset. The division of responsibilities between the two bodies is briefly spelt out in the TOR for LRRD2, whereby the MG handles the day-to-day management and the JSC's inputs relate to endorsing the TOR, advising on criteria for selection of the consultants, deciding on the design of the field work part of the evaluation, and commenting on the draft reports.

The MG comprised Sida, Danida, and unusually, representatives from the Indonesian and Sri Lankan governments. This group met a couple of times in the region, although in practice it was the Sida and Danida members that were most active in decision-making, especially on contractual issues. The JSC met face-to-face three times, in Bangkok, Jakarta and Colombo, with a final teleconference meeting when the draft report had been circulated. See Annex 2 for the timeline for LRRD2. Engagement by JSC members varied considerably. A small number were particularly active in commenting on the draft reports while others played no role at all, not even participating in JSC meetings to which they had been invited (and were consequently not mentioned in the final evaluation report).

Sida had planned to establish an Advisory Group for LRRD2. In the end the advisory role was played by one person – the same adviser as LRRD1 who had also authored the synthesis from LRRD1, and whose input was valued by Sida.

How it worked in practice

One of the well-documented learnings from earlier joint evaluations is the importance of taking time to plan the joint evaluation and to draw up the TOR⁷. This is the critical phase for negotiating and developing a consensus between the different stakeholders with different interests, and to build a sense of collective ownership⁸. LRRD2 appears to have taken this lesson on board by really investing in the planning stage which took more than a year (see Annex 2). An approach paper for LRRD2 was written and circulated in March 2007. During the rest of that year the approach paper was shared and consultations ongoing. Two events were particularly important:

- (1) a visit by Sida to the region in December 2007 to consult with the respective national governments to seek their views and potential support for LRRD2
- (2) the first JSC meeting in Bangkok in June 2008 when the TOR were discussed amongst the main stakeholders, including representatives of the governments in the region.

As discussed in the next section, these were key to engaging partner country governments. However, one of the consequences of wide consultation on the TOR is that they can keep expanding as different stakeholders add questions to cover their particular interest, and this is what happened with LRRD2. Originally designed with four themes, a fifth – capacity-building – was added after the first JSC meeting in Bangkok at the request of a number of members. The final TOR contained more than 50 questions, although this list was differently interpreted by different actors. Sida, and to an extent the evaluation team, saw this as an indicative list, whereas some JSC members saw it as a comprehensive list to be answered. With hindsight, a number of members of the JSC reflected that the scope of the TOR had become too big for the time available. In the words of one JSC member, 'we're drowning the fish; when we try to be comprehensive we lose focus in an evaluation'. Some would have welcomed stronger executive decision-making to reduce the scope of the evaluation, and/ or more negotiation by Channel Research to prioritise and ensure the evaluation was manageable. Another JSC member commented:

⁷ See Beck and Buchanan-Smith, 2008; DAC, 2005

⁸ Failure to do this for the TEC was documented in one of their 'After Action Reviews, of 9th September 2005.

'I wonder if more time spent up front on the TOR, evaluation questions and converting these into a schedule and report outlines would have meant less time re-doing the final documentation'.

Ideally this issue would have been addressed with the JSC when the inception report was submitted, but as mentioned below, the timing was tight for this part of the evaluation.

The management structure of a joint evaluation is one of the most critical determinants of whether the multi-stakeholder engagement actually works. Once again, LRRD2 had learned from the northern-dominated TEC, for example by holding all three face-to-face JSC meetings in the region (apart from the last teleconference call). This helped to shift the centre of gravity from northern donor capitals to the region, although as documented below there was still limited engagement from the regional government participants during some phases of the evaluation. Most of the JSC members were representatives of donor governments, UN agencies and the governments of the three countries covered by the evaluation, as well as IFRC. Only one NGO was engaged, CARE International⁹. This was surprising as NGOs have played such an active role in the tsunami response¹⁰. Their limited engagement, however, is in line with earlier findings that the transaction costs of being involved in joint evaluations are often too high for many INGOs, let alone national NGOs (Buchanan-Smith, 2007).

In practice, the organisation of LRRD2 was quite 'light'. There were no meetings between the whole evaluation team and the JSC, for example (nor a face-to-face meeting for the evaluation team alone – see section 5 below). Instead, the main focal point for the MG and JSC was the evaluation team leader. Indeed, a number of JSC members were 'sleeping members'. To some extent this is inevitable when so many stakeholders are involved, although it also reflects agency preoccupation with more recent events. The more active JSC members were involved principally in commenting on the TOR and on the draft report. The time it took to do the latter 'surprised' and even 'flabbergasted' some; they reported it took at least one day to read and comment on each draft. The limited contact between the evaluation team and the JSC hints at a missed opportunity to exchange views and debate issues that arose during the evaluation (although this missed opportunity is not unique to LRRD2). As this was a learning-oriented exercise, such ongoing discussions could have encouraged the learning process. The varied levels of evaluation expertise among the JSC members was also commented upon by some interviewees¹¹. In practice this meant they had different expectations of the quality of the final product, and disagreements arose about the acceptability of the final report.

Strong leadership of a joint evaluation is acknowledged to be a key factor determining its success, and Sida stepping into this role was widely appreciated. The leadership style that characterised this joint evaluation was a highly consultative one that focussed on building consensus. This approach was generally welcomed, as was Sida Evaluation

⁹ Urgence, Réhabilitation, Développement (URD) was involved in the JSC during the early stages of the evaluation, but was less active in the latter part.

¹⁰ Sida had faced a dilemma in deciding which INGOs they should actively encourage to participate after the TEC when large numbers had been involved.

¹¹ This raises an interesting question about whether there is now a higher turnover of staff within evaluation departments of donor governments (and other agencies), which means that there are more generalists than evaluation specialists.

Department's familiarity with the topic, although there were some dissenting voices who would have liked stronger leadership from Sida, for example in reining in the scope of the TOR, in quality assurance when the first draft reports were submitted before they were circulated to the JSC, and in reconciling different views amongst JSC members (particularly an issue when the draft report was being circulated)¹². The evaluation team, however, welcomed Sida's lead in management decisions because it facilitated speedy decision-making.

Key learning

While investing time and resources in the planning phase is critical to the success of a joint evaluation, especially allowing wide consultation on the TOR, there also needs to be a clear process of executive decision-making to ensure that the TOR do not become too broad, but define a coherent and do-able piece of work.

4. Involvement of partner government representatives

Involvement of national governments in joint evaluations of humanitarian action has been weak and has been identified as a key challenge for the future (Beck and Buchanan-Smith, 2008). A recent publication by Danida on joint evaluations comments on the lack of explicit policy by donor governments in involving partner countries in evaluation work. It also presents varying views on the obstacles and issues to be addressed, including lack of demand for evaluations at country partner level (Jensen and Eriksson, 2009). On the development side, the evaluation process for the Paris Declaration is breaking new ground in terms of involvement of recipient governments; the humanitarian side is less advanced.

Commendably, involving partner governments was identified as an early priority for LRRD2. The intention was to gain government involvement and ownership of the evaluation and of its findings, to better 'judge' the relevance of the findings, and to facilitate the use and dissemination of the report. As already mentioned, there was considerable investment in the early planning stage of the evaluation, including the trip to the region by Sida's evaluation manager in December 2007 to explore interest in LRRD2 and ways of cooperating. The response to these early overtures was positive, and recent interviews with the government representatives on the JSC confirmed their appreciation of being involved. In the words of one, this is a 'step in the right direction – working collaboratively between international agencies and government in evaluation'.

The peak of engagement of partner government representatives appears to have been the first JSC meeting in Bangkok in June 2008 to discuss and agree upon the TOR. The government representatives felt listened to and therefore able to feed into this crucial planning stage of LRRD2, although there was later some concern with the final product, in that it did not adequately reflect upon government policy (mentioned in section 2 above) nor deal sufficiently with the complexity of the situation. Thereafter, however, the engagement of the government representatives was much more limited despite Sida's efforts to hold all key meetings in the region. For example, the Indonesian government representatives were not present at the Jakarta JSC meeting and workshop to kick-off

¹² These are notoriously difficult issues to deal with in a joint evaluation, especially finding a way to deal with several rounds of comments. It is worth noting Jensen and Eriksson's (2009) recommendation that comments be discussed in a workshop, although this may be a more costly and time-consuming approach.

the second field-based phase of the evaluation, nor at the Colombo workshop in December 2008 when the preliminary findings of the evaluation were presented even though all travel expenses would have been covered.

Although the evaluation team said that the involvement of government representatives in LRRD2 had helped them during implementation, for example in gaining access to other government officials and in providing background information, LRRD2 did not develop into a joint evaluation where there was equal engagement of international agencies and national governments. After the planning stage of LRRD2 national government involvement really waned. This reflective exercise has revealed a number of reasons why this was the case:

1. As mentioned in section 2 above, the LRRD concept was of less interest to governments in the region for whom there were other pressing political priorities than it was to donor governments and to international aid agencies. As a result, the government representatives that were members of the JSC and/ or MG were often engaged in other business that was regarded as a higher priority than participating in LRRD2 evaluation meetings.
2. Evaluation has become a well-established activity within the international aid sector, but most developing country governments do not have an evaluation culture¹³. This throws up a number of challenges. First, what is the appropriate government department to engage with? One donor representative on the JSC commented that it could have been any one of around seven or eight different government departments in Sri Lanka, although in the end the decision was made to engage with the respective planning departments in each of the partner country governments. Second, there are few if any government officers with experience of evaluations which mitigates against the joint evaluation being conducted as an equal partnership. This was the case in Sri Lanka where the most experienced government officer was already working on the evaluation of the Paris Declaration; neither of the two government officers selected for LRRD2 felt they had sufficient knowledge or experience to contribute as actively as they would have liked. They also had difficulty with some of the language and evaluation jargon that was used.
3. Although meetings were held in the region, there was still a lot of reliance on e-mail and written communication during the LRRD2 process. But as some of the partner government representatives observed, this is less culturally appropriate in south-east Asia: 'communication by e-mail may work in the western world, but in the developing world you must meet people'. This clash of communication cultures was recognised by some donor representatives on the JSC, particularly at the point of discussing the final report. The requirement for written comments can encourage a more bureaucratic response that must be approved by higher levels of authority.
4. Although the intention of including government representatives in the Management Group was to give them a central role in the evaluation, the government representatives themselves felt that the discussions and decisions

¹³ Interestingly, the Government of India is one of the exceptions, with a dedicated and independent government evaluation section in the National Planning Council, although it was not part of LRRD2.

were mostly of a contractual and financial nature which was of less interest and relevance to the governments in the region. The fact that this did not feel like an equal partnership is common to many joint evaluations (Jensen and Eriksson, 2009).

As a consequence of all of the above, there was little exchange and learning between governments in the region throughout this LRRD2 process. Instead, the exchanges tended to be between donor governments and regional representatives.

So what are the *key learnings* that can be drawn from this experience?

- 1. First, the lead time to engage partner country governments can be considerable¹⁴. This was an important and appropriately slow step in LRRD2, although it is still questionable whether partner governments and donor agencies had the same expectations of what it meant to engage in the evaluation, despite early discussions on this topic. Danida's recent review of joint evaluations with partner countries demonstrates how this first stage can take even longer where there is some resistance or breakdown in trust between donor and recipient governments (Jensen and Eriksson, 2009). It also provides a useful checklist of good practices for the initiation stage – see Box 2*
- 2. A critical part of this first planning stage is to identify which government department(s) should be involved. This should be determined partly by the purpose of the evaluation. Thus, for an accountability-oriented evaluation it may be appropriate to link up with the national audit office (or equivalent), while for a learning-oriented evaluation like LRRD2 it may be more appropriate to link with the implementing departments involved in programming (although a challenge may still arise in choosing which implementing department, as in LRRD2). One partner government representative urged greater involvement with local government representatives in this kind of joint evaluation; they are more familiar with the situation on the ground and may benefit from the learning opportunity. This was confirmed by members of the evaluation team who did not feel that the TOR adequately reflected the reality on the ground.*
- 3. LRRD2 was relatively successful in shifting the centre of gravity of the evaluation to the region, particularly by holding the first three JSC meetings in partner countries. But to ensure that decision-making is also driven by regional concerns, it may be important to consider different levels of steering committees, both an international one and a regional or national one¹⁵, and to facilitate greater involvement of local government.*
- 4. Partner country involvement needs to be more than participation in meetings for it to be a truly joint evaluation. Although there are many obstacles to making it an equal partnership with international agencies (including capacity constraints,*

¹⁴ It is worth noting that it took around 2.5 years for TRIAMS to achieve country-level buy-in in Indonesia, Sri Lanka, Maldives, Thailand and India

¹⁵ The Interagency Health Evaluations (IHE), initiated in 2003 by WHO and UNHCR, provide some positive examples of how in-country steering groups, in addition to an international core management group, helped to create national-level ownership.

language barriers, timetabling, power relations etc¹⁶), the evaluation team can play an important role in maintaining the interest and engagement of partner governments. This implies playing a more facilitative role (eg through holding workshops as part of the evaluation process) and paying attention to stakeholder relationships as well as to the technical aspects of the evaluation. This is particularly important for a learning-oriented evaluation, and will have budgetary consequences.

5. When it comes to commenting on the evaluation report, Jensen and Eriksson's findings of Danida's experience are very relevant:

'For all the evaluations reviewed it was found that the conduct of workshops was the most effective means to obtaining comments to draft reports for several reasons:

- *Firstly, workshops were seen as building on the strong oral tradition of countries such as Ghana and Mozambique;*
- *Secondly, participation in workshops was regarded as more efficient by the participants since reading requirements are limited and findings are presented and discussed during the workshop;*
- *Thirdly, the conduct of workshops was seen as having the added benefit of explaining and interpreting findings to participants; and*
- *Fourthly, workshops were regarded as vehicles for enhanced country and sector dialogue, especially to a large but disparate stakeholder community, as in the case of the Tanzania health sector evaluation'.*

(Jensen and Eriksson, 2009: 29)

This chimes with feedback on LRRD2. The evaluation team felt that there was a good understanding of the issues when discussed in workshops, and the evaluation team felt energised by the engagement in these discussions.

Box 2 Good practices identified with regard to evaluation initiation

Promote evaluations initiated by partner country stakeholders by:

- Calling for suggestions about evaluation topics
- Supporting calls for (joint) evaluations on overall national development strategies including national poverty reduction strategies
- Supporting the establishment of a joint inter-sectoral evaluation committee
- Programming for joint evaluations at sector level

For donor HQ initiated evaluations:

- Program for evaluations already during the development of country strategies and/or country programmes
- Initiate evaluations through established government/development partner mechanisms where they exist
- Approach targeted national partner early to allow for adequate resourcing and strong ownership
- Communicate clearly the value-added of the evaluation, especially for sector evaluations in countries where advanced review structures are already in place

¹⁶ See Jensen and Eriksson, 2009

- Be candid about whose accountability needs the evaluation is seeking to meet (donor and/or partner constituencies in the form of e.g. senior management, executive boards, ministers, and parliaments)
- Be patient, a proposed evaluation might have to be launched years after initiation in order to align with partner country processes

Source: Jensen and Eriksson, 2009: 17

5. Implementation

Channel Research fielded a large evaluation team: 14 members plus three peer reviewers/ advisers. Whilst this may mean a wider range of expertise within the team, there is a trade-off in managing and coordinating the work of such a large team. One of the learnings from LRRD2 is the importance of the whole team meeting face-to-face, which unfortunately did not happen. Not only does this put a lot of pressure on the team leader to hold the evaluation together, it can also impair the coherence of the team's work and the opportunity for debate and exchange of findings between team members. During the After Action Review in Colombo in December 2008, the evaluation team members reflected that it would have been beneficial to have met together before the field work started and half-way through the field work to exchange early findings. This would also have promoted communication and coherence between theme leaders. Shortage of time was partly responsible for this (discussed below).

More positively, the evaluation team members felt that the chemistry within the team had worked well, and appreciated the strong participation by national consultants. There was a sense of having learned some of the lessons from LRRD1 (also carried out by Channel Research), for example in drafting clearer questions for the quantitative survey, and in producing a more accessible and better annotated documentation review.

As with any multi-country evaluation there is a dilemma in deciding how to split the team – by country or by theme. In LRRD2 this was mainly done by country to accommodate the relatively short time inputs of individual consultants and the overall time pressure. However, this inhibited comparisons between countries and in-depth analysis of some systemic issues, particularly where the theme leaders and the team leader were not able to visit all countries. The combination of a large team, over-ambitious TOR, and therefore the parcelling of the work into discrete chunks, led some interviewees to conclude that the fieldwork was 'too thin without enough depth'.

This was exacerbated by the very tight timescale for the implementation of LRRD2. There was a trade-off to the generous amount of time dedicated to the planning phase of the evaluation – over twelve months. In contrast, only about five months was allocated to the implementation phase, which was not enough for the team to cover four themes in three countries, and for the evaluation to be carried out in two phases. This pressure was felt keenly by the evaluation team members (as reported in an After Action Review in Colombo in December 2008). As is often the case, the timeframe appears to have been driven by institutional and budgetary pressures in donor headquarters which compromises the evaluation process.

Some of the consequences of this tight timeframe were the following:

- Very limited time between the awarding of the contract (in September 2008) and the start of the field work (in early November 2008) which meant that not all team members were available at the same time for the field work phase.
- The documentation review had to be carried out rapidly in October, and there was inadequate time before the field work for the findings to be fully taken on board in designing the field work.
- Similarly, an inception phase of one month after signing of the contract was felt to be too short. The evaluation team had not managed to meet and the quick progression into field work put unhelpful pressure on JSC members and on the LRRD2 adviser to be able to comment and influence the design of phase 2.
- Inadequate time for fieldwork is a common complaint in evaluations, especially in complex multi-country joint evaluations. It does seem to have been a major constraint for LRRD2, recognised by both the evaluation team members and by members of the JSC. There was a sense that the field work was rushed, and that there was inadequate time between the qualitative and quantitative surveys. For example some of the detailed analysis from the quantitative survey was too late to inform the report, and there was no opportunity for the leader of the quantitative survey to discuss the findings with the lead writers. The tight timeframe also discouraged more reflective and process-oriented methods, for example facilitated workshops with key stakeholders.
- Although the time allocated to writing up was similar to the writing time allocated to many evaluations, it was short for a joint evaluation of this kind where there is a large team and the need for team members to meet in person to complete the analysis and harmonise findings.

A number of interviewees felt that the quality of the evaluation suffered as a consequence of the above.

Key learning

Many joint evaluations are learning-oriented. This focuses the spotlight on process during the joint evaluation, in particular creating opportunities for the evaluation team to engage with key stakeholders as much as possible during the evaluation eg in facilitated, reflective workshops, and to be debating findings whilst the evaluation is ongoing.

The time for carrying out a joint evaluation that covers multiple actors and in this case a number of countries must be carefully planned and will almost always be more than the time needed for a single agency evaluation at each stage. For the best results, the implementation phase should not be rushed, and there should be adequate time and financial resources for the evaluation team to meet in person on at least two occasions: when the fieldwork is being planned, and to discuss and agree the preliminary findings. The management and coordination trade-off of a large evaluation team must also be recognised.

6. Dissemination and utilisation

Dissemination and communication were well-thought through for LRRD2. A communication plan was drawn up by the Chairperson of the MG well in advance of the final report being available, towards the end of 2008. This was circulated and discussed amongst the MG. Interviewees highlighted some positive aspects of the dissemination process, specifically:

- translating the final report into several languages (Bahasa, Indonesian, Tamil, Sinhalese) to increase its accessibility within the region
- holding a series of dissemination workshops in the region, in Jakarta, Bangkok and Male, as well as in Geneva and Stockholm in Europe.

The process of finalising the draft report drew mixed comments. A number of JSC members struggled with the accessibility of the report, partly because of language, and partly because of style which some felt presented the findings in too general and abstract a form. The report went through two drafts before it was finalised enabling comments to be taken on board and the report to become more accessible. Whilst JSC members felt that the evaluation report successfully highlighted pertinent systemic issues, they also questioned how many of the findings were really new. LRRD is a well-researched topic: to what extent did the evaluation deepen insights and identify how the system can be changed, and by whom? This was an issue also faced by the wider TEC (Beck and Buchanan-Smith, 2008). For a topic that has long been debated, the challenge for the evaluators is to ensure that new findings and ideas are spotted amongst the well-worn ones by the readers and potential users.

LRRD2 faces a common challenge of joint evaluations (DAC, 2005a). Compared with single agency evaluations there is no clear locus of responsibility for taking forward its findings and recommendations. The dissemination workshops can play an important role in promoting learning if they encourage debate and reflection by actors to whom the recommendations are directed. However, at least a couple of JSC members felt that key stakeholders should be encouraged to do a management response; indeed, some stakeholders, like SIDA, will automatically do a management response to this evaluation.

7. Conclusions – a milestone in joint evaluations?

LRRD2 faced many challenges that are common to joint evaluation, for example keeping all stakeholders fully engaged, allocating adequate time for such a complex (and in this case, multi-country) exercise, and in ensuring the findings and recommendations are taken up. But it was also an important step forward in applying the learnings from the recent TEC experience, the most significant of which was ensuring that the evaluation was not northern driven by the international humanitarian sector alone, but really engaged stakeholders in the region.

On this latter point LRRD2 has most to contribute to the growing body of knowledge and experience on joint evaluations in both the humanitarian and development sectors because of the considerable efforts by Sida to involve government representatives from the region. This got off to a good start, in large part due to Sida's visit to the region and holding the first JSC meeting in Bangkok. Thereafter government participation waned, despite the fact that two further workshops were held in the region during the evaluation, and only really picked up again during the dissemination phase. This experience reveals

some of the challenges to promoting real partnership in a joint evaluation particularly with governments where there is no evaluation culture. The learnings are presented in section 4.

LRRD2 ran into the common problem of a rushed and inadequate timeframe. This was unfortunate as it meant that the opportunity to do a more in-depth piece of work was missed. It also indicates how much more time is needed for a joint evaluation (especially one that covers a number of countries) than for a single agency evaluation. Given the resources that go into joint evaluations, not just financial but also human resources, this raises a big question about the most efficient and effective use of those resources. It may also be a determining factor in the appetite for future joint evaluations. Indeed, in this exercise some interviewees suggested giving greater consideration to more research-oriented and longitudinal pieces of work.

References

Beck, T. and Buchanan-Smith, M. (2008) 'Joint Evaluations coming of age? The quality and future scope of future evaluations', *ALNAP 7th Review of Humanitarian Action*, London: ODI

Buchanan-Smith, M. (2007) 'Joint Evaluations', report on a workshop on joint evaluations at ALNAP Biannual, Rome, December 2006

Christoplos, I. (2006) 'Links between relief, rehabilitation and development in the tsunami response. A synthesis of initial findings', TEC

DAC (2005a) 'Joint Evaluations: Recent Experiences, Lessons Learned and Options for the Future', DAC Evaluation Network Working Paper

DAC (2005b) 'Workshop on Joint Evaluations. Challenging the Conventional Wisdom – the View from Developing Country Partners. Nairobi 20-21 April 2005. Workshop Report'

DAC (2006) 'Guidance for Managing Joint Evaluations'

ECB (2007) 'What We Know About Joint Evaluations of Humanitarian Action. Learning from NGO Experiences.' Draft (v4), July

Jensen, R.I., and Eriksson, J. (2009) 'Experiences with Conducting Evaluations Jointly with Partner Countries', Danida Evaluation Study 2009/3, Ministry of Foreign Affairs of Denmark, May

Sida (2007) 'Evaluation of Links between Relief, Rehabilitation and Development (LRRD), Phase Two', Draft Approach Paper, March 18

Annex 1 List of people interviewed

Members of Management Group (also members of JSC)

Stefan Dahlgren	Deputy Head of Evaluation, Sida
Lars Elle	Senior Evaluation Adviser, Evaluation Department, Danida
Dr Suprayoga Hadi	Director for Special Areas and Disadvantaged Regions, National Development Planning Agency, BAPPENAS, Jakarta, Indonesia
A. Abeygunasekara	Additional Secretary (Development), Ministry of Plan Implementation, Colombo, Government of Sri Lanka (informal discussions in Colombo, Dec 08)

Members of Joint Steering Committee

Ted Kliest	Senior Evaluation Adviser, Evaluation and Inspection Division, Netherlands Ministry of Foreign Affairs
Satianan Debidin	Evaluation Division, CIDA
Claude Hilfiker	Evaluation Office, OCHA
Robert Ondrusek	Adviser, TRIAMS, IFRC Regional Delegation, Bangkok (by e-mail)
Jock Baker	Head of Evaluations, CARE International
Eddy Purwanto	Former Chief Operating Officer, BRR, Banda Aceh, Indonesia
R. Tharmakulasingam	Additional Secretary (Planning & Development), Ministry of Nation Building & Estate Infrastructure Development, Colombo, Government of Sri Lanka (informal discussions in Colombo, Dec 08)

Adviser to Evaluation

Ian Christoplos	Independent consultant
-----------------	------------------------

Evaluation team members

Emery Brusset	Evaluation team leader
John Cosgrave	Member of evaluation team (documentation review & disaster risk reduction)

Annex 2 Timeline for LRRD2 evaluation

2007

March	Approach paper for LRRD2 drafted and shared by Sida (at TRIAMS workshop in Bangkok, with ALNAP members, and subsequently more widely)
December	Stefan Dahlgren of Sida visited regional capitals to canvas opinion and seek support for LRRD2

2008

April	Sida circulates progress report and design outline for LRRD2
June	First meeting of JSC in Bangkok. Management Group appointed. TOR discussed and finalised
July	Invitation to tender for LRRD2 issued by Sida
September	Contract for LRRD2 awarded to Channel Research Evaluation began
October	Inception report submitted First draft of documentation review completed
November	Management Group meets Second JSC meeting, Jakarta. Presentation of documentary review by evaluation team, and plans for stage 2 of evaluation After Action Review by evaluation team, Jakarta Fieldwork commenced
December	Stakeholder workshop in Colombo for presentation of preliminary evaluation findings Third JSC meeting, Colombo After Action Review by evaluation team, Colombo

2009

February	Submission of first draft evaluation report
March	Fourth JSC meeting by teleconference
May	Acceptance of final evaluation report
May	Workshop in Geneva to disseminate evaluation findings
June to August	Workshops in Jakarta and Bangkok to disseminate evaluation findings
September	Workshop in Stockholm to discuss evaluation process

Acronyms

ALNAP	Active Learning Network for Accountability and Performance in Humanitarian Action
CIDA	Canadian International Development Agency
DAC	Development Assistance Committee
IFRC	International Federation of the Red Cross and Red Crescent Societies
JSC	Joint Steering Committee
MFA	Ministry of Foreign Affairs
MG	Management Group
Sida	Swedish International Development Cooperation Agency
TEC	Tsunami Evaluation Coalition
TOR	Terms of Reference
TRIAMS	Tsunami Recovery Impact Assessment and Monitoring System
URD	Urgence, Réhabilitation, Développement

