Evaluation of Humanitarian Action (EHA)

Course Reference Manual

This manual is for the beginner/intermediate and advanced EHA courses run by Channel Research on behalf of ALNAP. It is supported by a course specific set of case studies and exercises and by a bibliography of evaluation references.

ALNAP

http://www.alnap.org

and

Channel Research

http://www.channelresearch.com

The material in this manual has been developed by the training team of Margie Buchanan Smith and John Cosgrave. Some of the material is based on earlier work by John Telford.

Version 1.30

TABLE OF CONTENTS

INTRODUCTION

This manual was developed to support the training programme on the Evaluation of Humanitarian Action (EHA), delivered in collaboration with ALNAP. This training is facilitated by Margie Buchanan-Smith and John Cosgrave, who are responsible for developing this manual to support the trainings.

The trainings are further supported by a set of presentations, exercises, and case studies specific to each training, and by a bibliography of evaluation references.

There are two trainings. The first is an introductory to intermediate level training which runs for three days. This training is open to evaluators, evaluation managers and users. It presents an opportunity for learning and exchange between participants from different backgrounds and who play different roles. This 3-day training programme has been developed drawing on a variety of sources and is constantly updated.

The second training focuses on advanced topics. This training is for evaluators and evaluation managers who already have extensive experience of evaluations of humanitarian action. The course focuses on areas which often pose challenges for evaluators and evaluation managers. It is an opportunity for participants to reflect on their own experience and the experience of peers to improve their work.

This manual contains most of the reference materials that are used on both trainings, and more, and is therefore constantly being updated. For ease of use and accessibility it has been divided into eight sections.

Section one:
Deciding to do an evaluation

When deciding to do an evaluation there are a number of factors to consider.

This section first looks at a definition of what humanitarian evaluation is and then discusses some of the factors to be considered in this decision. It explores the purpose of evaluations – accountability versus lesson-learning – stressing the importance of being clear about the purpose and the implications for how the evaluation is designed. Accountability is unpacked, and evaluation contrasted with monitoring and audit functions.

The section ends with a description of real-time evaluations and joint evaluations, both growing trends in the humanitarian aid sector, indicating when such evaluations are worth considering.

WHAT IS THE EVALUATION OF HUMANITARIAN ACTION?

HUMANITARIAN ACTION

The best current definition of Humanitarian Action is that provided by the Good Humanitarian Donorship initiative (Good Humanitarian Donorship, 2003).

1. The objectives of humanitarian action are to save lives, alleviate suffering and maintain human dignity during and in the aftermath of man-made crises and natural disasters, as well as to prevent and strengthen preparedness for the occurrence of such situations.

2. Humanitarian action should be guided by the humanitarian principles of humanity, meaning the centrality of saving human lives and alleviating suffering wherever it is found; impartiality, meaning the implementation of actions solely on the basis of need, without discrimination between or within affected populations; neutrality, meaning that humanitarian action must not favour any side in an armed conflict or other dispute where such action is carried out; and independence, meaning the autonomy of humanitarian objectives from the political, economic, military or other objectives that any actor may hold with regard to areas where humanitarian action is being implemented.

3. Humanitarian action includes the protection of civilians and those no longer taking part in hostilities, and the provision of food, water and sanitation, shelter, health services and other items of assistance, undertaken for the benefit of fected people

EVALUATION

The OECD DAC glossary of key terms in Evaluation and Results Based Management defines evaluation as:

The systematic and objective assessment of an on-going or completed project, programme or policy, its design, implementation and results. The aim is to determine the relevance and fulfillment of objectives, development efficiency, effectiveness, impact and sustainability. An evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision–making process of both recipients and donors.

Evaluation also refers to the process of determining the worth or significance of an activity, policy or program. An assessment, as systematic and objective as possible, of a planned, on-going, or completed development intervention.

Note: Evaluation in some instances involves the definition of appropriate standards, the examination of performance against those standards, an assessment of actual and expected results and the identification of relevant lessons (OECD/DAC 2002 p21-22).

Different donor administrations sometimes offer slightly different definitions of evaluation, but the current trend is to try to reflect the OECD/DAC standard terms as much as possible.

HUMANITARIAN EVALUATION

The 2005 ALNAP guide on the use of the OECD/DAC Criteria in Humanitarian defines the Evaluation of Humanitarian Action as the

" systematic and impartial examination of humanitarian action intended to draw lessons to improve policy and practice and enhance accountability" (Beck, 2006 p14).

REFERENCES AND RESOURCES

Beck, T. (2006). *Evaluating humanitarian action using the OECD-DAC criteria.* London: ALNAP. Last viewed on 8 June, 2008. URL: www.odi.org.uk/alnap/publications/eha_dac/pdfs/eha_2006.pdf

Good Humanitarian Donorship. (2003). *Principles and Good Practice of Humanitarian Donorship*. Stockholm: Germany, Australia, Belgium, Canada, the European Commission, Denmark, the United States, Finland, France, Ireland, Japan, Luxembourg, Norway, the Netherlands, the United Kingdom, Sweden and Switzerland. Last viewed on 14 May 2009. URL: http://www.reliefweb.int/ghd/a%2023%20Principles%20EN-GHD19.10.04%20RED.doc

OECD/DAC. (2002). *Glossary of key terms in evaluation and results based management* (Evaluation and Aid Effectiveness: 6). Paris: OECD/DAC Working Party on Aid Evaluation. Last viewed on 21 January 2009. URL: http://www.oecd.org/dataoecd/29/21/2754804.pdf

WHAT IS THE PURPOSE OF EVALUATION: 'LESSON LEARNING' AND/OR 'ACCOUNTABILITY'?

As the definition of evaluation implies, it is usually intended to play an accountability as well as a learning role. But is it possible to fulfil both roles? Evaluation as an investigative process often seeks to attribute responsibility and blame which may not be conducive to lesson-learning. In contrast, participatory and facilitated evaluations are often seen as a more appropriate style for lesson-learning. At best, the tension between accountability and lesson-learning functions can be a creative one. At worst, one objective dominates and it is impossible to fulfil both roles.

Some characteristics of accountability-oriented versus lesson-learning oriented evaluations

Characteristic	Accountability-oriented	Lesson-learning oriented
Terms of reference	Likely to be set by those external to the programme	Likely to be set by those directly involved in the programme
Team membership	Independent external team	Internal team of employees, or mixed team of employees and independent members
Emphasis in approach	Methodology of data collection and analysis emphasised – more objective	Process of reflection and reaching conclusions emphasised – more subjective
Style of management	More directive	More facilitative
Circulation of report	In public domain	Internal to organisation/ restricted

Most important is to be clear about the purpose of the evaluation, and to design it accordingly. Is it principally to fulfil an accountability function? If so, accountability to whom? Is it principally for lesson learning? If so, lesson learning by whom? Answering either of these questions means clearly identifying your key stakeholders for the evaluation, *at the outset,* and designing the evaluation process accordingly.

REFERENCES:

ALNAP, Borton, J., Robertson, K., Kawalek, J., Hammond, R., & Beck, T. (2002). *ALNAP Annual Review 2002: Humanitarian Action: Improving Performance through Improved Learning*. London: Active Learning Network on Accountability and Performance in Humanitarian Action.
http://www.alnap.org/publications/rha.htm

ALNAP, Houghton, R., Robertson, K., Borton, J., Carver, L., Beck, T., & Apthorpe, R. (2001). *ALNAP Annual Review 2001: Humanitarian action: Learning from Evaluation* (ALNAP annual review ; 2002). London: Active Learning Network on Accountability and Performance in Humanitarian Action.
http://www.alnap.org/publications/rha.htm

EVALUATION STAGES AND RESPONSIBILITIES

Stage	Tasks	Responsibility	Outputs
Planning and commissioning	Drafting, circulation and approval of concept note Selection and briefing of Steering Group	Evaluation Manager	Concept paper
	Drafting, circulation and approval of TOR Consultant selection	Evaluation Manager (+ Steering Group)	TOR Tender documents
Inception	Briefing consultants Initial research Drafting Inception report Circulation and approval of Inception Report	Evaluation Manager Consultants Consultants Evaluation Manager and Steering Group	Inception report
Preparation and research	Project/programme research Interviews (UK and email) Planning country visits and dissemination strategy	Consultants Consultants Consultants, Evaluation Manager and Steering Group	
Country visits	Preparatory visit: Planning activities Draft and circulate visit report	Consultants (implementation)	Visit reports
	Main visit: Evaluation research Workshop(s) Draft and circulate visit report	Evaluation Manager (Quality of process)	In-country workshops
Reporting	Draft report and evaluation summary Circulate report for comment Edit and revise report Circulation and comment on revised report Submission to Development Committee Final amendments to report	Consultants Evaluation Manager and Steering Group Consultants & editor Evaluation Manager and Steering Group Evaluation Manager Evaluation Manager and Consultants	Draft Report Revised Report Final Report
Dissemination	Publication and distribution of report and evaluation summary Workshops	Evaluation Manager Consultants, Steering Group and Evaluation Manager	Published Report Evaluation summary Workshop

Source: Evaluation guidelines DFID, Evaluation Department

DEFINING ACCOUNTABILITY

Peter Raynard has defined the 'process' of accountability in terms of four inter-connecting stages:

(1)   agreement of clear roles & responsibilities of the organisation (and its staff)
(2)   taking action for which an organisation is responsible
(3)   reporting on and accounting for those actions
(4)   responding to and complying with agreed standards of performance and the views and needs of stakeholders

A number of challenges emerge when applying this model to the international aid sector. First, there is usually a lack of clearly defined responsibility between different actors. Second, there are often a large number of stakeholders who may have conflicting needs and agendas. And third, some of the most important stakeholders – the conflict-affected – may also be the weakest. Finally, the objectives of humanitarian programmes are often unclear or implicit, and may change which can further hinder attempts to establish accountability.

There are a number of different types of accountability:

(1) Strategic accountability ie to overall mandate & objectives
(2) Managerial accountability, eg within an agency re using resources to achieve overall objectives, to the board
(3) Financial accountability eg to funders re effectiveness & efficiency of resources
(4) Contractual accountability eg to donor
(5) Relational accountability eg to other agencies
(6) Legal accountability eg to Charity Law, human rights law

Being clear about which type of accountability is most relevant in an evaluation will help to identify the key stakeholders.

REFERENCE:

Raynard, P., 2000, 'Mapping Accountability in Humanitarian Assistance', Report presented to ALNAP at the bi-annual meeting in April 2000

THE MONITORING, EVALUATION, AUDIT 'SCALE'

What distinguishes monitoring, evaluation and auditor? There are no hard and fast, mutually exclusive definitions for these terms. They are increasingly dynamic. They are applied in different ways by different organisations and in different times. Increasingly, they overlap, share and adapt similar aims, methodologies and approaches.

Auditors, for example, increasingly include conclusions and recommendations in their reports that are in some cases similar in scope to those commonly addressed in evaluations. Auditors increasingly go beyond auditing compliance and enquiring into the substantive use of resources (examining 'results' and 'outcomes').

Yet, identical they are not. Useful de-limiting lines can be identified. A very rough, relative scale, as follows, might be helpful.

Approach	Activity
Facilitated/internal	Monitoring
Routine/lessons learned	
Systematic/including accountability	Evaluation
Control/Compliance	
Investigation/External	Audit/Inspection

The main issue is not what distinguishes the methods or approaches. Overlaps will occur, as in most project-cycle management activities (e.g. assessment and monitoring). The important challenge is how to best manage and apply evaluation, in order to improve learning and accountability.

WHY DO A REAL TIME EVALUATION?

UNHCR which has made more use of Real Time Evaluations (RTEs) than any other organisation sums up the three key advantages of RTEs as *timeliness, perspective, and interactivity* (Jamal and Crisp, 2002, pp 1-2)

Timeliness: A real-time evaluation is undertaken in the early phase of an operation, at a time when key operational and policy decisions are being taken. They can flag up important issues that have been overlooked, either by the field or headquarters in the heat of an emergency response.

Interactivity: The RTE format is interactive in that sense that real-time evaluations take place in time for them to influence emergency response decisions. They are also engaged in a sustained dialogue with staff, both in the field and at headquarters, and can provide a channel for communication between field staff and headquarters that bypasses the normal bureaucratic layers.

Perspective: A real-time evaluator is able to approach an emergency from a number of different angles. He or she should be a repository of knowledge on lessons from past emergency evaluations, and should incorporate such knowledge into the evaluation process and outputs. Unlike staff in the field or at headquarter the evaluator does not have to focus on solving immediate problem, but has the scope to analysis the critical issues - a space for analysis that the staff directly responsible will only have in retrospect. The evaluator will also have a broader perspective than most other actors as the evaluator will be able to talk to staff at all levels, in different countries, to the affected population, as well as to partners, and government officials

All organisations can benefit from RTEs in that the breath of information and the analysis provided in a good RTE can clarify issues both in the field and headquarter far faster than might otherwise happen.

Real time evaluations:

- Bring in an external perspective, analytical capacity and knowledge at a key point in the response.

- Reduce the risks that early operational choices bring about critical problems in the longer term.

- Enable programming to be influenced as it happens and for agencies to make key changes at an intermediate point in programming.

- Can facilitate improved communication and understanding between HQ and the field.

ADVANTAGES FOR THE FIELD

- RTEs can provide the agency with a chance to send messages to headquarters to reinforce your traditional reporting, whether it is about a

strategic programming issue or simply to let headquarters know that work is being done well.

- An RTE team can conduct an after-action review for the operational team.

- An RTE team will have more time to consult with beneficiaries than the operational team will.

- An RTE can answer specific research questions, such as what approaches have worked in the past for any particular problem that is being faced, or what approaches are working well in other areas.

- An RTE can provide some analytical insight into the issues that the operational teams face, generating insights which the teams do not have the time to develop because they are too busy dealing with the day to day problems of the response.

ADVANTAGES FOR HEADQUARTER

- RTEs can provide HQ with a quick overview of the whole programme that is far richer than that presented in progress reports as RTEs will normally include the perspectives of beneficiaries, partners, and the independent evaluation team.

- An RTE can reduce the risk of problems arising later because of the need to take quick decisions with poor information in the early stages of the emergency response.

- An RTE can answer specific research questions for HQ, such as what approaches have worked in the past for any particular problem being faced, or what approaches are working well in other areas.

- An RTE can provide some analytical insight into the issues that a particular programme faces, an insight which HQ staff does not have the time to develop

LIST OF PREVIOUS RTEs

Many parts of the humanitarian system have used RTEs over the last five years. A list of the better known RTEs is as follows:

- The Inter-Agency Standing Committee piloting of three inter-agency RTEs in Mozambique and Pakistan has led to some key lessons, both for ongoing humanitarian reform processes, and for improving RTE quality.

- UNHCR have worked with RTE-like mechanisms since 1992, in 10 countries. They continue to use and promote RTEs, for example a 2007 five-country review of the cluster approach and IDPs, as well as making available guidance material which has proved invaluable for other agencies piloting RTEs.

- UNICEF has played an important role in bringing different actors together around RTEs, generating discussion and debate, reviewing past experience, and pushing for clarity on the minimum requirements for an RTE.

- WFP has commissioned two RTEs – in Southern Africa and Asia, as well as promoted lesson learning.

- FAO has carried out RTEs related to assessment of avian influenza and tsunami programming.

- ECHO has commissioned RTEs in the food aid and health sectors, and has experimented with longer-term approaches.

- IFRC has commissioned RTEs in Southern Africa and Asia.

- OXFAM has started the process of institutionalising RTEs as a part of their response framework, with five RTEs completed and rich comparative experiences.

- CARE International has carried out RTEs in Darfur and Iraq.

- HAP has carried out RTEs as part of the process of piloting Humanitarian Accountability Frameworks.

- The Groupe URD Iterative Evaluation Process with Mini-Seminar, while not termed an RTE, uses a similar methodology.

- The UK's Disasters Emergency Committee has used monitoring missions which have much in common with RTEs.

REFERENCES AND RESOURCES

Jamal, A., & Crisp, J. (2002). *Real-time humanitarian evaluations: some frequently asked questions (EPAU /2002/05)*. Geneva: United Nations High Commissioner for Refugees, Evaluation and Policy Unit. URL: http://www.unhcr.org/research/RESEARCH/3ce372204.pdf

JOINT EVALUATIONS[1]

DEFINING AND CATEGORISING JOINT EVALUATIONS

The ALNAP meta-evaluation on joint evaluations of humanitarian action defines them as 'an evaluation carried out by two or more agencies, evaluating the work of two or more agencies' (Beck and Buchanan-Smith, 2008).

At a workshop on joint evaluations during the 20th ALNAP Biannual meeting in December 2006, Niels Dabelstein provided a useful categorisation of joint evaluations (based on the DAC categorization) as:

• *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 and the Tsunami evaluation.

• *qualified joint evaluations:* participation is open only to those who qualify, through
membership of a certain grouping (e.g. DAC, Nordic governments) or through active participation in the activity (e.g. 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.

The ALNAP meta-evaluation went a step further in developing a typology for joint evaluations in the humanitarian sector, presented in Table 1 below:

Table 1: A typology for joint evaluations in the humanitarian sector

		FOCUS OR SCOPE OF EVALUATION				
HOW ACTORS		*Program focus*	*Institutional*	*Sectoral or thematic focus*	*Multi-sectoral focus, related to a particular humanitarian crisis (usually bounded geographically)*	*Global eg global policy*

---

[1] The material is drawn from the ALNAP meta-evaluation of joint evaluations (Beck and Buchanan-Smith, 2008), and from the report of a workshop on joint evaluations, held at the 20th ALNAP Biannual Meeting in Rome, in December 2006.

'Partnership': donor & recipient agencies evaluate together as equal partners					
'Club' (or qualified): agencies with similar characteristics coming together					
'Hybrid multi-partner': disparate actors coming together playing variable roles (eg active/ passive)					
'System-wide': open to all actors in the system					

Source: Beck and Buchanan-Smith, 2008: 97

A SHORT HISTORY OF JOINT EVALUATIONS IN THE HUMANITARIAN SECTOR

Joint evaluations have a longer history in the development than in the humanitarian sector; they were originally pioneered by donor governments through the OECD/ DAC. Interest in joint evaluations in the humanitarian sector is more recent, and part of the growing trend towards 'jointness' in the aid world. Once again it was donor governments that took the lead in joint evaluation of humanitarian action (for example the Dutch and British governments evaluating WFP's programme in Sudan in 1999; five donors evaluating humanitarian assistance and reconstruction in Afghanistan in 2005). The Nordic governments have played a key leadership role.

During the last decade, however, there has been a growing trend of UN agencies carrying out joint evaluations of their combined work in a particular country or region (for example WFP and UNHCR coming together to evaluate 'protracted relief and recovery operations' in Sudan in 2001, and again to evaluate pilot food-distribution projects in five countries in 2005/06). For many years the Disasters Emergency Committee in the UK was the main pioneer of joint evaluations in the NGO sector, evaluating the use of appeal funds raised collectively by the member NGOs. (The DEC joint evaluations have now been replaced by an accountability framework that emphasizes single agency evaluations instead). The baton for joint evaluations has since been picked up by a group of NGOs that have come together under the Emergency Capacity Building project.

There have been only two system-wide joint evaluations of humanitarian action. The first was the seminal multi-agency Rwanda evaluation in 1996. Ten years passed before another system-wide evaluation was launched, by the Tsunami Evaluation Commission (TEC) in 2006.

In short, there appears to be a growing momentum behind the joint evaluation approach, engaging donors, UN agencies and NGOs. The most obvious gap is recipient governments and national NGOs. The second phase 'Linking Relief, Rehabilitation and Development' evaluation of the TEC, in 2008/09, attempted to break this mould by engaging the governments of Sri Lanka, Indonesia and the Maldives as key partners in the evaluation process. Their record of engagement was mixed, and offers some useful learnings for the future[2].

REASONS FOR DOING JOINT EVALUATIONS

The reasons for doing joint evaluations fall into five broad categories[3]:

(1) overarching policy reasons: the benefit that a joint evaluation provides in 'seeing the big picture', and evaluating the programme or range of interventions against this big picture

(2) evaluation strategy motives: for example to do with increased credibility and legitimacy that joint evaluations can provide, which can be useful in advocating for change, especially if there are sensitive issues to be covered

(3) learning motives: so partners understand each other's approaches and exchange good practice

(4) managerial, administrative and financial motives: for example sharing funds if evaluation resources are scarce, or redressing a lack of sufficient evaluation capacity within an agency

(5) developmental motives: for example reducing transaction costs for developing countries, and building ownership and participation of developing countries (this has been a stronger motivation in the development than humanitarian sector so far).

The reasons against doing joint evaluations are often to do with the complexity of the process and of the subject, which can be time-consuming and expensive, not least in transaction costs, and involving complicated management structures.

The stated purpose of joint evaluations of humanitarian action is usually both learning and accountability, but in practice learning wins out. In a survey of participating agencies in the TEC, the majority cited 'learning' (for improved performance) as their principal reason for involvement in the TEC; 'accountability' ranked fifth as a reason for involvement. By the very nature of the exercise, however, peer accountability is a key part of joint evaluations, requiring a high level of transparency. As it is more likely that joint evaluations will end up in the public domain compared with single agency evaluations, there can be an element of accountability. Other benefits of doing a joint evaluation may emerge during the process, for example building trust and social capital between agencies. This

---

[2] See Buchanan-Smith (2009) 'Reflections on the LRRD2 Joint Evaluation Process', Channel Research, September

[3] See DAC (2005) 'Joint Evaluations: Recent Experiences, Lessons Learned and Options for the Future', DAC Evaluation Network Working Paper: Paris: DAC

was the experience of some of the NGOs engaged in joint evaluations through the ECB project.

PLANNING AND MANAGING JOINT EVALUATIONS

More time is needed for joint evaluations than for single-agency evaluations, because of their broader scope and greater transaction costs, negotiating between multiple partners. Yet the planning time allocated to joint evaluations is frequently under-estimated. Indeed, the complex management structures and high transaction costs can be a barrier to NGO involvement.

A two-tier management system has become a popular approach and appears to have worked well, whereby there is a small management group and a larger steering group. Three tier systems have also been used successfully with an additional reference group in-country, for example in the inter-agency health evaluations (IHEs) that were first launched in 2003.

Some general learnings to guide the management of joint evaluations emerged during the 2006 ALNAP workshop:
(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).

However, the pool of sufficiently skilled evaluators – with the necessary technical, political and inter-personal skills – to do an effective joint evaluation appears to be small compared with the rising demand for joint evaluations.

THE OVERALL QUALITY OF JOINT EVALUATIONS

The ALNAP meta-evaluation concluded that joint evaluations are overall of higher quality than single agency evaluations, for example:

(i)   their terms of reference are generally clearer and more useable
(ii) consultation with local populations and beneficiaries is stronger
(iii) more attention is paid to international standards; and
(iv) the EHA criteria are more rigorously used.

However, striking gaps and weaknesses in the joint evaluations reviewed also emerged, especially inadequate attention to cross-cutting issues such as gender equality, protection and advocacy.

The meta-evaluation also concluded that the debate about whether joint evaluations replace or reduce the need for single agency evaluations is a distraction: they are very different exercises with different purposes. Joint evaluations can fulfil accountability purposes, for example to peers and through stronger consultation with beneficiaries, but this may be at a different level to the accountability needs of a single agency. Where they clearly complement single agency evaluations is by placing the response in the bigger picture/ wider context, exploring how agencies work together, and addressing wider policy issues.

Section two:
Designing the evaluation

When designing an evaluation there are a number of steps to be taken.

What kind of evaluation are you going to choose? If learning is the primary purpose of the evaluation, what kind of approach will foster learning by the staff and organisation involved?

This section guides the reader through these steps, and introduces a tool for stakeholder mapping. Drawing up a clear and do-able Terms of Reference (TOR) is a key part of evaluation design. A guide to TOR, and a sample 'good practice' example of TOR are provided, as well as some useful definitions of the difference between outputs, outcomes and impact. This section introduces the OECD/ DAC evaluation criteria that are normally used in EHA, and some alternative evaluation criteria.

This section provides a number of frameworks that are worth considering when designing an evaluation, ranging from the Red Cross & Red Crescent Code of Conduct to the principles of Good Humanitarian Donorship. It ends with a review of the advantages and disadvantages of using internal versus external evaluators.

EVALUATION TYPES

(These are not all mutually exclusive):

Cluster	An evaluation of a set of related activities, projects and/or programs.
Country Programme	Evaluation of one or more of the agency's portfolio of interventions and the assistance strategy behind them, in a particular country
Synthesis	A synthesis of the findings from a number of evaluations of individual projects or programmes.
Joint	An evaluation to which different donor agencies and/or partners participate.
Impact	Focuses on the impact of the aid, rather than on aid delivery. Usually, but not always, carried out some time after project completion.
Meta-evaluation	The term is used for evaluations designed to aggregate findings from a series of evaluations. It can also be used to denote the evaluation of an evaluation to judge its quality and/or assess the performance of the evaluators
Mid-term evaluation	Evaluation performed towards the middle of the period of implementation of the intervention.
Participatory	Evaluation method in which representatives of agencies and stakeholders (including beneficiaries) work together in designing, carrying out and interpreting an evaluation.
Process	An evaluation of the internal dynamics of implementing organizations, their policy instruments, their service delivery mechanisms, their management practices, and the linkages among these.
Programme	Evaluation of a set of interventions, marshaled to attain specific global, regional, country, or sector development objectives.
Project	Evaluation of an individual intervention designed to achieve specific objectives within specified resources and implementation schedules, often within the framework of a broader programme.
Policy	Examines the framework of understanding, beliefs and assumptions that make individual projects possible as

	well as desirable. Policy evaluations seek out the inherent tensions or contradictions in policy objectives through tools such as discourse analysis.
Real time	Evaluation of an ongoing operation to evaluate events as they unfold.
Self-evaluation	An evaluation by those who are entrusted with the design and delivery of an intervention.
Sector	Evaluation of a cluster of interventions in a sector within one country or across countries, all of which contribute to the achievement of a specific development goal.
Thematic	Evaluation of a selection of interventions, all of which address a specific priority that cuts across countries, regions, and sectors.

Source: DAC (2001) *Glossary of Evaluation and Results Based Management Terms* DCD/DAC/EV Paris: OECD http://www.oecd.org/dataoecd/29/21/2754804.pdf

EVALUATION ACTIVITY WITH A LEARNING-ORIENTATION

There are a range of approaches that can be adopted in an evaluation that is more learning-oriented. The following are some of the most common.

AFTER ACTION REVIEWS FOR LEARNING DURING/ AFTER

After Action Reviews (AARs) originated with the US Army in the 1970s. They are a simple and straightforward facilitated process to encourage reflection and learning, usually on a 'no attribution, no retribution' basis. . The key questions in an AAR are:

- What was the objective/ intent of the action?

- What went well?

- What could have gone better?

- What would we do differently next time?

(See later in the manual for a fuller description of AAR)

REFERENCE:

Sexton, R., & McConnan, I. (2003). *A Comparative Study of After Action Review (AAR) in the context of the Southern Africa Crisis*. London: ALNAP. Last viewed on 23 June 2009. URL: http://www.alnap.org/pdfs/other_studies/aar_case_study.pdf

APPRECIATIVE INQUIRY

Appreciative Inquiry originated in the field of organizational development. It is based upon building organizations around what works, with a constructive approach that asks positive questions. It adopts a solution-oriented rather than a problem solving approach.

'This approach to personal change and organization change is based on the assumption that questions and dialogue about strengths, successes, values, hopes, and dreams are themselves transformational. Appreciative Inquiry suggests that human organizing and change, at its best, is a relational process of inquiry, grounded in affirmation and appreciation.' (Whitney and Trosten Bloom, 2003: 1)

Appreciative Inquiry utilizes a cycle of four processes, designed to trigger the imagination as follows:

1) Discover: the identification of organizational processes that work well.

2) Dream: the envisioning of processes that would work well in the future.

3) Design: Planning and prioritizing processes that would work well – encourages innovation.

4) Destiny (or deliver): the implementation (execution) of the proposed design

In recent years there have been a number of examples of Appreciative Inquiry applied to evaluations.

REFERENCES:

Coghlan, A.T., Preskill, H., and Tzavaras Catsambas, T. (2003) 'An Overview of Appreciative Inquiry in Evaluation'
http://blogs.ubc.ca/evaluation/files/2009/02/appreciative20inquiry.pdf

> Whitney, D. and Trosten-Bloom, A. (2003) *The Power of Appreciative Inquiry: A Practice Guide to Positive Change* Berrett-Koehler Publishers.

MOST SIGNIFICANT CHANGE/ STORIES OF CHANGE APPROACHES

> The 'Most Significant Change' technique (MSC) captures change in the form of stories rather than pre-defined indicators.

> 'Essentially, the process involves the collection of significant change (SC) stories emanating from the field level, and the systematic selection of the most significant of these stories by panels of designated stakeholders or staff. The designated staff and stakeholders are initially involved by 'searching' for project impact. Once changes have been captured, various people sit down together, read the stories aloud and have regular and often in-depth discussions about the value of these reported changes. When the technique is implemented successfully, whole teams of people begin to focus their attention on program impact.' (Davies and Dart, 2004: 8)

> Thus, it is a participatory approach that builds from the field level upwards, around stories that capture what, when and why, and the reasons the event was important.

REFERENCES:

> Davies, R. and Dart, J., (2005) 'The 'Most Significant Change' (MSC) Technique'. A Guide to Its Use.
> http://www.mande.co.uk/docs/MSCGuide.pdf

STAKEHOLDER MAPPING

Identifying the key stakeholders should be a critical step in the planning of an evaluation. Here are some questions to ask yourself:

1) Who are the primary stakeholders for this evaluation?

2) How can you engage them, encourage their ownership, and ensure relevance of the evaluation to their needs?

3) Which other stakeholders do you need to take into account?

4) Which stakeholders' needs are you going to prioritise in order to make the evaluation do-able and to retain focus?

'The decision to evaluate should be the result of a considered and negotiated process involving the key stakeholders and especially the intended users' (Sandison, 2006)

One way of facilitating this process is to do some stakeholder mapping, distinguishing between the inner core of stakeholders who are to be fully engaged in the evaluation, the stakeholders who will just be consulted during the evaluation, and the stakeholders you wish to influence with the evaluation findings. These can be visually mapped as a series of concentric circles, as indicated below.

Stakeholders to be influenced by the evaluation

Stakeholders to be consulted

Primary stakeholders
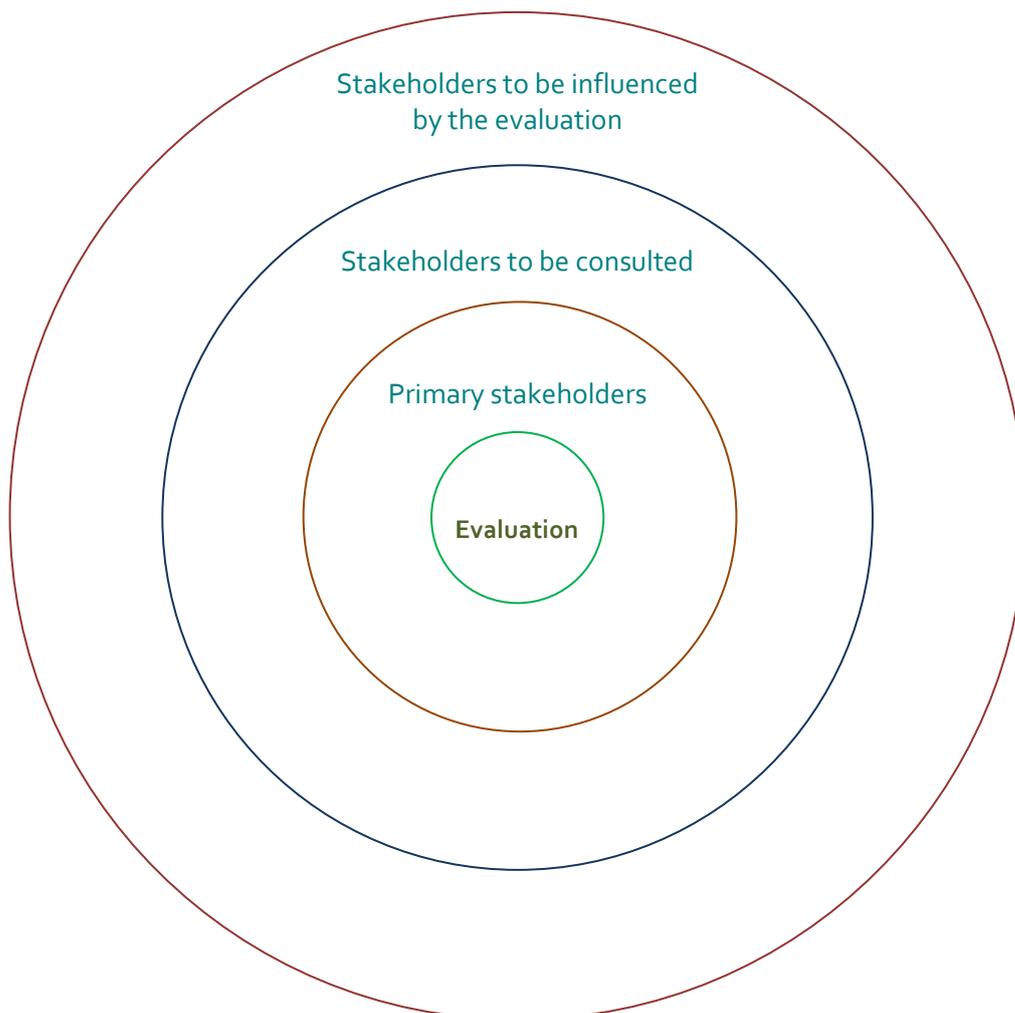
Evaluation

OUTPUTS, OUTCOMES & IMPACT

Here are some definitions to distinguish between outputs, outcomes and impact:

Outputs are defined deliverables provided by the actor or groups of actors evaluated. They are the result of the inputs received and activities developed by that actor. An output must be fully attributable to an actor or a group of similar actors.

Outcomes describe the use made of the outputs by the beneficiaries. An outcome is only partly attributable to an actor, even if ideally it should still be the predominant cause of the outcome.

Impact describes the longer term effects of the use of outputs by beneficiaries, or, in other words, the consequences of the outcomes. The actors evaluated only contribute to the overall impact, which must be attributed to a variety of factors outside the actors' control.

Most evaluations focus on outputs and outcomes. Impact-oriented evaluations are still rare.
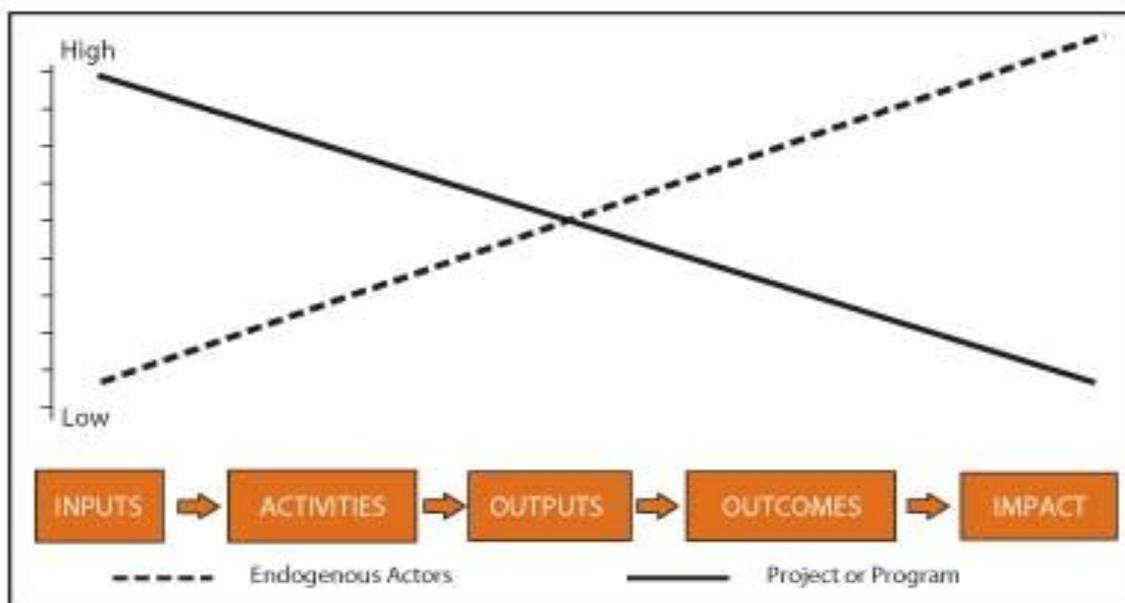
Figure 2. Relative Influence Along the Results Chain
(source: Smutylo 2001)

REFERENCES

Smutylo, T. (2001). *Crouching impact, hidden attribution: overcoming threats to learning in development programs. Paper presented at the Block Island Workshop on Across Portfolio Learning, 22-24 May 2001.* Last viewed on 9 November 2009. URL: http://www.idrc.ca/uploads/user-S/10905186681Crouching_Impact,_Hidden_Attribution.pd.pdf

FUNCTIONS OF TERMS OF REFERENCE

The TOR represent agreed expectations in terms of:

- Scope and parameters

- Process (including timing)

- Role of each key stakeholder

- Obligations of evaluation team, and of other stakeholders

- Key questions to be answered

The TOR provides a formal record of agreement as to what will be done. They are just as important for internal teams as for external teams (although TOR for external teams may require more detail on background context and on intended audiences and uses).

NB. The TOR are a critical tool for linking the evaluation's design with its intended use (but this is often overlooked in the planning phase!)

TERMS OF REFERENCE (TOR) CONTENT

DETAILED TERMS OF REFERENCE GENERALLY INCLUDE:

- The reasons for the evaluation and its objectives (why evaluate).

- A statement of the scope and specific issues to be addressed (what to evaluate – policy, programme, operation, issue).

- Objectives – the extent to which the evaluation is expected to provide accurate measures of impact and contribute to accountability should be carefully considered.

- The questions to be answered (criteria or focus) and tasks to be carried out, including, if possible, what information is to be collected and how.

- The locations to be visited (where); access to information and people.

- Which people are responsible for which tasks (who) – to what extent is the evaluation independent?

- A statement of the expected output & style of the report.

- A timetable (when to evaluate) indicating when various tasks will be completed as well as the due dates and recipients of any periodic reports or outlines. The TOR should specify that an initial report will be submitted in draft and provide time for corrections or changes to the draft once it has been reviewed.

- A budget indicating the costs associated with the evaluation.

- What happens after the evaluation (follow up, ownership).

OTHER ISSUES TO BE CONSIDERED:

- Need to consider translation of TOR for in-country use by evaluators.

- The evaluation team often draws up a detailed work plan of the evaluation, once the TOR has been decided on.

(Adapted from: Planning and Organising Useful Evaluations. UNHCR Inspection and Evaluation Service, January 1998)

SAMPLE TERMS OF REFERENCE (TOR)

The following is a sample TOR. It may be helpful as a model outline. Areas which could have been covered explicitly are, however: an indication of how the report will be used and provisions for follow-up action.

Disasters Emergency Committee TOR for the Independent Evaluation of Expenditure of DEC Mozambique Floods Appeal Funds

Background

Around 8th February 2000, Mozambique, one of the world's poorest countries, suffered its worst floods for half a century. Following the unusually heavy rains over the region and the tropical storms that accompanied cyclones Connie and Eline, hundreds of square miles around the Limpopo and Save river basins were left under water, with hundreds of thousands of people homeless and at risk.

The Mozambique Government estimated that up to 1 million people had been affected with hundreds feared dead. Overwhelmed by the devastation it appealed to the international community for assistance.

The Disasters Emergency Committee (DEC) is a UK charity which launches and co-ordinates a National Appeal in response to a major disaster overseas. It brings together a unique alliance of aid, corporate, public and broadcasting sector services to rally compassion and ensure that funds raised go to those amongst its fourteen UK member aid agencies best placed to bring effective and timely relief. DEC agencies have been present in Mozambique for many years working directly or supporting local partners.

The DEC 'Mozambique Floods Appeal' was launched on 2nd March 2000. To date, the appeal has generated some £20m pooled funds to be disbursed to the 11 DEC agencies participating in the appeal. These funds are supporting activities in: search and rescue; water and sanitation; food aid; medicine and health care; clothing and household items; shelter and housing; seeds and tools etc *(see Appendix 2 for summary of agencies' activities and locations)*. There will be an initial nine-month expenditure period following the launch of this appeal, during which participating agencies will have access to a pre-determined share of pooled appeal funds *(see Appendix 1 for summary of disbursement shares)*. Participating agencies will submit a 'Final Declaration of Expenditure' reports in the tenth month following the launch.

DEC rules require an independent evaluation of the expenditure of appeal funds. This provides an important mechanism for DEC transparency and accountability to fundraising partners and the British public. Evaluation also enables DEC agencies to extend individual and collective learning on good-practice in response to humanitarian emergencies. The final report will be made public on completion of the evaluation.

Main purpose of the evaluation

To provide independent assessment of the effectiveness and impact of the DEC-funded responses following the Mozambique floods.

APPROPRIATENESS

Were assessments undertaken appropriate to identification of need? Were the actions undertaken appropriate in the context of the needs of the affected population and the

context in which the agencies were operating? Was sufficient attention given to the identification of clear objectives and activities that would ensure objectives were met?

Was the assistance appropriate in relation to the customs and practices of the affected population?

To what extent were potential and actual beneficiaries consulted as to their perceived needs and priorities? What was the level of beneficiary involvement in project design, implementation and monitoring? How effective and appropriate were these processes in ensuring relevant and timely project delivery in support of the most needy and vulnerable?

Was the assistance provided in a timely manner?

EFFICIENCY

Were resources used efficiently? For instance, were more expensive forms of response (such as air transport) used longer than was necessary? Would greater investment in preparedness measures have resulted in more effective and less costly responses?

IMPACT

What direct and indirect evidence is available that the action taken contributed to the reduction of mortality, morbidity and suffering and that the affected population was assisted in maintaining or resuming basic dignity and livelihoods? In the absence of much baseline data, it is suggested this might best be measured against the level of satisfaction of beneficiaries and their perception of appropriateness and effectiveness of the response.

What systems or indicators did agencies use to evaluate the effectiveness of their work?

COVERAGE

Was DEC assistance provided to all major population groups facing life-threatening situations?

What efforts were made to ensure that particular populations, vulnerable groups and areas were not overlooked?

Were beneficiaries correctly and fairly identified and targeted?

CONNECTEDNESS

Was the assistance provided in a way that took account of the longer-term context?

Did the assistance seek to strengthen the capacity of local agencies and personnel?

COHERENCE

What steps were taken by participating agencies to ensure their responses were co-ordinated with each other and with other humanitarian agencies?

Were other actions, such as advocacy work, undertaken by the member agencies to complement their immediate relief actions?

These criteria take into account 'standard' evaluation questions, and also reflect the DEC's Objective, the NGO/Red Cross Code of Conduct and those disaster response objectives of DEC member agencies that are broadly shared. Thus, objectives such as achieving a co-ordinated response, ensuring that relief activities take account of longer-term considerations and that the capacity of local organisations and personnel is strengthened during the response, are explicitly included in the criteria.

Following the field visits the evaluation team should be in a position to comment on the adequacy of management, accounting, monitoring and reporting processes of the DEC agencies and their field-level partners. They should also be able to comment on the key constraints that affected the DEC supported programs, and how the different DEC agencies dealt with them.

Specific Issues for Consideration

- What was the added value of DEC appeal funds in the context of the overall humanitarian response? Did DEC funds facilitate a quick response?

- Was gender considered in the agencies' emergency assessments? Did relief provision include special components for women and, if so, were these systematically monitored?

- Were the special needs of acute vulnerable groups (e.g. children/elderly/disabled etc) considered in the agencies' emergency assessments and were they consulted in the same way as other groups? Did relief provision include special components for them and if so were these appropriate and systematically monitored?

- Did the response of the DEC agencies strengthen and complement the response of local organisations and coping mechanisms, or hinder them?

- What was the level of co-operation in the field? Could more have been done to help improve the effectiveness of DEC agencies' responses in terms of co-ordination, joint-logistics, communications packages, and information flows between the key relief players?

- Was there appropriate geographical coverage within the affected region?

- To what extent did responses reflect lessons-learned from previous flood disasters?

- To what extent did the DEC agencies' limited involvement in 'search & rescue' and the difficulties around transport affect the impact of the DEC funded response? The evaluation team is requested to dedicate one section of the final evaluation report to this issue.

Method

Participating DEC agencies are required to submit the following material (in both hard copy and electronic format) to the Secretariat to assist the evaluation team's work:

- a summary chronology

- key documents on the agency's response to the emergency and their use of DEC funds - e.g. '48 Hour Plan of Action'; '4 Week Plan of Action'.

- names, contact details and roles during the response of key agency and partner personnel in the head office and in the relevant field offices.

- List of indicators used by the agencies to monitor and evaluate their DEC funded activities.

The Secretariat will prepare a package of materials on each participating agency to be given to the evaluation team, as well as minutes of appeal related decision-making meetings – e.g. decision to appeal; decision to extend/reduce the period of joint action; decisions affecting DEC rules governing appeals.

The evaluation team will begin with a review of available documentation.

The evaluation team will be responsible for ensuring appropriate data-collection is undertaken in the field following their appointment, so that key information, that may no longer be available in the later stage of the DEC funded response, is not lost to the evaluation process.

Following submission of DEC agencies' '10th Month Declaration of Expenditure' reports, member(s) of the evaluation team will visit the head office of each participating agency to undertake interviews and collect and review supplementary documentation. Evaluators should be allowed full access to relevant files. The schedule of the subsequent fieldwork will be discussed during these visits. Since certain operations will already have closed down by the time the evaluation proper is underway, it will be appropriate to undertake preliminary fieldwork during the expenditure period. The evaluation team's schedule, accommodation and transport arrangements will be finalised and communicated to the Secretariat and all agencies at least one week prior to any visit.

In the field the evaluation team will seek to spend a period with each agency that is roughly proportional to the share of DEC pooled funds received by each agency. During their work the evaluators will fill out the chronology of decisions and actions so as to understand the context and the level of information that was available to the agency in deciding on a particular action. During their time with each agency the team will interview key personnel remaining in-country (contacting others prior to the field visits or on their return) and undertake visits to selected project sites/areas. The field visit must include at least one DEC funded project for each participating agency. The evaluators will have to make extensive use of agency reports and their own preliminary data collection, where later site visits would prove pointless. It should be noted that in the case of agencies that are part of larger organisations UK assistance might not be distinguishable from that of global counterparts, nevertheless, every effort should be made to distinguish DEC funding.

As well as interviewing the agencies' project officers, key officials in co-ordinating agencies (e.g. UNICEF, OCHA, central and state governments), and partner agencies, a sample of beneficiaries will be selected and interviewed by the evaluators. These interviews will be conducted without agency personnel being present, using interpreters (where necessary) hired directly by the evaluation team. The beneficiaries will be questioned on their views of the assistance provided, the way they were selected and their overall views of the agency. Interviews with individuals may be complemented by discussions with groups of beneficiaries. So as to assess the agency's targeting and beneficiary selection methods the evaluation team will also interview a selection of potential beneficiaries who did not receive assistance.

It is expected that the evaluation team will use gender-aware and participatory approaches to seek the views of beneficiaries and, where appropriate, non-beneficiaries. Inclusive techniques will be expected of the evaluators, to seek active participation in the evaluation by members of local emergency committees, staff of implementing partner agencies and member agencies, and representatives of local and central governments.

Agencies' '10th Month Declaration of Expenditure' reports will be examined to assess direct and indirect project costs, and, in conjunction with beneficiary/team assessment of direct and indirect benefits, and to compare the cost-effectiveness of different approaches.

The evaluation will be undertaken with due reference to the Red Cross/Red Crescent NGO Code of Conduct, which all agencies have signed. Reference should also be made to the Sphere Standards.

Before leaving the country, members of the team will indicate their broad findings to Country Representative and senior staff of each agency and note their comments.

A meeting should then be held in London to disseminate a draft report of the evaluation. The report should be circulated one week prior to the workshop to allow for preliminary review by agencies and their partners, and followed by a two-week formal agency comment period.

The evaluation team should allow for a follow-up workshop in-country within a month of the release of the final evaluation report. The aim of this workshop will be to discuss the evaluation recommendations and major lessons of the Mozambique floods, and how agencies might seek to implement. It is suggested that participants include UK, regional, and in-country representatives from the agencies and their implementing partners, and other key stakeholders as appropriate.

The Report

The evaluation report should consist of:

- executive summary and recommendations (not more than six pages)

- main text, to include index, emergency context, evaluation methodology, appeal management, commentary and analysis addressing evaluation purpose and outputs to include a section dedicated to the issue of particular lesson-learning focus, conclusions (not more than thirty pages)

- appendices, to include evaluation terms of reference, maps, sample framework, summary of agency activities, sub-team report(s), end notes (where appropriate) and bibliography. (All material collected in the undertaking of the evaluation process should be lodged with the Secretariat prior to termination of the contract)

EVALUATION TEAM AND TIMEFRAME

It is anticipated there will be a core team of at least three people, with others drawn in as necessary. The Team Leader should have a relevant skill and a proven background in emergency evaluations. The appropriate balance of professional and analytical skills amongst the remaining team members should be determined following a preliminary examination of agency activities. It is likely, however, that sector expertise in areas such as water and sanitation, public health and shelter will be required. At least one person from the region should be included in the team that makes the field visits.

All team members should be gender aware, and a reasonable gender balance within field teams is desirable.

Consultants or independent evaluation teams short-listed in the tendering process should seek DEC approval for any proposed changes to the composition of the team originally submitted.

The evaluation timeframe should allow for the circulation of a first draft by early March 2001, followed by presentation of the draft by the evaluation consultant(s) to member agencies a week later. A formal comment period, of at least two weeks, for participating agencies and their partners will then follow. The completion date for the Final Evaluation Report will be 15th April 2001, the consultants having addressed agencies' comments as appropriate.

Tenders and Evaluation Management

Tenders should be submitted to the DEC Secretariat by the closing date of 15th May 2000. A maximum 5 page summary should be submitted with appendices of team member CVs

(each CV a maximum of 3 pages) and an indication of availability. The DEC may wish to see substantive pieces work or to take up references of short-listed consultants.

The final decision on tenders will be taken by the DEC Executive Committee, following short-listing and interviews. Key factors will include:

Provisional framework, methodology, team balance, local experiences, distinctive competencies, timeframe and budget, an appreciation of key constraints and comments on the above terms of reference.

Professionalism of the bid, team experience (professional and analytical), degree of parity with the terms of reference, likelihood of achieving the DEC timetable, and realism not just competitiveness in the cost submission.

Tenders will be accepted from "freelance" teams as well as from company, PVO or academic teams. Tenders are particularly welcome from regional teams.

Administration and overall co-ordination, including monitoring progress, lies with the DEC Secretariat. The evaluation Team Leader must, from the commencement of the evaluation, submit a weekly report on actual against anticipated progress. The Steering Committee (DEC Operations Sub-Committee) will via the Secretariat undertake to respond to weekly submissions as necessary. In addition, the Team Leader should alert the Secretariat immediately if serious problems or delays are encountered. Approval for any significant changes to the evaluation timetable will be referred to the Steering Committee.

DEFINITIONS OF THE MAIN EVALUATION CRITERIA

Criterion	Definition	Source and comments
Effectiveness	Effectiveness measures the extent to which an activity achieves its purpose, or whether this can be expected to happen on the basis of the outputs. Implicit within the criterion of effectiveness is timeliness.	One of the five original OECD/DAC criteria.
Efficiency	Efficiency measures the outputs – qualitative and quantitative – achieved as a result of inputs. This generally requires comparing alternative approaches to achieving an output, to see whether the most efficient approach has been used.	One of the five original OECD/DAC criteria.
Impact	Impact looks at the wider effects of the project – social, economic, technical, environmental – on individuals, gender- and age-groups, communities and institutions. Impacts can be intended and unintended, positive and negative, macro (sector) and micro (household).	One of the five original OECD/DAC criteria.
Relevance/ Appropriateness	Relevance is concerned with assessing whether the project is in line with local needs and priorities (as well as donor policy). Appropriateness is the tailoring of humanitarian activities to local needs, increasing ownership, accountability and cost-effectiveness accordingly.	Relevance is one of the five original OECD/DAC criteria. Appropriateness was introduced as an alternative in the 1997 OECD/DAC guidelines
Sustainability/ Connectedness	Sustainability looks at whether the benefits of an activity are likely to continue after donor funding has been withdrawn. Connectedness look at the extent to which activities of a short-term emergency nature take longer-term and	Sustainability was one of the five original OECD/DAC criteria. The alternative definition of Connectedness was added in the 1999 OECD/DAC guide.

	interconnected problems into account	
Coverage	The need to reach major population groups facing life-threatening suffering wherever they are.	Added in the 1999 OECD/DAC guide. Reflected concerns that some humanitarian operations were not inclusive.
Coherence	The need to assess security, developmental, trade and military policies as well as humanitarian policies, to ensure that there is consistency and, in particular, that all policies take into account humanitarian and human-rights considerations	Added in the 1999 OECD/DAC guide. Coherence can be analysed purely within humanitarian operations or between humanitarian operations and the wider activities of the same actors.
Coordination	The extent to which the interventions of different actors are harmonised with each other, to promote synergy, and to avoid gaps, duplication, and resource conflicts.	Added as a potential sub-criterion in the 1999 OECD/DAC guide. Now one of the eight standard criteria used by DFID. There is so agreed standard short definition for coordination.
Protection	The extent to which the interventions increases or reduces the risk that members of the affected population will be murdered, raped, physically abused, or suffer other violations of their human rights.	Added as a potential sub-criterion in the 1999 OECD/DAC guide. This is rarely used as evaluation criteria other than as a cross-cutting theme
Cross-cutting themes	Cross-cutting themes are often included as additional criteria that the evaluators should pay attention to. There are a large number of potential cross-cutting themes including: • Gender • Protection • The natural environment • Participation • Resilience	Typically only a few cross-cutting themes will be included in the terms of reference of any one evaluation

- Human Resources

- HIV/AIDS

- The use of local resources

- Partnership

- Power relations

- Poverty reduction

- Accountability

- Management issues

- Corruption

THE HISTORY OF THE STANDARD OECD/DAC EVALUATION CRITERIA

In 1991 The Development Assistance Committee of the Organisation for Economic Cooperation and Development (OECD/DAC) defined evaluation as follows:

> *An evaluation is an assessment, as systematic and objective as possible, of an on-going or completed project, programme or policy, its design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, developmental efficiency, effectiveness, impact and sustainability. An evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision-making process of both recipients and donors.* (OECD Development Assistance Committee, 1991)

This definition encapsulates what are now the standard criteria for evaluating overseas development assistance.

- Efficiency

- Effectiveness

- Impact

- Sustainability

- Relevance

THE 1999 REVISION FOR EVALUATING HUMANITARIAN ACTION

In 1997 the OECD/DAC sponsored a project to develop a modified set of criteria more appropriate for the particular challenges of evaluating humanitarian action in complex emergencies. This study culminated in an ODI Good Practice Review (Hallam, 1998), and a guidance note (OECD Development Assistance Committee, 1999). These proposed a modified set of these criteria for such evaluations. The changes proposed were as follows.

- Efficiency – unchanged

- Effectiveness - to specifically include Timeliness or Timeliness addressed separately as a separate criterion

- Impact - unchanged

- Sustainability replaced by Connectedness (dealing with whether the emergency response takes the long-term situation into account)

- Relevance (of the overall goal and purpose) to be supplemented with Appropriateness (of activities and inputs)

The following additional criteria were proposed.

- Coverage of the response

- Coherence (of policies with humanitarian action)

- Coordination

- Protection

Most evaluations of humanitarian action since the late 90s have been based on a subset of either the DAC criteria or of the modified criteria. These criteria are the mainstream criteria for the evaluation of humanitarian action.

Current best practice in the evaluation of humanitarian action is set out in the ALNAP guide on evaluating humanitarian action using the OECD-DAC criteria (Beck, 2006). However Beck deals with coordination under the heading of effectiveness, and includes protection as a cross-cutting issue.

REFERENCES

Beck, T. (2006). Evaluating humanitarian action using the OECD-DAC criteria. London: ALNAP. Last viewed on 8 June, 2008 at: www.odi.org.uk/alnap/publications/eha_dac/pdfs/eha_2006.pdf

DFID. (2009). Building the evidence to reduce poverty: The UK's policy on evaluation for international development. London: Department for International Development. Last viewed on 2 July 2009 at: http://www.oecd.org/dataoecd/39/54/43183649.pdf

Hallam, A. (1998). Evaluating Humanitarian Assistance Programmes in Complex Emergencies (RRN Good Practice Review 7). London: Overseas Development Institute. Last viewed on 8 June, 2008 at: http://www.odihpn.org/documents/gpr7.pdf

OECD/DAC. (1991). Principles for evaluation of development assistance [OCDE/GD(91)208]. Paris: Development Assistance Committee of the Organisation for Economic Cooperation and Development. Last viewed on 16 June 2008 at: http://www.oecd.org/dataoecd/31/12/2755284.pdf

OECD/DAC. (1999). Guidance for Evaluating Humanitarian Assistance in Complex Emergencies (1999). Paris: Organisation for Economic Co-operation and Development, Development Assistance Committee. Last viewed on 09 July 2009 at: http://www.oecd.org/dataoecd/9/50/2667294.pdf

ALTERNATIVE EVALUATION CRITERIA

NOTE

Before considering alternative evaluation criteria, remember that a strong advantage of the OECD/DAC criteria is that their widespread use makes it easier to compare evaluations of different responses, or of the same response at different times. Evaluators as well are likely to be very familiar with them and may be better able to apply them than other criteria. However, other criteria, based on standards or ideal types, may be more familiar to operational staff.

WHAT ARE THE FUNDAMENTAL EVALUATION CRITERIA?

These criteria are so familiar to evaluators of humanitarian action that it is sometimes assumed that these are fundamental criteria, but this is not the case. These are not fundamental criteria but are criteria that reflect recurring areas of weakness in development aid or humanitarian assistance. Many early overseas development assistance projects suffered from being unsustainable; having little relevance to the populations they were meant to assist; being ineffective at bringing about change; having no clear positive impact; or represented poor value for money.

Similarly, coverage, coherence, coordination, connectedness, timeliness, appropriateness, and protection, are all issues that have challenged the quality of different humanitarian aid responses. These criteria reflected the concerns in the humanitarian aid sector at the time that they were written, shortly after the Rwanda crisis. If the humanitarian action evaluation guidelines were rewritten today, they would doubtless include accountability as a criterion. They might also comment on the difference between the evaluation of complex emergencies and natural disasters.

The two fundamental criteria in evaluating anything are:

- quality, a measure of how good or bad something is; and

- value for money or return for effort, a measure of the benefits received against their cost.

Some may argue that value for money is one of the components of quality, but there is a key difference between measures of quality and measures of value for money. Measures of quality are intrinsic, and look at the intervention in its own terms. Value for money is extrinsic; it requires a valuation of the benefits received which we can only make by comparison to alternative means of spending the same amount of money in other ways.

Evaluations in the sector typically look at quality thorough the lens of common problem areas and at value for money through the "efficiency" criterion.

ALTERNATIVES TO THE TRADITIONAL CRITERIA

Realising that the traditional criteria are not fundamental, but merely measures of quality and value of money in areas identified as problematic allows us to

develop different evaluation criteria. Using the traditional OECD-DAC criteria means that meta-evaluation is much easier, and it is easier to synthesise the evaluation. There are three basic approaches:

1. Developing criteria around some generally agreed definition of quality, such as those provided by codes of practice.

2. Evaluation against an ideal type.

3. Basing criteria around specific issues identified for a particular humanitarian response.

EVALUATING AROUND QUALITY CRITERIA

Humanitarian action has seen a very large number of quality initiatives since the mid 90's[4]. These initiatives include: the Sphere Project; People in Aid; the Humanitarian Accountability Project; the Compass Quality Initiative and many others. However, the very first such quality initiative was the *Code of conduct for the International Red Cross and Red Crescent Movement and non-governmental organizations (NGOs) in disaster relief* (Steering Committee for Humanitarian Response and ICRC, 1994).

The Code of Conduct was first agreed in 1994, having grown out of a 1991 proposal from the French Red Cross in response to the Balkans Crisis (Relief and Rehabilitation Network, 1994). The Code of Conduct has been very widely adopted by humanitarian agencies, and adherence to it is required by some agency collaborative mechanisms as well as by some donors.

> **The ten principles of the Code of Conduct**
>
> 1. Humanitarian imperative
>
> 2. Basis of need
>
> 3. No proselytising
>
> 4. Not foreign agents
>
> 5. Respect culture
>
> 6. Build on local capacities
>
> 7. Involve beneficiaries
>
> 8. Reduce vulnerability
>
> 9. Accountable both ways
>
> 10. Respect victims as human beings

The main advantage of the using the Code of Conduct for evaluation is that it provides a set of criteria which agencies have themselves subscribed to. Evaluations against these criteria effectively test compliance with the standards that agencies have themselves signed up to. The evaluation of the DEC 2003 Southern Africa Crisis Appeal was an example of this approach (Cosgrave et al., 2004).

Evaluation against the Code of Conduct may be particularly appropriate for a real-time humanitarian evaluation. Sometimes evaluators are asked to perform an evaluation against quality criteria as well as against the OECD-DAC criteria. However, such double working would not be suitable for the limited time available in a real time evaluation.

EVALUATION AGAINST AN IDEAL TYPE

---

[4] One aid agency, Oxfam, is said to be a member of no less than 23 different quality initiatives in the sector (Polastro, 2007).

Evaluation against codes of conduct and sphere concentrate more on the outcomes of the action of agencies. It is also possible to evaluate agency performance against an ideal type. This can be appropriate where an ideal type is defined in strategic plans or other documents. This may be appropriate for

Assessment matrix used by AusAid to evaluate their partners who responded to the 2005 Pakistan earthquake (simplified) (Crawford et al., 2006).

Performance Dimension	Performance criteria
Organisational capacity	Ability to deliver emergency assistance
	Quality of relationships with partners and beneficiaries
Planning Capacity	Quality of analysis and of initial strategy
	Quality of funding proposal
Implementation	Efficiency of emergency response
	Capacity for learning and accountability
	Effectiveness of emergency response
	Connectedness/Sustainability

donors looking at their partners.

EVALUATING AROUND SPECIFIC ISSUES

Another alternative set of criteria is offered by selecting a set of issues for a particular humanitarian response and framing these as hypotheses to be tested by the evaluators. This approach demands significant pre-evaluation work on the identification of suitable issues. It has the advantage the evaluators are answering questions which may be more directly relevant to the client.

This approach has the advantage that it may promote the direct utilisation of the evaluation results. This was the approach taken with the 2006 evaluation of WFP's Darfur operations where the evaluation set out to answer a series of hypotheses developed from the pre-evaluation studies (Cosgrave et al., 2006). However this approach is less suitable for a real time evaluation because developing of the need for some prior research to highlight possible hypotheses for analysis.

REFERENCES

Cosgrave, J., Oman, A., Fadel, A. J. A., & Babiker, S. A. (2006). *Full Report of the Evaluation of EMOP 10339.0/1: Assistance to populations affected by conflict in greater Darfur, West Sudan: A report from the Office of Evaluation (OEDE/2007/002).* Rome: World Food Programme. Last viewed

on 19 September 2008. URL: http://documents.wfp.org/stellent/groups/public/documents/reports/wfp1173 35.pdf

Hallam, A. (1998). *Evaluating Humanitarian Assistance Programmes in Complex Emergencies.* London: Overseas Development Institute

OECD/DAC. (1991). *Principles for evaluation of development assistance [OCDE/GD(91)208].* Paris: Development Assistance Committee of the Organisation for Economic Cooperation and Development. Last viewed on 16 June 2008. URL: www.oecd.org/dataoecd/secure/9/11/31779367.pdf

OECD/DAC. (1999). *Guidance for evaluating humanitarian assistance in complex emergencies.* Paris: Organisation for Economic Co-operation and Development, Development Assistance Committee. Last viewed on 16 June, 2008. URL: http://www.oecd.org/dataoecd/9/50/2667294.pdf

Relief and Rehabilitation, N. (1994). *Code of conduct for the International Red Cross and Red Crescent Movement and NGOs in disaster relief*. London: Relief and Rehabilitation Network, Overseas Development Institute

SCHR, & ICRC. (1994). *Code of conduct for the International Red Cross and Red Crescent Movement and non-governmental organizations (NGOs) in disaster relief*. Geneva: Disaster Policy Department, International Federation of Red Cross and Red Crescent Societies. Last viewed on 8 June, 2008. URL: http://www.icrc.org/web/eng/siteengo.nsf/htmlall/code-of-conduct-290296

THE RC CODE OF CONDUCT

Principles of Conduct for the International Red Cross and Red Crescent Movement and NGOs in Disaster Response Programmes

1 The humanitarian imperative comes first

The right to receive humanitarian assistance, and to offer it, is a fundamental humanitarian principle which should be enjoyed by all citizens of all countries. As members of the international community, we recognise our obligation to provide humanitarian assistance wherever it is needed. Hence the need for unimpeded access to affected populations is of fundamental importance in exercising that responsibility. The prime motivation of our response to disaster is to alleviate human suffering amongst those least able to withstand the stress caused by disaster. When we give humanitarian aid it is not a partisan or political act and should not be viewed as such.

2 Aid is given regardless of the race, creed or nationality of the recipients and without adverse distinction of any kind. Aid priorities are calculated on the basis of need alone

Wherever possible, we will base the provision of relief aid upon a thorough assessment of the needs of the disaster victims and the local capacities already in place to meet those needs. Within the entirety of our programmes, we will reflect considerations of proportionality. Human suffering must be alleviated whenever it is found; life is as precious in one part of a country as another. Thus, our provision of aid will reflect the degree of suffering it seeks to alleviate. In implementing this approach, we recognise the crucial role played by women in disaster-prone communities and will ensure that this role is supported, not diminished, by our aid programmes. The implementation of such a universal, impartial and independent policy, can only be effective if we and our partners have access to the necessary resources to provide for such equitable relief, and have equal access to all disaster victims.

3 Aid will not be used to further a particular political or religious standpoint

Humanitarian aid will be given according to the need of individuals, families and communities. Notwithstanding the right of NGHAs to espouse particular political or religious opinions, we affirm that assistance will not be dependent on the adherence of the recipients to those opinions. We will not tie the promise, delivery or distribution of assistance to the embracing or acceptance of a particular political or religious creed.

4 We shall endeavour not to act as instruments of government foreign policy

NGHAs are agencies which act independently from governments. We therefore formulate our own policies and implementation strategies and do not seek to implement the policy of any government, except in so far as it coincides with our own independent policy. We will never knowingly – or through negligence – allow ourselves, or our employees, to be used to gather information of a political, military or economically sensitive nature for governments or other bodies that may serve purposes other than those which are strictly humanitarian, nor will we act as instruments of foreign policy of donor governments. We will use the assistance we receive to respond to needs and this assistance should not be driven by the need to dispose of donor commodity surpluses, nor by the political interest of any particular donor. We value and promote the voluntary giving of labour and finances by concerned individuals to support our work and recognise the independence of action promoted by such voluntary motivation. In order to protect our independence we will seek to avoid dependence upon a single funding source.

5 We shall respect culture and custom

We will endeavour to respect the culture, structures and customs of the communities and countries we are working in.

6 We shall attempt to build disaster response on local capacities

All people and communities – even in disaster – possess capacities as well as vulnerabilities. Where possible, we will strengthen these capacities by employing local staff, purchasing local materials and trading with local companies. Where possible, we will work through local NGHAs as partners in planning and implementation, and cooperate with local government structures where appropriate. We will place a high priority on the proper co-ordination of our emergency responses. This is best done within the countries concerned by those most directly involved in the relief operations, and should include representatives of the relevant UN bodies.

7 Ways shall be found to involve programme beneficiaries in the management of relief aid

Disaster response assistance should never be imposed upon the beneficiaries. Effective relief and lasting rehabilitation can best be

achieved where the intended beneficiaries are involved in the design, management and implementation of the assistance programme. We will strive to achieve full community participation in our relief and rehabilitation programmes.

8 Relief aid must strive to reduce future vulnerabilities to disaster as well as meeting basic needs

All relief actions affect the prospects for long-term development, either in a positive or a negative fashion. Recognising this, we will strive to implement relief programmes which actively reduce the beneficiaries' vulnerability to future disasters and help create sustainable lifestyles. We will pay particular attention to environmental concerns in the design and management of relief programmes. We will also endeavour to minimise the negative impact of humanitarian assistance, seeking to avoid long-term beneficiary dependence upon external aid.

9 We hold ourselves accountable to both those we seek to assist and those from whom we accept resources

We often act as an institutional link in the partnership between those who wish to assist and those who need assistance during disasters. We therefore hold ourselves accountable to both constituencies. All our dealings with donors and beneficiaries shall reflect an attitude of openness and transparency. We recognise the need to report on our activities, both from a financial perspective and the perspective of effectiveness. We recognise the obligation to ensure appropriate monitoring of aid distributions and to carry out regular assessments of the impact of disaster assistance. We will also seek to report, in an open fashion, upon the impact of our work, and the factors limiting or enhancing that impact. Our programmes will be based upon high standards of professionalism and expertise in order to minimise the wasting of valuable resources.

10 In our information, publicity and advertising activities, we shall recognise disaster victims as dignified humans, not hopeless objects

Respect for the disaster victim as an equal partner in action should never be lost. In our public information we shall portray an objective image of the disaster situation where the capacities and aspirations of disaster victims are highlighted, and not just their vulnerabilities and fears. While we will cooperate with the media in order to enhance public response, we will not allow external or internal demands for publicity to take precedence over the principle of maximising overall

relief assistance. We will avoid competing with other disaster response agencies for media coverage in situations where such coverage may be to the detriment of the service provided to the beneficiaries or to the security of our staff or the beneficiaries.

The Working Environment

Having agreed unilaterally to strive to abide by the Code laid out above, we present below some indicative guidelines which describe the working environment we would like to see created by donor governments, host governments and the inter-governmental organisations – principally the agencies of the United Nations – in order to facilitate the effective participation of NGHAs in disaster response. These guidelines are presented for guidance. They are not legally binding, nor do we expect governments and IGOs to indicate their acceptance of the guidelines through the signature of any document, although this may be a goal to work to in the future. They are presented in a spirit of openness and cooperation so that our partners will become aware of the ideal relationship we would seek with them.

Annex I: Recommendations to the governments of disaster-affected countries

1 Governments should recognise and respect the independent, humanitarian and impartial actions of NGHAs

NGHAs are independent bodies. This independence and impartiality should be respected by host governments.

2 Host governments should facilitate rapid access to disaster victims for NGHAs

If NGHAs are to act in full compliance with their humanitarian principles, they should be granted rapid and impartial access to disaster victims, for the purpose of delivering humanitarian assistance. It is the duty of the host government, as part of the exercising of sovereign responsibility, not to block such assistance, and to accept the impartial and apolitical action of NGHAs. Host governments should facilitate the rapid entry of relief staff, particularly by waiving requirements for transit, entry and exit visas, or arranging that these are rapidly granted. Governments should grant over-flight permission and landing rights for aircraft transporting international relief supplies and personnel, for the duration of the emergency relief phase.

3 Governments should facilitate the timely flow of relief goods and information during disasters

Relief supplies and equipment are brought into a country solely for the purpose of alleviating human suffering, not for commercial benefit or gain. Such supplies should normally be allowed free and unrestricted passage and should not be subject to requirements for consular certificates of origin or invoices, import and/or export licences or other restrictions, or to importation taxation, landing fees or port charges.

The temporary importation of necessary relief equipment, including vehicles, light aircraft and telecommunications equipment, should be facilitated by the receiving host government through the temporary waiving of licence or registration restrictions. Equally, governments should not restrict the re-exportation of relief equipment at the end of a relief operation.

To facilitate disaster communications, host governments are encouraged to designate certain radio frequencies, which relief organisations may use in-country and for international communications for the purpose of disaster communications, and to make such frequencies known to the disaster response community prior to the disaster. They should authorise relief personnel to utilise all means of communication required for their relief operations.

4 Governments should seek to provide a coordinated disaster information and planning service

The overall planning and coordination of relief efforts is ultimately the responsibility of the host government. Planning and coordination can be greatly enhanced if NGHAs are provided with information on relief needs and government systems for planning and implementing relief efforts as well as information on potential security risks they may encounter. Governments are urged to provide such information to NGHAs. To facilitate effective coordination and the efficient utilisation of relief efforts, host governments are urged to designate, prior to disaster, a single point-of-contact for incoming NGHAs to liaise with the national authorities.

5 Disaster relief in the event of armed conflict

In the event of armed conflict, relief actions are governed by the relevant provisions of international humanitarian law.

Annex II: Recommendations to donor governments

1 Donor governments should recognise and respect the independent, humanitarian and impartial actions of NGHAs

NGHAs are independent bodies whose independence and impartiality should be respected by donor governments. Donor governments should not use NGHAs to further any political or ideological aim.

2 Donor governments should provide funding with a guarantee of operational independence

NGHAs accept funding and material assistance from donor governments in the same spirit as they render it to disaster victims; one

of humanity and independence of action. The implementation of relief actions is ultimately the responsibility of the NGHA and will be carried out according to the policies of that NGHA.

3 Donor governments should use their good offices to assist NGHAs in obtaining access to disaster victims

Donor governments should recognise the importance of accepting a level of responsibility for the security and freedom of access of NGHA staff to disaster sites. They should be prepared to exercise diplomacy with host governments on such issues if necessary.

Annex III: Recommendations to intergovernmental organisations

1 IGOs should recognise NGHAs, local and foreign, as valuable partners

NGHAs are willing to work with UN and other inter-governmental agencies to effect better disaster response. They do so in a spirit of partnership which respects the integrity and independence of all partners. Inter-governmental agencies must respect the independence and impartiality of the NGHAs. NGHAs should be consulted by UN agencies in the preparation of relief plans.

2 IGOs should assist host governments in providing an overall coordinating framework for international and local disaster relief

NGHAs do not usually have the mandate to provide the overall coordinating framework for disasters which require an international response. This responsibility falls to the host government and the relevant United Nations authorities. They are urged to provide this service in a timely and effective manner to serve the affected state and the national and international disaster response community. In any case, NGHAs should make all efforts to ensure the effective coordination of their own services. In the event of armed conflict, relief actions are governed by the relevant provisions of international humanitarian law.

3 IGOs should extend security protection provided for UN organisations to NGHAs

Where security services are provided for inter-governmental organisations, this service should be extended to their operational NGHA partners where it is so requested.

4 IGOs should provide NGHAs with the same access to relevant information as is granted to UN organisations

IGOs are urged to share all information, pertinent to the implementation of effective disaster response, with their operational

NGHA partners.

THE RC CODE OF CONDUCT AND DAC CRITERIA

The following is an ad hoc linking of the RC Code of Conduct principles with standard DAC evaluation criteria.

PRINCIPLES OF CONDUCT FOR THE INTERNATIONAL RED CROSS AND RED CRESCENT MOVEMENT AND NGOS IN DISASTER RESPONSE PROGRAMMES

1 The humanitarian imperative comes first (Humanitarian Imperative: Relevance)

2 Aid is given regardless of the race, creed or nationality of the recipients and without adverse distinction of any kind. Aid priorities are calculated on the basis of need alone (Impartiality and need: Relevance)

3 Aid will not be used to further a particular political or religious standpoint

(Unconditional aid: Impact and Relevance)

4 We shall endeavour not to act as instruments of government foreign policy (Independence: Impact and Relevance)

5 We shall respect culture and custom (Respect for cultures: Effectiveness, Impact)

6 We shall attempt to build disaster response on local capacities (Build on local capacities: Effectiveness, Impact, Efficiency)

7 Ways shall be found to involve programme beneficiaries in the management of relief aid (Participation: Effectiveness, Relevance)

8 Relief aid must strive to reduce future vulnerabilities to disaster as well as meeting basic needs (Vulnerability Reduction: Impact and Relevance)

9 We hold ourselves accountable to both those we seek to assist and those from whom we accept resources (Accountability: Impact and Relevance)

10 In our information, publicity and advertising activities, we shall recognise disaster victims as dignified humans, not hopeless objects (Disaster victims as dignified humans: Effectiveness, Impact and Relevance)

***

EVALUATING PROTECTION

Four Modes of Protective Practice

Denunciation is pressuring authorities through public disclosure into fulfilling their obligation and protecting individuals or groups exposed to abuse;

Persuasion is convincing the authorities through more private dialogue to fulfil their obligations and protect individuals and groups exposed to violations;

Substitution is directly providing services or material assistance to the victims of violations;

Support to structures is empowering existing national and/or local structures through project-oriented aid to enable them to carry out their functions to protect individuals and groups.

Sectorally-based Protective Assistance

*Aid as protection

In themselves, the provision of healthcare, water, sanitation, food security, livelihood support and psychosocial programming can play an enormous part in the practical protection of civilians who have suffered from massive violations of their rights in war. Such programmes are usually based directly on aid as a substitution for state provision or support to state services. They can help people realise their social and economic rights in war and help them recover their personal dignity.

*Aid as Protective

If aid can help to ensure people's immediate protection, it can also help to keep people safe. Used strategically protective assistance can function preventively to protect people actively from further attacks.

> Every humanitarian programme of protective assistance for past violations should also be designed with a protective edge that consciously seeks to prevent future violations.

As much as possible, all humanitarian aid programmes need to 'think safety' and focus on ways in which their sectoral programming can reduce people's vulnerability to attack and violation.

> *Note: Taken from Alnap guidance booklet Humanitarian Protection, pilot version 2004, H Slim, Luis Enrique Eguren*

Types of questions that could be included in TOR to cover protection:

What modes of protective practice did the organisation adopt at various stages of the humanitarian response? With what effect?

How well did the programming team understand the affected population's coping mechanisms, its will and capacities for self-protection and organisation?

What methods were used to assess violence against women throughout the programme cycle?

How was the agency's presence targeted to get close to particular groups of vulnerable people at particular high-risk times and in high-risk places?

How much are existing guidelines on protection disseminated within the organisation concerned i.e. UNHCR guidelines to counter sexual violence?

How aware were staff and field partners as to how to refer women and men seeking redress for human rights violations? How able were staff to link people's experiences of violations to specific legal standards?

What specific steps were taken to ensure and increase personal safety and security of women, girls, boys and men?

What measures were put in place regarding accommodation, transportation and security to enable women workers to do their jobs as safely as possible?

How much did staff include humanitarian values and principles in educational programmes in health promotion, schools and literacy groups?

Other useful sources of information:

- Growing the Sheltering Tree, - protecting rights through humanitarian action, IASC.

- http://www.icva.ch   NGO Statement on International protection 4 July 2003

- Protecting Refugees: A Field Guide for NGOs. UNHCR

PRINCIPLES AND GOOD PRACTICE OF HUMANITARIAN DONORSHIP (GHD)

Endorsed in Stockholm, 17 June 2003 by Germany, Australia, Belgium, Canada, the European Commission, Denmark, the United States, Finland, France, Ireland, Japan, Luxemburg, Norway, the Netherlands, the United Kingdom, Sweden and Switzerland.

OBJECTIVES AND DEFINITION OF HUMANITARIAN ACTION

1. The objectives of humanitarian action are to save lives, alleviate suffering and maintain human dignity during and in the aftermath of man-made crises and natural disasters, as well as to prevent and strengthen preparedness for the occurrence of such situations.

2. Humanitarian action should be guided by the humanitarian principles of *humanity*, meaning the centrality of saving human lives and alleviating suffering wherever it is found; *impartiality*, meaning the implementation of actions solely on the basis of need, without discrimination between or within affected populations; *neutrality*, meaning that humanitarian action must not favour any side in an armed conflict or other dispute where such action is carried out; and *independence*, meaning the autonomy of humanitarian objectives from the political, economic, military or other objectives that any actor may hold with regard to areas where humanitarian action is being implemented.

3. Humanitarian action includes the protection of civilians and those no longer taking part in hostilities, and the provision of food, water and sanitation, shelter, health services and other items of assistance, undertaken for the benefit of affected people and to facilitate the return to normal lives and livelihoods.

GENERAL PRINCIPLES

4. Respect and promote the implementation of international humanitarian law, refugee law and human rights.

5. While reaffirming the primary responsibility of states for the victims of humanitarian emergencies within their own borders, strive to ensure flexible and timely funding, on the basis of the collective obligation of striving to meet humanitarian needs.

6. Allocate humanitarian funding in proportion to needs and on the basis of needs assessments.

7. Request implementing humanitarian organisations to ensure, to the greatest possible extent, adequate involvement of beneficiaries in the design, implementation, monitoring and evaluation of humanitarian response.

8. Strengthen the capacity of affected countries and local communities to prevent, prepare for, mitigate and respond to humanitarian crises, with the

goal of ensuring that governments and local communities are better able to meet their responsibilities and co-ordinate effectively with humanitarian partners.

9. Provide humanitarian assistance in ways that are supportive of recovery and long-term development, striving to ensure support, where appropriate, to the maintenance and return of sustainable livelihoods and transitions from humanitarian relief to recovery and development activities.

10. Support and promote the central and unique role of the United Nations in providing leadership and co-ordination of international humanitarian action, the special role of the International Committee of the Red Cross, and the vital role of the United Nations, the International Red Cross and Red Crescent Movement and non-governmental organisations in implementing humanitarian action.

GOOD PRACTICES IN DONOR FINANCING, MANAGEMENT AND ACCOUNTABILITY

(a) Funding

11. Strive to ensure that funding of humanitarian action in new crises does not adversely affect the meeting of needs in ongoing crises.

12. Recognising the necessity of dynamic and flexible response to changing needs in humanitarian crises, strive to ensure predictability and flexibility in funding to United Nations agencies, funds and programmes and to other key humanitarian organisations

13. While stressing the importance of transparent and strategic priority-setting and financial planning by implementing organisations, explore the possibility of reducing, or enhancing the flexibility of, earmarking, and of introducing longer-term funding arrangements.

14. Contribute responsibly, and on the basis of burden-sharing, to United Nations Consolidated Inter-Agency Appeals and to International Red Cross and Red Crescent Movement appeals, and actively support the formulation of Common Humanitarian Action Plans (CHAP) as the primary instrument for strategic planning, prioritisation and co-ordination in complex emergencies.

(b) Promoting standards and enhancing implementation

15. Request that implementing humanitarian organisations fully adhere to good practice and are committed to promoting accountability, efficiency and effectiveness in implementing humanitarian action.

16. Promote the use of Inter-Agency Standing Committee guidelines and principles on humanitarian activities, the Guiding Principles on Internal Displacement and the 1994 Code of Conduct for the International Red Cross and Red Crescent Movement and Non-Governmental Organisations (NGOs) in Disaster Relief.

17. Maintain readiness to offer support to the implementation of humanitarian action, including the facilitation of safe humanitarian access.

18. Support mechanisms for contingency planning by humanitarian organisations, including, as appropriate, allocation of funding, to strengthen capacities for response.

19. Affirm the primary position of civilian organisations in implementing humanitarian action, particularly in areas affected by armed conflict. In situations where military capacity and assets are used to support the implementation of humanitarian action, ensure that such use is in conformity with international humanitarian law and humanitarian principles, and recognises the leading role of humanitarian organisations.

20. Support the implementation of the 1994 Guidelines on the Use of Military and Civil Defence Assets in Disaster Relief and the 2003 Guidelines on the Use of Military and Civil Defence Assets to Support United Nations Humanitarian Activities in Complex Emergencies.

(c) Learning and accountability

21. Support learning and accountability initiatives for the effective and efficient implementation of humanitarian action.

22. Encourage regular evaluations of international responses to humanitarian crises, including assessments of donor performance.

23. Ensure a high degree of accuracy, timeliness, and transparency in donor reporting on official humanitarian assistance spending, and encourage the development of standardised formats for such reporting

ACCOUNTABILITY AND QUALITY INITIATIVES IN HUMANITARIAN ASSISTANCE

This is only a partial list. There are many other initiatives in the sector. Some of the products of the initiatives can be used as evaluation frameworks. Sources for the list include (Borton, 2008) and (Macrae et al., 2002).

Initiative	Date	Origin	Key features
Code of Conduct for the International Red Cross & Red Crescent Movement and NGOs in Disaster Relief	1994	Formulated by the IFRC and other NGO representatives	States ten basic principles for agencies implementing humanitarian aid. Hundreds of signatories. (Borton, 1994)
Sphere Project	1996 First full version 2000	Developed by a coalition of European & US NGOs	Includes a Humanitarian Charter. Sets minimum standards in five sectors. Currently under revision. (Sphere Project, 2004).
People In Aid	1996 Code published in 1997, revised in 2003.	Established by a group of UK organisations	Sets out seven guiding principles on: human resources strategy; staff policies and practices; managing people; consultation and communication; recruitment and selection; learning, training and development; health, safety and security. (People in Aid, 2003)
ALNAP	1997	Supported by a wide range of agencies (donors, UN, NGOs, Red Cross movement etc), and hosted by ODI	Inter-agency forum to promote learning in the sector. Produces annual reviews of humanitarian action as well as lessons learned products and evaluation guidance. (http://www.alnap.org)
Humanitarian Ombudsman	1997 - now defunct	Concept developed by a group of UK organisations	Explored how a Humanitarian Ombudsman might act as an impartial and independent voice for people affected by emergencies.
Humanitarian Accountability Partnership (HAP)	2000	Developed out of the Humanitarian Ombudsman project, to become an international project based in Geneva	Dedicated to improving accountability within the humanitarian sector, initially through a programme of action-research. Now has benchmarks and certification. (HAP International, 2008)

Initiative	Date	Origin	Key features
Groupe URD Quality Compass	2000	Developed by French NGOs opposed to Sphere	Quality model and tools tailored to needs of humanitarian agencies. (Groupe URD, 2005a, 2005b)
Good Humanitarian Donorship	2003	Can be thought of a humanitarian parallel to the Paris Declaration process	Sets out 23 principles of good practice for donors to follow. Probably the shortest initiative of them all. (Good Humanitarian Donorship, 2003)
Global Accountability Framework	2003 first report Framework developed in 2005	One World Trust promotion of good governance for a better world	Produces and annual Global Accountability Report that compares the accountability of NGOs, UN and similar organisations, and transnational corporations. (Lloyd et al., 2007; Lloyd et al., 2008)
WANGO Code of Ethics and Conduct	2004	Developed by the World Association of Non Governmental Organizations	The code covers all NGOs and not just humanitarian or development ones. (WANGO, 2004a, 2004b)
Fritz Institute Beneficiary Surveys	2005	Began to address the issues of the extent to which humanitarian action met the needs of the affected population	Rigorous quantitative surveys in the aftermath of natural disasters establishing that the way in which aid was delivered was as important to affected populations as the aid itself. (Bliss and Campbell, 2007a, 2007b; Bliss et al., 2006; Fritz Institute, 2005a, 2005b, 2005c, 2006)
Emergency Capacity Building Project	2005 - now lapsed	Project by six leading NGOs. Conducted joint evaluations.	Intended to improve the speed, quality, and effectiveness of the disaster and emergency response work within the humanitarian community. Produced the Good Enough Guide. (ECB, 2007)
Cluster Coordination Approach	2005	Part of the UN's Humanitarian Reform initiative	Designed to improve the coordination of humanitarian response and make leadership more predictable and more effective. (http://clustercoordination.org/)

Initiative	Date	Origin	Key features
International Non-Governmental Organisations Accountability Charter	2005	A reaction by humanitarian, development, human rights, and environmental NGOs to curtail their freedom of action citing accountability concerns.(Action Aid International et al., 2005)	Intended to promote further the values of transparency and accountability that the members stand for through addressing issues of their own transparency and accountability. (ActionAid International et al., 2005)
IASC IARTEs	2006	Developed to evaluate the implementation of the cluster coordination approach	Conducting a series of real time evaluations of the operation of the cluster coordinator approach in major disasters. (Cosgrave et al., 2007; Turner et al., 2008; Young et al., 2007)
Global Humanitarian Platform (GHP)	2006	This is a forum bringing together NGOs, the Red Cross, and the UN and related international organisations.	The overall goal of the Global Humanitarian Platform (GHP) is to enhance the effectiveness of humanitarian action. It is premised on the belief that no single humanitarian agency can cover all humanitarian needs and that collaboration is, therefore, not an option, but a necessity. (Global Humanitarian Platform, 2007)
DARA Humanitarian Response Index	2007	Privately funded index to measure donor compliance with Good Humanitarian Donorship principles	Based on both quantitative and qualitative criteria, established an annual ranking of donor performance against the Good Humanitarian Donorship criteria. (Hidalgo and Development Assistance Research Associates, 2009; Hidalgo et al., 2008)

REFERENCES AND RESOURCES

ActionAid International, Amnesty International, CIVICUS World Alliance for Citizen Participation, Consumers International, Greenpeace International, Oxfam International, International Save the Children Alliance, Survival International, International Federation Terre des Hommes, Transparency International, & The World YWCA. (2005). *International Non-Governmental Organisations Accountability Charter*. London: International Non-Governmental Organisations Accountability Charter Founding Signatories. Last viewed on 11 June, 2008. URL: http://www.ingoaccountabilitycharter.org/download/ingo-accountability-charter-eng.pdf

Bliss, D., & Campbell, J. (2007a). *The immediate response to the Java Tsunami: Perceptions of the Affected* San Francisco: Fritz Institute. Last viewed on 8 June, 2008. URL: http://www.fritzinstitute.org/PDFs/findings/JavaTsunami_Perceptions.pdf

Bliss, D., & Campbell, J. (2007b). *Recovering from the Java earthquake: Perceptions of the affected*. Last viewed on 8 June 2008. URL: http://www.fritzinstitute.org/PDFs/findings/JavaEarthquake_Perceptions.pdf

Bliss, D., Larsen, L., & Fritz Institute. (2006). *Surviving the Pakistan Earthquake: Perceptions of survivors one year on.* San Francisco: Fritz Institute. Last viewed on 8 June, 2008. URL: http://www.fritzinstitute.org/PDFs/findings/PakistanEarthquake_Perceptions.pdf

Borton, J. (2008). *Summary of the Emerging Results from the ALNAP Humanitarian Performance Project: Exploratory Phase* London: ALNAP. Last viewed on 28 February 2009. URL: http://www.alnap.org/meetings/pdfs/23_jborton.pdf

Borton, J. (Ed.). (1994). *Code of conduct for the International Red Cross and Red Crescent Movement and NGOs in disaster relief: Network Paper 7*. London: Relief and Rehabilitation Network, Overseas Development Institute. Last viewed on 21 January 2009. URL: http://www.odihpn.org/documents/networkpaper07.pdf

Cosgrave, J., Gonçalves, C., Martyris, D., Polastro, R., & Sikumba-Dils, M. (2007). *Inter-agency real-time evaluation of the response to the February 2007 floods and cyclone in Mozambique*. Geneva: Inter-agency Standing Committee. Last viewed on 8 June, 2008. URL: http://www.reliefweb.int/rw/RWFiles2007.nsf/FilesByRWDocUnidFilename/CFF788904EBDC1D7852572E6005E1600-Full_Report.pdf/$File/Full_Report.pdf

ECB. (2007). *The Good Enough Guide: Impact Measurement and Accountability in Emergencies*. Oxford: Oxfam for the Emergency Capacity Building Project. Last viewed on 1 July 2009. URL: http://www.oxfam.org.uk/what_we_do/resources/downloads/Good_Enough_Guide.pdf

Fritz Institute. (2005a). *Lessons from the Tsunami: Survey of Affected Families in India and Sri Lanka*: Fritz Institute. Last viewed on 9 February 2009. URL: http://www.fritzinstitute.org/PDFs/findings/AffectedFamilies.pdf

Fritz Institute. (2005b). *Lessons from the Tsunami: Top Line Findings*. San Francisco: Fritz Institute. Last viewed on 8 June 2008. URL: http://www.fritzinstitute.org/PDFs/Programs/Findings_Sept2605.pdf

Fritz Institute. (2005c). *Recipient perceptions of aid effectiveness: rescue, relief and rehabilitation in tsunami affected Indonesia, India and Sri Lanka*. San Francisco: Fritz Institute. Last viewed on 8 June 2008. URL: http://www.fritzinstitute.org/PDFs/findings/NineMonthReport.pdf

Fritz Institute. (2006). *Hurricane Katrina: Perceptions of the Affected*. San Francisco: Fritz Institute. URL: http://www.fritzinstitute.org/PDFs/findings/HurricaneKatrina_Perceptions.pdf

Global Humanitarian Platform. (2007). *Principles of Partnership*. Geneva: Global Humanitarian Platform. Last viewed on 7 July 2009. URL: http://www.globalhumanitarianplatform.org/doc00002628.doc

Good Humanitarian Donorship. (2003). *Principles and Good Practice of Humanitarian Donorship*. Stockholm: Germany, Australia, Belgium, Canada, the European Commission, Denmark, the United States, Finland, France, Ireland, Japan, Luxembourg, Norway, the Netherlands, the United Kingdom, Sweden and Switzerland. Last viewed on 14 May 2009. URL: http://www.reliefweb.int/ghd/a%2023%20Principles%20EN-GHD19.10.04%20RED.doc

Groupe URD. (2005a). *Project Evaluation Companion Book*. La Fontaine des Marins: Groupe URD. Last viewed on 07 July 2009. URL: http://www.projetqualite.org/compas/outil/uk/docs/project_evaluation_companion_book.doc

Groupe URD. (2005b). *Project Management Companion Book*. La Fontaine des Marins: Groupe URD. Last viewed on 07 July 2009. URL: http://www.projetqualite.org/compas/outil/uk/docs/project_management_companion_book.doc

HAP International. (2008). *The Guide to the HAP Standard: Humanitarian Accountability and Quality Management*. Oxford: Oxfam. Last viewed on 6 July 2009. URL: http://www.oxfam.org.uk/resources/downloads/HAP/HAP_book.pdf

Hidalgo, S., & Development Assistance Research Associates (Eds.). (2009). *Humanitarian response index 2008 : Donor accountability in humanitarian action*. Basingstoke: Palgrave Macmillan

Hidalgo, S., López-Claros, A., & Development Assistance Research, A. (Eds.). (2008). *Humanitarian response index 2007 : measuring commitment to best practice*. Basingstoke: Palgrave Macmillan

Lloyd, R., Oatham, J., & Hamme, M. (2007). *2007 Global Accountability Report*. London: One World Trust. Last viewed on 30 December 2008. URL: http://oneworldtrust.org/index.php?option=com_docman&task=doc_view&gid=169&tmpl=component&format=raw&Itemid=55

Lloyd, R., Warren, S., & Hamme, M. (2008). *2008 Global Accountability Report*. London: One World Trust. Last viewed on 30 December 2008. URL: http://oneworldtrust.org/index.php?option=com_docman&task=doc_view&gid=226&tmpl=component&format=raw&Itemid=55

Macrae, J., Collinson, S., Buchanan-Smith, M., Reindorp, N., Schmidt, A., Mowjee, T., & Harmer, A. (2002). *Uncertain Power: The Changing Role of Official Donors in Humanitarian Action: HPG Report 12*. London: Overseas

Development Institute. URL: http://www.odi.org.uk/resources/download/242.pdf

People in Aid. (2003). *Code of Good Practice in the managment and support of aid personnel*. London: People in Aid. Last viewed on 7 July 2009. URL: http://www.peopleinaid.org.uk/pool/files/publications/code.pdf

Sphere Project. (2004). *Humanitarian charter and minimum standards in disaster response* (2004 ed.). Geneva: Sphere Project. Last viewed on 8 June, 2008. URL: http://www.sphereproject.org/component/option,com_docman/task,doc_download/gid,12/Itemid,26/lang,English/

Turner, R., Baker, J., Oo, Z. M., & Aye, N. S. (2008). *Inter-Agency Real Time Evaluation of the Response to Cyclone Nargis: 17 December 2008*. Geneva and New York: Inter-Agency Standing Committee

WANGO. (2004a). *Code of Ethics and Conduct or NGOs*. Tarrytown: World Association of Non Governmental Organisations. Last viewed on 7 July 2009. URL: http://www.wango.org/codeofethics/COEEnglish.pdf

WANGO. (2004b). *Code of Ethics and Conduct or NGOs: Compliance Manual*. Tarrytown: World Association of Non Governmental Organisations. Last viewed on 7 July 2009. URL: http://www.wango.org/codeofethics/ComplianceManual.pdf

Young, N., Khattak, S. G., Bengali, K., & Elmi, L. (2007). *IASC inter-agency real time evalaution of the Pakistan floods/Cyclone Yemyin: Final version*. New York: OCHA for the Inter-Agency Standing Committee. Last viewed on 8 June, 2008. URL: http://ocha-gwapps1.unog.ch/rw/RWFiles2007.nsf/FilesByRWDocUnidFilename/KHII-78Q5FS-full_report.pdf/$File/full_report.pdf

POTENTIAL ADVANTAGES AND DISADVANTAGES OF INTERNAL AND EXTERNAL EVALUATORS

Internal Evaluators	External Evaluators
+ Know the organisation	+ Objective
+ Understand organisational behaviour and attitudes	+ No organisational bias
+ Are known to staff	+ Fresh perspectives
+ Are less threatening	+ Broader experience
+ Often a greater chance of adopting recommendations	+ More easily hired for longer periods of time
+ Are less expensive	+ Can serve as an outside expert
+ Build internal evaluation capability	+ Not part of the power structure
+ Contribute to programme capacity	+ Can bring in additional resources
− Objectivity may be questioned	+ Trained in evaluation
− Structure may constrain participation	+ Experienced in other evaluations
− Personal gain may be questioned	+ Regarded as an 'expert'
− Accept the assumptions of the organisation	− May not know the organisation
− Full participation may be constrained by usual workload	− May not know of constraints affecting recommendations
− May not be trained in evaluation methods	− May be perceived as an adversary
− May lead to the evaluation not having acceptable outside credibility	− Expensive
− May have difficulty avoiding bias	− Contract negotiations may take time
− May lack special technical expertise	− Follow up on recommendations is not always there
	− Unfamiliar with environment

Section three:
Planning, budgeting, and managing the evaluation

This section provides some practical tools and checklists for planning, budgeting and managing evaluations, both for the overall evaluation manager in the commissioning agency, and for the evaluation team leader.

It presents a checklist for the management of the evaluation, offers practical hints on timing, as well as offering a budget checklist and framework.

The section also offers a checklist for evaluation selection. A final element is a note on the special issues that arise in managing a multi-agency evaluation.

EHA MANAGEMENT CHECKLIST

Preparatory action by evaluation team.

❑ Commenting on the TOR.

❑ Preparing others – who else needs to prepare for the evaluation.

❑ Making sure that the basic data will be available and sufficient copies.

❑ Making sure that important data, trends, insights are well documented, organised, presented.

❑ Preparing visits – check what dates and times are possible. When selecting people or organisations to be visited, evaluators need to make sure that they consider more than their own priorities.

❑ Organisational logistical issues – analysing chronologically what will be needed.

❑ Preparing the team itself – ensure that each member knows the objectives for the evaluation, preparatory meetings.

Elements of the document set

❑ Original and subsequent project emails and memoranda

❑ Any previous reviews or evaluations of the projects/themes

❑ Original and revised programme proposals or log frames

❑ Relevant policy and strategy documents

❑ General literature related to the sector/issue/theme

❑ Relevant evaluations and reviews by other donors

❑ Policies and strategies of the partner government/institutions

Planning the fieldwork

- ❏ Time: length, share of total evaluation time

- ❏ Timeliness: relative to emergency assistance, relative to other variables affecting effectiveness

- ❏ Field contacts: beneficiaries, field actors, donor representatives, agency headquarters

- ❏ Structure of fieldwork: single stage, multi stage.

- ❏ Division of responsibilities: sectoral, country/area.

- ❏ Limits: access, security, beneficiary recall, attrition of staff, availability of records

Logistics of country visits and fieldwork

- ❏ First contact by commission agency with field offices, partner governments, and others.

- ❏ All necessary permits and clearances issues.

- ❏ Staff security briefing and any special requirements.

- ❏ Planned travel dates.

- ❏ National holidays

- ❏ Any rules on working days, accommodation and per diems, and the applicable rates.

Briefing by evaluation manager to include:

- ❏ TOR

- ❏ Initial work plan

- ❏ Advisory group

- ❏ Roles and responsibilities (incl. Evaluation Manager)

- ❏ Support provided

❑ Document and files required

❑ Confidentiality and security clearance

❑ Best practice examples of evaluation reports

❑ Any requirements for regular reporting

❑ Procedures for dealing with any exception issues (the team coming across fraud for example)

❑ Constraints to the evaluation, security, political, weather factors.

Questions that the evaluator should know the answer to

❑ What actions did the evaluation manager undertake prior to your arrival on the scene?

❑ What decisions were made as to the timing and content of the evaluation?

❑ What political realities and contextual issues affected the process?

❑ Was cooperation gained from all stakeholders?

❑ Are stakeholders aware of how the evaluation will benefit them? Was the exercise painted as a contribution to dialogue and not a judgement?

❑ Were the stakeholders involved in planning – discussing goals, building consensus, planning the evaluation approach – and to what degree?

❑ Was the TOR submitted to stakeholders for their approval to help in gaining commitment to the evaluation?

❑ Were the constraints (security, lack of data, etc.) to the evaluation made clear to the stakeholders?

❑ What will the review process for the report draft be?

❑ Can the team share the draft with interviewees for their comments and corrections or will the evaluation manager do this?

❑ What format will dissemination take?

TIMING OF AN EVALUATION – PRACTICAL CONSIDERATIONS

When scheduling evaluations, consider the timing carefully, from a number of different perspectives:

1) How can the timing maximise the chances of the evaluation being used? For example, pay attention to:

- policy-making cycles and the timing of key policy decisions
- the project cycle and how the evaluation can feed into programming decisions
- funding decisions, perhaps in relation to annual assessments and appeals

2) What do you need to be aware of in terms of the political, social and livelihoods context. For example:

- are there political events such as elections that you want to avoid?
- pay attention to the timing of public holidays and festivals
- in a conflict context, are there seasonal patterns to the conflict that you should be aware of?

3) What institutional factors should you take account of in scheduling the timing? For example:

- what are the patterns of staff holidays?
- how can the evaluation best fit into work cycles and the timing of the fiscal year?

4) How might timing affect the logistics and methodology of the field work? For example:

- what is the timing of the agricultural season? It is best to avoid fieldwork during periods of peak agricultural work eg harvest time, and the rainy season may hamper travel
- what else may affect the availability of local people for interviewing eg seasonal migration patterns
- what are the seasonal patterns of malnutrition and certain diseases?

A PRACTICAL BUDGET CHECKLIST AND MATRIX

CHECKLIST FOR FIELDWORK BUDGET

	Budget Item	Comments
☐	*Personnel costs and benefits*	This may arise when staff of the commissioning agency have to have salaries (rare) or accommodation and per-diem costs paid from the evaluation budget. This can sometimes arise with the travel or per-diem costs for partners.
☐	*Consultant costs*	Consultants are often paid at different rates. Typically the team leader is the most expensive and national consultants are the cheapest. It is best to break down the number of consultant days by the evaluation tasks which may include: • Briefing • Document assembly • Document analysis • Pre –mission visits • Fieldwork • Travel time (can be significant for some studies) • Writing • Debriefing • Commenting • Finalising • Workshops to present report • Other dissemination tasks
☐	*Support staff costs*	Obviously you want to have your expensive consultants focussing on what only they can do and try and have basic administration work done by others. Support staff costs can include: • Interpreter costs • Researcher costs for document assembly or other tasks • Secretarial support for making travel arrangements • Secretarial support for setting up meetings

	Budget Item	Comments
❑	*Travel costs*	Here you need to think about travel costs not only for the consultants, but also for anyone who is to accompany them (such as drivers or the staff of local partners) if this is not covered by other budgets. Travel costs will include: - Travel to briefing, debriefing and other meetings. - Visas and any mandatory medical costs - Travel to the country or countries where the field work is to be conducted - Internal travel within the fieldwork country, this may sometimes include the costs of escorts etc - Travel for any feedback or dissemination activities - Travel by interviewees when conducting a remote evaluation of countries where it is not safe for the evaluation team to travel
❑	*Data collection costs*	Some types of data collection may be contracted out. Surveys and their associated data processing can be a significant cost. A typical professional survey in the developing world can cost 5 to 10 USD per interviewee, for a survey of 600 to 1,000 interviewees. There may also be translation costs where data is in a language that is not accessible to the evaluation team. Some methods like focus group discussions may bring minor costs for refreshments etc.
❑	*Accommodation and per diems*	These can often be a significant cost. These cost may apply not only for the evaluation team but also agency or local partner staff accompanying the evaluation team. They may also apply to other stakeholders (such as government) who attend workshops or meetings (depending on local practice).
❑	*Meeting venue fees*	The costs of any meeting venue, including any catering or services (such as simultaneous interpreting) provided. You also need to consider the likely cost of any dissemination events.
❑	*Special equipment or software*	If the evaluation requires any special equipment (e.g. body armour for the evaluation team) or software (specific software to work with agency records in a particular format) then these costs should be built in here.

	Budget Item	Comments
❑	*Communications, mail and teleconferencing costs*	Communication costs are rarely comparable with other costs, given that the price of telecommunications has fallen so much. However, they must be allowed for, especially if planning to use expensive facilities such as video-conferencing suites.
❑	*Overhead*	A consultant company will need to cover its overheads, typically this is done by loading the rates paid to the consultants (so that this cost is often not transparent) and applying a percentage management charge to all other expenses.

CHECKLIST FOR REPORT PRODUCTION BUDGET

The elements above include the elements needed to produce a draft report and to disseminate the findings. However, the physical production of the report can also be a significant cost and needs to be considered. This is presented separately as it may not be part of the main evaluation contract.

	Budget Item	Comments
❑	*Translation*	If the report is to be translated, this can be a significant cost, depending on the size of the report and the number of languages. It is to be translated into. Very sensitive reports may require back translation by a separate translator to verify the translation.
❑	*Report editing*	A major report will need to be copy edited to be suitable for placing in the public domain. Translated versions will also require copy editing.
❑	*Design and layout*	Design, layout, and typesetting costs need to be allowed for. Again, producing a report in multiple language versions can increase these.
❑	*Printing*	Printing costs vary with the number of reports, the number of pages, the number of colours, and the report design.
❑	*Electronic media*	If it is planned to produce a CD Rom of the report, or a website to present the main findings, these costs needs to be included unless they are covered from other budgets.
❑	*Licence fees*	In some cases, licence fees or royalties may arise from the use of copyright images, material, or other intellectual property.

Budget Item	Comments
☐ *Legal fees*	In some rare cases, it may be necessary to seek legal advice before a report is published.
☐ *Report Distribution*	The mundane task of distributing a large report to a wide range of stakeholders can be challenging. Posting a large number of heavy reports can be a non-trivial cost, especially as courier delivery may be the only option for many.

BUDGET MATRIX FOR COSTS FOR EVALUATION WORK

Typically the detailed budget will be drawn up by looking at the cost of each of the evaluation tasks, such as documentary study, inception report, fieldwork visit or visits, draft report and so on. This can often be many more than the five task used for this sample budget matrix.

Budget Item	Task 1	Task 2	Task 3	Task 4	Task 6
Personnel costs and benefits Own Staff					
Partners Staff					
Others staff					
Consultant costs Team Leader					
International consultants					
National consultants					
Other Consultants					
Support staff costs Admin					
Background researchers					
Interpreters					
Others					
Travel costs Visas etc					
Flights for team					
Flights for others					
Internal travel for team					

Budget Item	Task 1	Task 2	Task 3	Task 4	Task 6
Internal travel for others					
Other Travel					
Data collection costs					
Accommodation and per diems Team					
Agency staff					
Partner staff					
Other staff					
Meeting attendees					
Interviewees					
Meeting venue fees					
Special equipment or software					
Communications, mail and teleconferencing costs					
Overhead costs					

BUDGET MATRIX FOR COSTS FOR THE REPORT AND OTHER PRODUCTS.

The budget matrix for the production cost of specific evaluation products is usually based around the products, be they reports, summaries, key message sheets or others. In some cases these costs will be met from general budgets for in-house staff, but where this is not the case, they costs need to be specifically allowed for.

Budget Item	Product 1	Product 2	Product 3	Product 4
Translation				
Report editing				
Design and layout				
Printing				

Budget Item	Product 1	Product 2	Product 3	Product 4
Electronic media				
Licence fees				
Legal fees				
Report Distribution				

EVALUATOR SELECTION CHECKLIST

Very loosely based on (Jones et al., 2004, pp. 38-39).

Candidate characteristics

❑ A sufficiently deep knowledge and experience of humanitarian action to be accepted as credible, both by field staff and by the managers who will implement the evaluation recommendations.

❑ The ability to interact with and to get useful information from a diverse range of actors including agency staff, the affected population, local authorities and the military.

❑ Ability to think systematically and rigorously.

❑ Open to new ideas.

❑ Systematic and thorough

❑ Sufficiently detached not to have an agenda.

❑ Able to thinks systematically and rigorously

❑ Experience of conducting humanitarian evaluations.

❑ Able complete evaluations to tight time constraints and to deliver reports on time.

❑ Organised enough to manage the data collection and synthesis.

❑ The leadership skills to lead the evaluation team.

❑ Good knowledge of the context of the region.

❑ Good knowledge of, and sensitivity to, the political context of the agency or structure to be evaluated.

❑ Good communication skills.

❑ Available for the time needed to complete the evaluation.

❑ The technical knowledge to deal with any specific technical aspects.

❑ Flexible and able to cope with changed plans.

REFERENCES AND RESOURCES

Jones, R., Young, V., & Stanley, C. (2004). *CIDA Evaluation Guide*. Ottawa: Canadian International Development Agency. Last viewed on 28 June 2008. URL: http://www.acdi-cida.gc.ca/INET/IMAGES.NSF/vLUImages/Performancereview5/$file/English_E_guide.pdf

TEAMWORK AND LEADERSHIP

The following approach to leadership was developed by John Adair in conjunction with the Industrial Society. In John Adair's view, there is no standard format for successful leadership and teamwork. It is therefore important to look at the actions that a leader and team members have to take in order to be effective. These actions relate to achieving a balance between the following three elements of teamwork:

- The task
- The team
- The individual

The three elements are interrelated, and neglecting one element can lead to the degeneration of all three. Below are some considerations which may be necessary for effective teamwork to complete a task: -

The task	*The team*	*The individual*
❖ Set clear objectives for the task ❖ Assess the needs of individuals, the team and the task ❖ Plan for completion of the task ❖ Do the task ❖ Set up systems for monitoring and evaluating the task ❖ Modify operational systems where necessary ❖ Complete the task	❖ Establish and agree a clear goal with the team ❖ Establish commitment to the task and system ❖ Set targets and standards ❖ Allocate roles, including any leadership roles ❖ Co-ordinate and co-operate with individuals and the team ❖ Set up clear communication and consultation systems ❖ Decide what the decision making process will be ❖ Ensure the whole team can participate	❖ Motivate ❖ Encourage contributions ❖ Give responsibility ❖ Facilitate ownership ❖ Recognise individual's skills

MANAGING MULTI-AGENCY EVALUATIONS (MAEs)

The following is a short note on managing Multi-Agency Evaluations (MAEs)[5]. It sets out some indicative considerations and is not comprehensive. It is not an evaluation guide, rather a complement to such guides focussing on the specificity of MAEs. The note is based on the assumption that the two overall differences between MAEs and other evaluations boil down to, **firstly**, the number of actors involved (especially in the management of the evaluation) and **secondly**, the potential breadth of the activities, content or programme to be evaluated (e.g. a multi-actor programme rather than a single agency programme, project or activity level).

The note takes a chronological approach, through the main evaluation management steps or phases (allowing, however, that phases may overlap and run parallel). It is based on the assumption that the evaluation is a classical evaluation exercise, involving a team of consultants (or mixed with agency staff), going to 'the field', returning and reporting. If the exercise involves a number of constituent evaluations or studies, each one might be managed in the following manner, within a larger, overall structure and process. Considerations and key questions are set out as bullet-points under each step:

❖ Assessment and scoping of the value of conducting conduct an MAE

- What is the added-value of a joint exercise as opposed to individual exercises, especially in function of its ultimate usage/user-focus? If so, what might those uses be, aimed at what target groups/audiences?

- Is there sufficient buy-in for an MAE, including understanding of the complexities, costs and benefits of an MAE?

- Can the multiple possible stakeholders to be mapped, prioritised and consulted and by whom and by when will they be consulted?

❖ Establishing a multi-agency management structure

- Can the main actors be identified and committed to the process, e.g. through the unambiguous provision of time and resources?

- Can a lead or host agency be identified? This is to provide a legally established umbrella organisation e.g. that the MAE be under the aegis of one of the constituent organisations, as a host, providing a physical 'home'; administrative and contracting body (e.g. for the team); accounts and legal status; etc.

- Based on the identification of those actors, what is the most effective and efficient management structure? This will probably be multi-layered, including: an overall MAE group, including an appropriate chairperson; a smaller management sub-committee (e.g. 3 – 5 people, ideally including the overall chairperson); a day-to-day manager (who would also sit on the management sub-committee, but without voting rights); an administrative and coordination secretariat (based in the host agency, to support the manager and entire under-taking); and the evaluation team.

- An explicit agreement on roles, responsibilities, rights and obligations of all concerned. This includes fundamentally who holds the 'ownership' of the process and its products (e.g. the report). Ownership entails a series of aspects, from legal rights and obligations, to decision-making authority, especially in the event of disputes.

- Agreement and formalization of procedures for relevant aspects of the process, including establishment of the management structure; dismantling of the structure

---

[5] This short note reflects comments made by the author, John Telford, in response to a number of requests and initiatives; a roundtable meeting on the subject of joint evaluations, sponsored by ALNAP, DEC, ECB2 and the TEC in 2006; a review of the Tsunami Evaluation Coalition (TEC) exercise, also held in 2006; and a draft guide being produced by the Emergency Capacity Building (ECB) project.

(including consideration of future ownership of the products of the MAE); a disputes-management mechanism; and financial, material and human resource management (including contracting the evaluation team and other resource provision, especially for day-to-day secretariat functions).

❖ Designing the MAE and TOR

- Managing and focusing multiple possible uses, needs and expectations (links to assessment and scoping of need above). This includes prioritization of the possible target users and audiences for the possible products coming from the evaluation.

- Delegating authority to the management sub-committee to make these decisions with the confidence and support of the broader MAE group, subject their review and approval.

❖ Team selection, preparation and planning

- Decision by the MAE group and/or sub-committee on the nature and size of the team.

- Delegation of selection to the management sub-committee, based on accepted standards of professionalism, independence and transparency.

- Close coordination, preparation and planning among the management sub-committee, contracting (host) entity, MAE manager and the evaluation team. In function of the scoping and TOR, this includes the transparent selection of locations to be visited, possible stakeholders to be involved,

❖ Conducting the MAE, including analysis and reporting

- Application of methods and availability of time and resources in accordance with the possible breadth of aspects, issues and locations to be covered.

- Allowing sufficient time and 'space' (e.g. workshops) for the above and for team and stakeholder analysis of what may be very considerable materials and findings emerging (possibly more than in a single-agency undertaking).

- Clear agreements on when, how and by whom the draft reports will be reviewed. Of particular importance is explicit agreement on the authority of reviewers, especially if they are from within the overall management group i.e. are certain types of comments to always to be acted upon (such as errors of fact or inadequate verification) while others are of an advisory nature only (e.g. interpretations or analysis).

❖ Dissemination and use

- Agreement (well in advance, preferably at the outset) on the number and type of products that will result from the MAE (according to the diverse sets of target groups/audiences).

- A usage, dissemination and communications plan for the outputs of the exercise. This would explain whether the process is centralized (managed by the sub-committee and/or the secretariat, or some such group), de-centralised (among all the agencies involved and possibly others), or a mixture of both. Given the complex range of possible stakeholders and locations involved, planning may need to be delegated to a sub-committee.

- Resources available for the implementation of the plan, and for unforeseen costs.

- A follow-up plan, including whatever activities are seen to be relevant, as decided by the overall MAE group. This would be in function of the initial scoping, TOR and

results of the MAE. It could imply a new structure and process involving the agencies wishing to take the results forward into a new review and action process.

- A review of the MAE itself, recording lessons on the exercise. This would probably require a workshop or one-day meeting of all main actors and stakeholders.

Section four:
Field work methods

The context of EHA can be challenging for the implementation and selection of appropriate field work methods.

This section begins with an ethical guide for the evaluator or evaluation manager as a reminder that all field work must be conducted in an ethical way.

It presents a list of data collection methods and then explores some of these in detail, for example carrying out focus group discussions, drawing up an interview guide, and carrying out an After Action Review. A checklist to help select different field work methods is provided.

ETHICAL CONSIDERATIONS

Evaluators should aspire to provide high quality information and assessment and to conduct a high quality evaluation process. Ethical considerations are integral to this.

Ethical guidance adapted from that provided by CIDA (1990), and as partially reproduced by Danida in its 'Evaluation Guidelines' (2006):

Cultural Intrusion	Local customs regarding dress, personal interaction, religious beliefs and practices should be respected.
Anonymity/confidentiality	Evaluators must respect people's right to provide information in confidence, and must ensure that sensitive information cannot be traced to its source.
Responsibility for evaluations	In some countries, criticism can have serious consequences for a national. Evaluators must take care that those involved as local evaluators either endorse a report, or that their restricted roles are clearly described in the report. Statements should not be made on behalf of the evaluation team if other team members have not had an opportunity to disagree.
Right to privacy	Evaluators should realise that people can be extremely busy and their participation in evaluations can be burdensome. Therefore, evaluators should provide ample notice and minimise demands on time.
Supremacy of fundamental values	There is a delicate balance between certain cultural practices and the deprivation of fundamental rights and freedoms. While evaluators are expected to respect other cultures, they must also be aware of the values affecting minorities and particular groups. In such matters the UN Universal Declaration of Human Rights (1948) is the operative guide.
Omissions	Ethically, evaluators have a responsibility to bring to light issues and findings which may not relate directly to the Terms of Reference. Certain other issues can cause difficulties for the evaluator and should be acknowledged and discussed with the Evaluation manager as necessary.
Evaluation of individuals	Performance evaluation is not normally a part of evaluations, though reports will touch on issues such as leadership and management competence that border on evaluation of individuals. The evaluator is not expected to evaluate individuals and must balance an evaluation of management functions with this general principle.

CONSENT AND CONFIDENTIALITY

Evaluations are a form of social research. Social research is normally conducted on the basis of informed consent. Interviewees should be informed of what the interview ground rules are. For key informant interview this may be done with a simple card. Humanitarian evaluation is normally conducted on the basis of a

variant of the Chatham House Rule, where the comments by interviews are not attributed to them, either directly or indirectly.

Evaluations may occur in complex emergencies or other contexts where interviewees could be at risk if they are identified. In some extreme cases interviewees, names should not be reported but simply replaced by *"Interviewee One", "Interviewee Two"* etc. in a list. However if one interviewee at a particular location is made anonymous in this way, it may be appropriate for all interviewees at that location to be made anonymous.

LIST OF DATA COLLECTION METHODS

1. Desk study/Literature search.   This is a key first step and is an economic and efficient way of obtaining information. However, it can be difficult to assess the reliability and validity of some sources. The volume of documents depends on the context.

2. Key-informant interviews This is a flexible, in-depth approach that forms the back-bone of data collection on many humanitarian evaluations. It is relatively easy to implement. Need to be aware that all interviewees have their own biases and agendas. Usually done using a semi-structured interview guide.

3. Observation   This is a very useful method, especially for triangulating the information from other sources. It can also be used as a primary data collection method (e.g. observing food distributions) but needs to be combined with other data collection methods (such as interviews) to ensure the observations are not misinterpreted.

4. Reviewing records   Almost all projects maintain records of some sorts. Records may be internal as opposed to public documents. Analysis of records may identify particular patterns that can answer evaluation questions or lead to further questions.

5. Group interviews   A low-cost and efficient means of collecting beneficiary views in a quicker but less rigorous way than a formal survey. Less suitable for sensitive issues, but the presence of a range of persons at the interview can help to ensure veracity.

6. Focus group interviews   An intensive facilitated discussion with 6-8 persons where the moderator used a prepared guide or question trajectory to steer the discussion to areas of interest. Can be very fruitful, but needs effort to get it right. Good for generating understanding of participants perceptions.

7. On-line survey   Obviously limited to those with direct internet access. Allows quick and cheap surveys that can be used to identify issues for further analysis.

8. Field survey   Statistically valid surveys where enumerators tabulated responses to a structured survey instrument in the field. Design of instrument and sample selection needs to be done carefully. Often subcontracted to specialist market research firm.

9. Case studies   In-depth review of one or a small number of selected cases. Well-suited for understanding processes and for formulating hypotheses to be tested later. Ideally done to test an existing theory.

10. Ethnographic interviewing In-depth interviewing of a limited number of individuals to provide a good picture of how a particular event has affected them. Helps to put human detail onto a larger canvas.

11. Story-telling/ collection   Obtaining participants' and communities' experiences of change by collating their observations and stories. The Most Significant Change technique is an example of this method.

12. After-action review A facilitated discussion that focuses on four questions:

    - What was planned?

    - What happened?

    - What went well?

    - What could have been better?

13. Participatory Rural Appraisal techniques   A wide range of techniques that enable those from outside the community to capture knowledge that is held by the community. PRA tools can be thought of as helping communities to overtly analyse issues and to translate their analysis into a format that those outside the community can understand. Examples of PRA techniques are:

    - Calendars (seasonal, 24 hour, multi-annual) and other calendars

    - Proportional piling

    - Ranking (wealth, seeds, coping strategies, etc)

    - Transect walks

    - Mapping (wealth, hazard, mobility, social, resource, risk, network, influence, relationship etc)

    - Asset wheel

    - Venn diagrams

    - Time lines/histories

    - Causal flow diagram

    - Entitlement matrix

    - Stakeholder analysis

14. Triangulation Triangulation is the key technique to ensure accuracy and reliability in qualitative and mixed method research. Essentially, triangulation consists of looking at data from different sources to see whether they support the same interpretation. Triangulation can be based on:

    - Method triangulation: Comparing data generated by different research methods (e.g. comparing observations with group interviews)

    - Source triangulation: Comparing information from key-informants at headquarters with information from key-informants in the field.

    - Researcher triangulation: Comparing information from different researchers.

- Analytical triangulation: Comparing the results of different analytical techniques to see how they compare, for example, you could compare how the number of references in official documents to a particular issue varies, against how the level of funding for that issue varies.

CONDUCTING THE FIELDWORK

RECORD KEEPING DURING FIELDWORK

The time demands of humanitarian evaluation mean that records must be kept up to date through the evaluation rather than being produced at the end by a forensic examination of notes and tickets etc. This applies to:

- Lists of persons met

- Itinerary

- Issues and findings

- References

Prepare proformas for all of these and distribute them to the team with the instruction to keep them up to date and to return them to you every few days. You can give different team members the task of collating different item and keeping the full team itinerary etc.

This ensures that you avoid the problem of having to reformat the information that they send you and that you time at the end of the mission can be devoted to analysis rather than to administrative work.

The golden rule for all evaluation is that no item of data should be typed in more than once.

INTERVIEWS

Key informant interviews are the backbone of data collection for humanitarian evaluation. Evaluation interviews typically last from 45 to 60 minutes (and longer in some cases). All interviews and similar interactions should be listed in the report so that readers can easily see what the key findings are based on.

SURVEYS

Surveys can take a long time to prepare, conduct, and process. Prior to a survey, the questionnaire has to be agreed and tested, enumerators have to be trained, and then the survey results have to be processed and analysed. A common problem with surveys is to ask so many questions that the results become impossible to analyse in any depth.

Sometimes, it may be possible to take advantage of monitoring surveys (such as food-basket monitoring) or even to undertake simple monitoring surveys on the evaluators own datasets. While it is not ideal to try and generalise from such small surveys - as can be done with a well designed survey - they can still provide useful indications of areas which need further investigation.

DOCUMENTARY RESEARCH

The level of documentation available depends on the stage at which the evaluation. Evaluations after a response has been running for a year may have a large range of analytical documents to draw on. Earlier evaluations may be restricted to agency situation reports and other primary accounts. However, the sheer volume of these can make them a difficult research resource.

An example of the post-disaster volume of reports

Documents posted on reliefweb after a disaster
Data for the 2005 Pakistan Earthquake
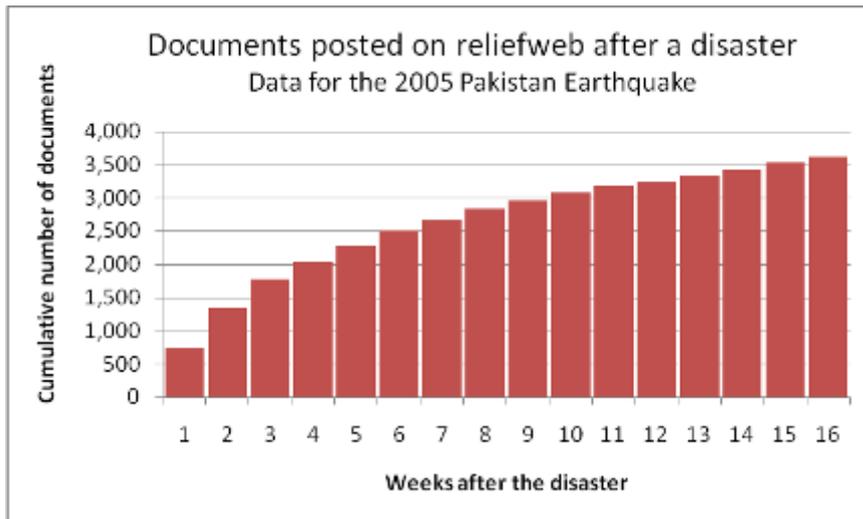
One way through the document maze is to concentrate on series of updates from donors, such as those often provided by USAID or DFID. These often provide a good summary overview of how the situation has developed from the donor viewpoint. For gaining a quick understanding they may be more useful than detailed situation reports.

If you are going to be dealing with a large document set, it can be worth investing in software that will index the documents and quickly find items of relevance. Such software can also enable you to conduct keyword searches in documents to see how these vary by time or across document sources. One piece of software that can do this is dtSearch (http://www.dtsearch.com).

AFTER-ACTION REVIEW

An after-action review is not a common tool in evaluations. After-action reviews are sometimes seen as an alternative to evaluation for organisational lesson learning. However they are also a good tool for humanitarian evaluations as they

may identify learning that would not emerge in key informant interviews as staff have not yet had time to reflect on their experience and gain explicit rather than implicit learning from their experience. They also help to emphasise that staff in the field are sources of learning and not just data-fodder for the evaluators.

An after-action review will typically take half a day or more. It may be difficult to convince management of their utility. However, conduction an after-action review at the start of an evaluation helps to focus the evaluation on the issues of concern to the field staff. It can also help to highlight key issues at the start, allowing interviews to focus on these.

BENEFICIARY CONSULTATION

Beneficiary consultation is often a weak area of humanitarian action. Field managers are often completely overloaded with all their tasks and may have very limited contact with beneficiaries. The evaluation team has far more opportunity to meet with beneficiaries and to get their views. It is essential that humanitarian evaluation teams engage in beneficiary consultation as this is one area where they can demonstrate that they are bringing real value to the field programme.

Consultation can take various forms, from key informant interview through to general meetings and focus group interviews. As with key informant interviews, preparation is essential to make the most out of the meetings and the team should prepare a topic guide for the meeting. Details of all meetings should be presented in the evaluation report, to demonstrate how extensive the beneficiary consultations were that underlie the team's conclusions.

OBSERVATION

Observation can play a particularly important role in humanitarian evaluation. In the early stages of the response, both the affected population and staff are still

This photograph (by Daryl Martyris) of a family's whole possessions in Mozambique was used to illustrate the point during feedback presentations that poverty rather than flooding was the critical issue for many families. Almost all the goods were post-flood donations.

learning about the context, and the evaluators may gather directly from observation that would not be available indirectly through key informants. Later on, observation is a very useful source of triangulation.

Photography is a very useful adjunct to observation for the real-time evaluator, as it allows the evaluator to show others what has been observed. Even if photographs are not used in the report, they are still very useful for briefing presentations.

FOCUS GROUP DISCUSSIONS

The term focus group interview is often incorrectly used for a wide range of group meetings that lack the structured approach of a true focus group

Focus group Discussions are facilitated in-depth discussions on three to four topics for six to eight people where the moderator guides the discussion based on a prepared discussion guide. The intent of focus group interviews is to promote self -disclosure among participants about what they really think and feel.

interviews.

Focus groups can provide evaluators with qualitative information on a range of issues. A facilitator guides six to eight people in a discussion of their experiences and opinions about a topic. The facilitator works from a pre-prepared discussion guide or a question trajectory and uses probing questions to elicit information. Focus group meetings typically last about two hours and cover two to four topics.

They are more important for evaluations that do not have beneficiary surveys or where a particular topic needs to be investigated in depth. However Humanitarian Evaluation focus groups interviews are different from the traditional model in that evaluators are rarely going to be able to sustain the normal ratio of one day or more of analysis per hour of focus group discussion.

TRIANGULATION

Given that humanitarian evaluators aspire to have their findings utilised directly, it is important that research results are triangulated so that they can clearly be shown to be based on more than one type of data source. Triangulation gives the evaluation authority.

The need for triangulation means that the team need to ask questions even when they think that they already know the answer. Sometimes, they will find that the answer is different from what they expect.

Triangulation can take be based on:

- Method triangulation: Comparing the results from different methods e.g. comparing between:
  - Key informant interviews
  - Observation
  - Group interviews

- o Survey

- o Focus group interviews

- o Documentary research

- Source triangulation: Comparing information from different sources e.g. comparing information from beneficiaries with that from Cordaid staff, with partner staff, with other stakeholders.

- Researcher triangulation: Comparing the analyses developed by the locally-based researchers. Drafts of the report will be shared with the team to ensure this form of triangulation.

- Geographic triangulation: Comparing the information from different localities studies to distinguish between broad generic issues and location specific ones.

- Treatment triangulation: Comparing the information from areas where the same problem has been addressed by different types of assistance or through different agencies. For example, you might compare weekly with monthly food distributions to see whether issues that came up were generic or specific to the distribution modality.

PLANNING THE FIELDWORK

Planning the teamwork needs to be done in a careful but flexible manner. It needs to be flexible because humanitarian response take place in a dynamic environment and a rigid fieldwork plan can sometimes prove unsuitable.

However, the fieldwork also needs to be well planned, so that it allows the team to develop their findings and conclusions in a considered way. In particular the following should be planned for:

- Time at the start of the evaluation for all the team to agree the approach that the team is going to take.

- Some reflection time at the midway stage so that the team can develop its initial analysis.

- Opportunities for reflection at the end of the fieldwork so that the team can discuss and agree their findings, conclusions, and recommendations.

It is vital that the fieldwork plan be agreed with the country team, and that it respects their concerns about over-burdening field staff.

An example of a work plan from an inception report. This was a lengthy evaluation ranging over a wide area, but with a requirement to submit a draft report before leaving the country.

Description of activity	ISO Week Number
1 Background Research	52
2 Background Research	1
3 Travel to Delhi	
4 Team meeting	
5 Initial Briefing	
6 Document research	
7 Meetings with stakeholders in Delhi	2
8 Sub-team 1 travel to West Bengal	
9 Sub-team 1 fieldwork in West Bengal	3
10 Sub-team 2 travel to Kerala	
11 Sub-team 2 fieldwork in Kerala	
12 Sub-team 1 travel to Sikkim	
13 Sub-team 1 fieldwork in Sikkim	
14 Sub-team 2 travel to Mumbai	
15 Sub-team 2 fieldwork in Gujarat	
16 Joint team travels to Assam	4
17 Field work in Assam	
18 Joint team returns to Delhi	
19 Sub-team 2 team visits another city	
20 Field work in additional city	5
21 Sub-team 2 team returns to Delhi	
22 Team leader follows up in Delhi	
23 Follow up and preparing draft	
24 Preparing draft report	6
25 Presenting draft at debriefing	
26 Depart Delhi	
27 Submit final draft for comments	7

THE SNOWBALL TECHNIQUE

The snowball technique refers to the technique used by evaluators to ensure that they interview all the available key informants. The evaluators start with a list of a small group of key informants, selected for the their "information richness" and

Figure 1: Initial Nucleus of key informants.

then asking them who else the evaluation team should interview.

The persons suggested by this initial nucleus are then interviewed and they are asked for further suggestions. The new names are then interviewed (Figure 2)
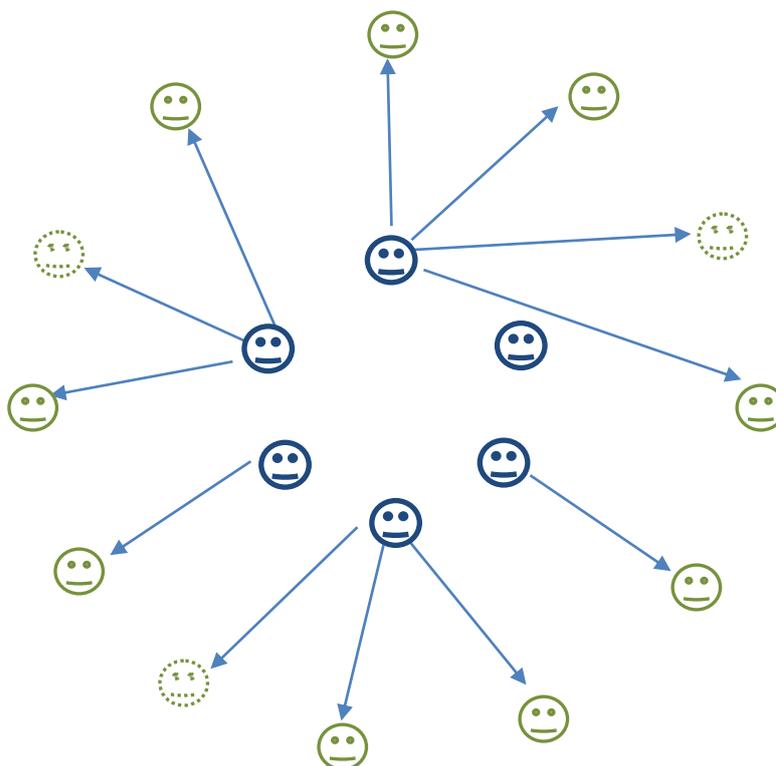
Figure 2: Second level key informants suggested by the members of the nucleus during interviews. Only the first reference to each person is shown, references to those in the nucleus or already on the list are not shown.

The process is continued with the next level of informants, building up a whole shell of informants. With time, more and more suggestions will be for people

already interviewed. The number of times that a particular individual is suggested is an indicator of their role within the response.

Eventually, interviewees will not be able to suggest anyone who is not already on the list - the snowball techniques is said to have reached saturation at this stage (Figure 3). Figure 3 is a simplification, as some of the suggested names will not be available for interview because of absence or time conflicts. Nevertheless the technique is quite powerful and is particularly useful in circumstances like humanitarian responses where there is rarely an up-to-date list from which interviewees can be sampled.
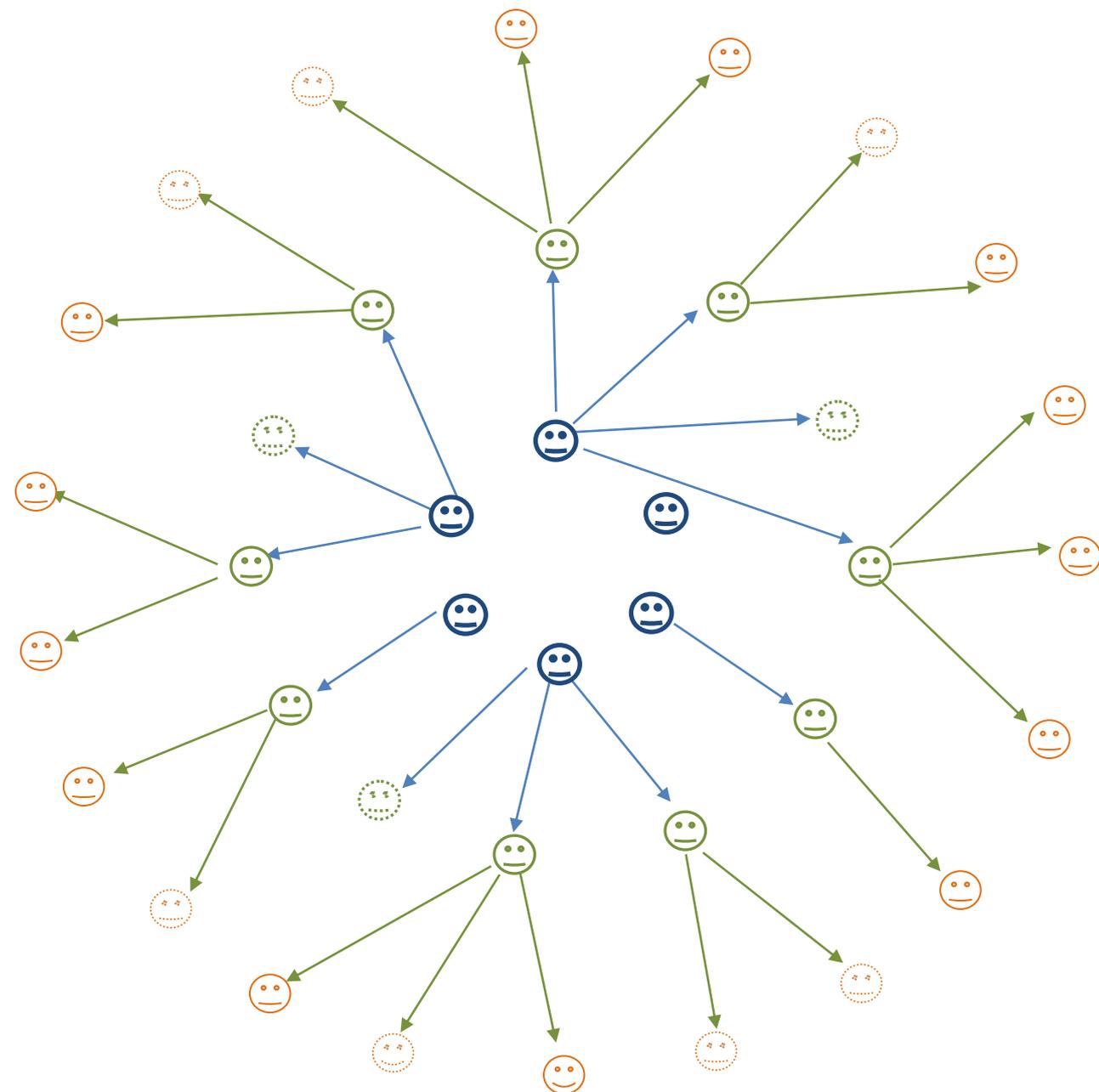
😐 Initial nucleus of key-informants

😐 Second order key-informants (unavailable informants shown dotted)

😐 Third order key-informants (unavailable informants shown dotted)

Figure 3: The final set of interviewees.

CHECKLIST FOR PLANNED EVALUATION METHODS:

Questions about planned evaluation methods

❏ Are the planned methods clearly set out?

❏ Are the planned methods likely to be able to answer the evaluation questions?

❏ Do the planned methods fit in with the available evaluation team time and resources?

❏ Do the planned methods fit in with the work-load of those being evaluated?

❏ Do the evaluation team appear to have the skills to use the planned methods?

❏ Are the planned methods in line with what was specified in the ToR?

❏ If not as specified, is the reason for the difference clearly explained?

❏ If not as specified, will the alternative methods be as effective as those originally specified?

❏ Will this team produce credible conclusions using these methods?

❏ Will the methods provide the type and quality of evaluations findings needed by stakeholders?

❏ Will the methods provide robust conclusions that are valid and reliable?

❏ Are the planned methods ethical?

❏ Are potential constraints, and prerequisite conditions, for the use of these methods clearly identified?

❏ Is the analysis procedure clear for the proposed methods?

❏ Is there a clear link between the proposed methods and how the team plan to develop their conclusions?

❏ Will the methods promote utilisation?

DRAWING UP YOUR INTERVIEW GUIDE

Using a topic guide is useful to ensure that you cover your key areas of interest in the interview and don't get so side tracked that you don't ask the questions that you need answers to so that you can in turn answer the evaluation questions. Your topic guide can be expected to have three types of questions.

- Warm-up questions.

- Questions based on the terms of reference.

- Wrap-up questions.

WARM-UP QUESTIONS

Before any questions it is important to state the evaluators' names, what the purpose of the evaluation is, what the interview rules are, and possibly, how long the participant can expect the interview to take. An interview protocol card (which can be business card size) can usefully set out the interview rules for interviewees who can read in the language of the card..

An interview protocol card.

YOUR RIGHTS AS AN INTERVIEWEE

➢ You have the right not to be interviewed or to terminate the interview at any time.

➢ You have the right not to answer any question.

➢ Nothing you say will be attributed to you directly or indirectly without your explicit permission.

➢ The notes on this interview will not be shared outside the evaluation team.

➢ If you provide an email address, we will send you draft of the report for your comments.

Warm-up questions held to break the ice at the start of the interview. Warm up questions should be factual rather than about opinions, and should be answerable by all interviewees. One of the most effective warm-up questions is about the role of the interviewee in the response. It seems that we all like talking about ourselves (a pitfall that the evaluator has to be careful to avoid - as with one colleague who, if not controlled by his interview partner, would launch into long stories about his own experience). A second warm-up question can be about the interviewees previous experience - this can be useful for asking the interviewee later in the interview to compare and contrast the present response with other responses.

QUESTIONS BASED AROUND THE TERMS OF REFERENCE

The bulk of any interview will be aimed at answering the questions raised in the terms of reference. Not all questions will be equally applicable to all interviewees,

and it may be useful to prepare different topic guides for different types of respondents.

When drawing up your interview guide it is crucial to draw a distinction between:

- The evaluation questions that you are trying to answer; and

- The interview questions that you put to interviewees.

Rubin and Rubin (2005) classify the questions that you put to interviewees as being of three types:

- Main questions. These should generally ask interviewees about their own direct experience and give them space to answer. More experienced interviewees can be asked evaluative questions (that ask them to value some aspect under study). However, this should be in the context of identifying key bits of underlying data rather than just establishing opinions, unless the evaluation questions are about what opinions and perceptions are held.

- Probes and control questions. These help to channel the interview exchange to the area of greatest interest.

- Follow-up questions. Questions that deepen understanding or expose something new that has cropped up in the interview.

The topic guide is a guideline rather than a questionnaire. The order of questions or the text of a question can be changed to suit the particular interview. More contentious issues should normally be left until the end of the interview.

WRAP UP QUESTIONS

Wrap up questions should include both questions about general learning, as well as questions about contact details such as the interviewees email address. Question about learning can take the form of direct or indirect questions such as:

- *When you look back on the response, what it the biggest lesson that you have learned, or had reinforced, by this experience?*

- *What have you personally learned from this experience?*

- *What was the thing that most surprised you in this operation?*

- *If you were back at the start of the operation with the knowledge you have now, what would you do differently?*

- *Is there any question that you were expecting which I have not asked?*

- *How do you see this response developing?*

- *Who else should we talk to?*

The wrap up should deal with any admin details, such as the email address for sending the draft report to.

The following is an edited version of the topic list used for semi-structured interviews with aid agencies for an evaluation of the response to the 2007 Mozambique floods. A different set of questions was used for beneficiary

interviews and yet another for focus groups. These questions are only a guide to the evaluator, who may drop some questions, ask them in a different way and change the order to suit the interview. The answers to some questions may lead to requests for clarification or confirmation from the interviewer.

Introduction of the subject of the evaluation, Chatham House rule etc.

Warm-up: What role did you play it the response?

1. About the Cluster Approach
   a. What do you know about the Cluster Approach?
   b. What role have you played in clusters? Outside clusters? If outside, reason?
   c. Which cluster is generally regarded as the best? What is the reason for that?
   d. Which cluster generally regarded as the weakest? What is the reason for that?
   e. What has the level of participation in clusters been like?
   f. For the clusters your most involved in, has the cluster approach
      a. helped fill sector or geographic gaps in the mobilisation of humanitarian agencies
      b. enhanced partnerships between UN/RC/NGOs and Govt on the ground
      c. improved strategic field-level coordination and prioritization
      d. made operational partners accountable to the Humanitarian Coordinator.
   g. For specific clusters you are involved in, has the cluster lead delivered on:
      a. coordinated the analysis of needs and information sharing
      b. securing and following up on commitments to respond
      c. acting as provider of last resort
      d. managed quality control in the cluster
      e. acting impartially without favouring their own operational role
   h. What is the biggest constraint the cluster approach has faced? Any way around?
   i. Have the clusters added anything to the response? What?
   j. How well has the Resident Coordinator functioned as the link between clusters and the Govt?
2. Funding
   a. Has the Flash Appeal and CERF contributed to a timely response?
   b. What affect, if any, have they had on your funding?
   c. How coherent has fundraising by the Humanitarian Country Team been?
3. The response
   a. Was the response to Cyclone Favio faster than for the Floods? If so or if not, Why?
   b. What problems did you have in shifting from development to relief? Ditto the UN?
   c. How does this response compare to previous such emergencies?
   d. Have you seen any impact from prior investment in Disaster Risk Reduction?
   e. Is the response to this emergency leaving people better prep'd for the next? How?
4. Quality

a. What needs of the affected population have been best met? Least met? Why?

b. What was the quality of needs assessments? Were they shared in the clusters?

c. How well have agencies worked with local institutions?

5. Learning

a. What would you do differently the next time? What lessons have you (re)learned/?

b. How appropriate is the cluster approach for a country with a strong government?

c. How could the cluster approach be improved?

d. What lessons have you learned, or had reinforced by the response?

e. What would you do differently in any future response like this?

f. Given what we are looking at, is there anything else we should have asked you?

6. Closing: Who else would you recommend that we talk to?

REFERENCES AND RESOURCES

Drever, E. (2003). *Using semi-structured interviews in small-scale research: a teacher's guide* (revised ed.). Glasgow: Scottish Council for Research in Education

Holstein, J. A., & Gubrium, J. F. (1995). *The active interview* Thousand Oaks: Sage

Kaiser, A. (1979). *Questioning Techniques: A practical guide to better communication* (N. M. Lechelt & U. Marten, Trans.). Farnham: Momenta Publishing

Kvale, S., & Brinkmann, S. (2008). *Interviews: An introduction to qualitative research interviewing* (2nd ed.). Los Angeles: Sage

McCracken, G. D. (1988). *The long interview*. Newbury Park, Calif.: Sage Publications

Merton, R. K., Fiske, M., & Kendall, P. L. (1990). *The focused interview: A manual of problems and procedures*. New York: Free Press

Rubin, H., & Rubin, I. (2005). *Qualitative interviewing: The art of hearing data* (2nd ed.). Thousand Oaks: Sage

USAID. (1996). *Conducting Key Informant Interviews* (Performance monitoring and evaluation tips: 2). Washington: USAID Center for Development Information and Evaluation. Last viewed on 20 September 2008 at: http://www.usaid.gov/pubs/usaid_eval/pdf_docs/pnaby233.pdf

Weiss, R. S. (1995). *Learning from strangers: The art and method of qualitative interview studies*. New York: Free Press

FACILITATING AN AFTER-ACTION REVIEW

After-action reviews can range from an informal self-examination to a formal after an event. Our focus here is on formal facilitated after-action reviews.

> **After-action review:** a structured discussion of an event the focus on drawing learning from that event. It looks at the divergence between the planned and the actual, and then identifies what went well or ill. The aim of an after-action review is to improve personal and collective performance in the future by identifying lessons from particular events.

After-action reviews can only work if there is full participation by the group in an open, learning atmosphere.

The role of the facilitator is to take the process forward and to maintain the learning nature of the event by avoid detours into self-justification or blame, while able to keep the discussion focused on emerging issues.

The first task for the facilitator it to create the right climate, emphasising that the focus is on learning and that what is said is confidential, in so far as it will not be attributed directly or indirectly to whoever said it. The climate needs to be one of trust and openness.

The facilitator needs to emphasis the professional and candid nature of the review, and that it is not a complaint session or an evaluation, but concentrates on what the different participants have learned during the operation.

In many cases this learning will not be explicit because people have not had a chance to reflect on it. It is the job of the facilitator to draw this learning out into the open.

THE FOUR KEY QUESTIONS

The four key questions for any humanitarian after action review are:

1. What did we plan? What was expected to happen?
2. What actually occurred?
3. What went well, and why?
4. What can be improved, and how?

WHAT DID WE PLAN? WHAT WAS EXPECTED TO HAPPEN?

The first question illustrates the military origins of the after-action review technique. The military invests a great deal of effort in planning whereas humanitarian action tends to be ad hoc. So asking what was planned is not such

Ask participants to individually cast their minds back to the start of the operation, then ask them to record what was planned and what they expected at that time briefly on a card.

Then ask participants to think back to when their expectations first changed in a major way and ask them to put their new expectations on a dated card. Have them do this for two to three changes of expectations or plans and then put the cards on flipchart paper on the walls with a timeline on it .

an appropriate question in humanitarian response.

Humanitarian operations tend not to be planned in a formal sense but to be managed by staff along the lines of the heuristic rules they have developed from their experience of other emergencies. One way of drawing out lessons is not to compare the planned with the actual, but the expected with the actual.

It is important that the after-action review be relaxing and a bit of fun to ensure full participation.

Ask anyone if they can think of anyone's expectations that are funny in retrospect - such as someone expecting that the operation would be over quickly enough for them to go on holiday. Use humour to break the ice.

The purpose of asking people about what was planned or expected is to remind them about how plans and expectations changed over time - such changes often reflect implicit lessons that the after-action review should draw out explicitly.

WHAT ACTUALLY OCCURRED?

The aim here is to get staff to construct a timeline for what happened. This serves a number of purposes, including providing a timeline for the evaluator, but as with the first question it helps to encourage people to look back and see what has changed since they began.

Have participants in small teams prepare a timeline (or a flowchart if they prefer) for what actually happened. You may need to go back to the plan timeline at this stage.

WHAT WENT WELL, AND WHY?

The two previous questions (what was planned and what happened) should have

You may find that this works best with all the participants seated in a circle around a standing facilitator. Ask what when well, and have these noted by a helper on a flipchart outside the circle. To check that you have understood, rephrase the point as a lesson and see if you have agreement on it.

laid the ground work for this question.

WHAT CAN BE IMPROVED, AND HOW?

Next we need to look at the other side of a coin. This can be a bit tense as staff may be afraid to identify areas of weakness that might be seen as criticisms of colleagues or that might reveal weak spots in their own performance. The

facilitator needs to emphasise that emergencies are very difficult contexts to work in as one cannot get it right all the time when one has to take decisions with poor information.

It can sometimes be difficult to have staff identify this as identifying problems may be seen as a criticism of others. The facilitator can also help create the right atmosphere by referring to some instance where he or she did not get it right and learned from it.

Another approach is to ask everyone to write down their rating (out of ten) for the response, and then ask them *"what would have made it a ten for you?"* (or an eleven is someone has rate it as a ten).

List these responses and then use this list to ask why these issues were problems and how they could be improved. Rephrase this as a lesson for anyone facing a similar situation and see if you have agreement for your reformulation.

Once you have your learning points, you should put them up on flipcharts.

Check to see if people have come up with other ideas during the workshop it is a good idea to try and capture these by revisiting the lists and asking people if they would like to make any changes or additions.

The facilitator can also check how the participants view the validity of any lessons generated.

At the end of the workshop, place the lessons that have been generated on flipcharts and then ask everyone to vote for what they regard as the three most important lessons.

During an evaluation, holding an after-action review as soon as possible is often useful as the evaluation team can then use the lessons identified as starting points for further questions.

"One of the lessons identified at the after-action review was the need to get staff transport sorted out more quickly. Do you think that this was a problem (and why)? What do you think the agency could do to improve this in the future?

USAID has published a useful guide for conducting after-action reviews (USAID, 2006). There is also useful short guidance on the UK's National Library for Health website (Robertson and Brún, 2005). There is also a comparative study of AARs after the Southern Africa Crisis of 2002-2003 (Sexton and McConnan, 2003) that

provides good practice guidance on the effective design and use of AAR processes .

REFERENCES

Robertson, S., & Brún, C. D. (2005, 12 January). National Library for Health: Knowledge Management Specialist Library: After Action Management Reviews.   Retrieved 16 September 2008, from http://www.library.nhs.uk/KnowledgeManagement/ViewResource.aspx?resID=70306&tabID=290&summaries=true&resultsPerPage=10&sort=TITLE&catID=10403.

Sexton, R., & McConnan, I. (2003). A Comparative Study of After Action Review (AAR) in the context of the Southern Africa Crisis. London: ALNAP. Last viewed on 23 June 2009. URL: http://www.alnap.org/pdfs/other_studies/aar_case_study.pdf.

USAID. (2006). *After-action review (Technical Guidance PN-ADF-360).* Washington: USAID.. Last viewed on 16 September 2008. URL: http://events.fcw.com/events/2007/KM/downloads/USAID_Wallace_The%20After%20Action%20Review.pdf

Focus group discussions

Introduction

Focus group interviews are good for investigating topics in depth. They are particularly useful for getting past the polite responses to questions about how the usefulness of assistance. However, unlike key informant interviews, where the issues of concern may emergency from the interview, the need to prepare a detailed discussion guide beforehand, and the limited range of topics that can be covered mean that the evaluation team should already have identified some of the key issues before engaging in such interviews.

Humanitarian evaluation can pose other problems for focus group interviews. A normal requirement for Focus Groups s is to have a comfortable environment where people can relax and concentrate on the issues under discussion. Relaxation is key to achieving the kind of self-revelation that focus group interviews strive for. Achieving a relaxed and comfortable atmosphere may be quite difficult under field conditions, especially in the early stages of a response.

A final problem is that of analysis. Traditionally, focus group interviews are recorded and transcribed for analysis, with a considerable time devoted to analysis. This is not realistic in the context of the typically limited time-frames and budgets for humanitarian evaluations. Here the analysis must accompany the interview itself with only limited opportunity for follow up analysis.

Two basic approaches

There are two fundamental approaches to guiding Focus Group Discussions. The first is to prepare a topic guide, the second (recommended for humanitarian evaluations) is to prepare a questioning route that sets out the questions to be answered in more detail.

Preparing the discussion guide

The first step is to identify your topic, and whom you should seek to interview to illuminate it. Nest you will need to prepare your focus group discussion guide. The following section contains an example of a discussion guide.

Examples in the discussion guide help to channel the thinking of participants, so need to be used with care. Traditional focus group discussion guide avoid providing categories to participants but develop the categories from the discussion.

You can only deal with 2 to four topics in a typical focus group interview. You can start the discussion on each topic with a round robin question, asking each member in turn to respond to the question.

This is followed by probing questions, a wrap up round robin question followed by a summary and a quick check that nothing has been missed.

PREPARING A QUESTIONING ROUTE

A questioning route for a focus group discussion should have a number of stages, that contribute to the development of a open atmosphere that promotes the free exchange of experience.

- The Opening Question is basically to get talking and to get them to feel comfortable. It should therefore be a question that everyone can answer, and involves facts rather than opinions. It is important to avoid questions about length of experience or position that might create perceptions of status differences that could inhibit later discussion.

- The Introductory Questions are to get the participants to start thinking about the topic. Typically the ask people how they see or understand the topic under investigation.

- Transition Question move the conversation towards the key questions for the study. They set the stage for productive key questions, often going deeper into the introductory questions and linking the participants with the topic

- Key Questions begin about one third to one half way through the discussions. These are the main questions for the focus group and are usually the first ones prepared by the evaluation team. Typically there are two to five questions in this category. They are given the most time and the facilitator should probe to ensure that the questions are fully discussed.

- Ending questions are intended to provide closure to the discussion. They can enable participants to reflect on previous comments. The "all things considered" asks participants for their positions on some key issue that was discussed but which for which there were conflicting viewpoints. The summary question is an alternative to this. Here the facilitation presents a brief summary of the discussion and asks if the summary adequate captures what was said in the discussion. The final question is an insurance question - after presenting a quick overview of the purpose of the discussion the facilitator asked "Have we missed anything?".

SELECTING YOUR PARTICIPANTS

The participants in any group should be relatively homogeneous. Participants should be from similar socioeconomic background and broadly share a common experience of the response. Men and women should be consulted in separate focus groups, similarly displaced and host population, adults and youth.

RUNNING THE FOCUS GROUP

Each focus group demands a facilitator and a recorder. The facilitator moderates the discussion and keeps it focused on the area of interest, while the recorder notes what is said, You can record the focus groups, this is essential if you want to undertake further analysis, but can also be useful if the facilitator and moderator disagree about what was said during the discussion.

The facilitator should begin the session by establishing rapport with the group. Introduce yourself as facilitator and introduce whoever is recording. Let participants introduce themselves. In many cultures some initial hospitality can help to establish a relaxed atmosphere. Provide tea and biscuits or whatever refreshments are traditional in that society to the participants.

It is useful to provide tea standing up and to avoid having people sit down until you have made an assessment of the group members. Ideally you want to place the most talkative people close to the facilitator (to give you maximum control over them) and the least talkative opposite the facilitator (to enable you to give them the maximum encouragement to contribute.

Once seated, you can explain the purpose of the interview, the sort of information you are looking for, how they were selected, and how the information will be used. If you are going to take the session, ask for permission.

Explain that participation is voluntary and that whatever anyone says will not be attributed to them.

The task of the facilitator is to:

- steer the discussion between the participants by;
  - ➢ Injecting new questions
  - ➢ Verbal cues and body language
  - ➢ Repeating the question
  - ➢ Broadening and narrowing discussion through asking questions
  - ➢ Take advantage of pauses to politely cutting off the over talkative
  - ➢ Encourage the less talkative by directing questions at them
  - ➢ Open up topics for the group by asking "does anyone else feel the same wary, or have a different view?"
- probe to bring out the meaning by;
  - ➢ Rephrasing the question
  - ➢ Rephrasing answers
  - ➢ Asking when, what, how, where, who, and which to deepen topics
  - ➢ Asking for details or examples
  - ➢ Prod with "anything else"
  - • Summarising and asking if the summary is right.

The recorder should record:

- Date, time, place
- Number and characteristics of the participants
- General description of the group dynamics (level of participation, presence of a dominant participant, level of interest)
- Opinions of participants, with a particular eye for strong quotes that could be used to illustrate those opinions in the report. Non-verbal opinions should be recorded also (as when group members' body language or non-verbal gestures indicate strong interest or agreement with what is being said.

- Key themes emerging from the discussion.
- Any relevant spontaneous discussion between the participants outside of the formal session.

If you have a three person team you might use two people to record what is said.

ANALYSIS

Immediately after the discussion ends, the recorder should read the notes back to the facilitator and between them they should deal with any differences of memory or understanding. The team should confirm that they have a common understanding of what was said. They should also identify:

They should then discuss if anything that was said surprised them, if the discussion has changed their view on any topic, or how it has opened up new issues for further research in the evaluation. They should consider if any of the issues raised can be triangulated through interviews, research or observation. In particular they should particularly consider if any points should be amended in the discussion guide for further focus groups, or in the interview guide for key informant interviews.

PAYMENT

In market research, focus group members are paid for their time - this might be inappropriate in the humanitarrian context, because agencies normally do not pay focus group participants. It might be appropriate however to give small gifts of a symbolic nature (such as a meal, refreshments, a tee-shirt, a cap, or a pen set) to focus group participants to acknowledge that they have given their time.

RESOURCES

Krueger gives a good short step by step guide to focus group interview (2002), and there is an older USAID guide on focus group interviews that is more orientated to the aid context (USAID, 1996). Kruger and Casey (2009) provide detailed guidance on all aspects of focus groups, including the analysis of data.

Krueger, R., & Casey, M. (2009). Focus groups: A practical guide for applied research (4th ed.). Thousand Oaks: Sage.

Krueger, R. A. (2002). *Designing and Conducting Focus Group Interviews*. St Paul: University of Minnesota. Last viewed on 20 September 2008. URL: www.shadac.umn.edu/img/assets/18528/FocGrp_Krueger_Oct02.pdf

USAID. (1996). *Conducting Focus Group Interviews* (Performance monitoring and evaluation tips: 10). Washington: USAID Center for Development Information and Evaluation. Last viewed on 20 September 2008. URL: http://www.usaid.gov/pubs/usaid_eval/pdf_docs/pnaby233.pdf

A SAMPLE QUESTIONING ROUTE

INTRODUCTION

Prior to the discussion, introduce yourself as facilitator and introduce whoever is recording.

Explain that:

- We are here today because we want to better understand how the funding agency's partners understand Community Managed Disaster Risk Reduction and how they work with it.
- There are no right or wrong answers, only differing views and that differing views flow from the different viewpoints and experience of each one of us.
- The role of the moderator is to guide the discussion and that participants should address their comments to each other.
- That whatever is said will not be attributed to the person directly.

Remember to explicitly get permission for recording the meeting, and explain why you want to record and who will have access to the recording.

Question type	Question	Notes
Opening	Tell us your name and a little bit about the communities that you work with.	Warm-up question
Introductory	How did you first hear about Disaster Risk Reduction (DRR)	To get people thinking about DRR and their own relationship to it
Transition	What do you think of when you hear the term Disaster Risk Reduction?	Getting people to think about what DRR means
	What does Community Management mean?	Getting people to think about what DRR means
Key question	What is different about the Community Managed Disaster Risk Reduction CMDRR from other approaches that you have used?	Possible use a flip-chart to list - look for consensus and divergence

Question type	Question	Notes
	How does the CMDRR compare with other approaches in terms of promoting community involvement?	Probe for in what way etc
	What is the biggest advantage of the CMDRR approach?	Probe for whether they are using the approach in other programmes.
	What is the biggest disadvantage with the CMDRR approach?	Prove for issues. These may include speed, the level of training need, the cost of meeting community plans etc.
Ending	If I understood the discussion correctly, the main points are: ………. Is this a good summary, is there anything that I am missing.	You could use the assistant here to present the summary.
	What one message could you to the funding agency about their CMDRR approach and programme.	Get everyone to comment
	We wanted to find out about partners understand CMDRR and their experience with it. Is there anything that we have missed?	Get everyone to comment

A SAMPLE FOCUS GROUP DISCUSSION GUIDE

INTRODUCTION

Introduce yourself as facilitator and introduce whoever is recording. Let participants introduce themselves.

Explain that:

- We are here today because aid agencies want to find out how they could have better served the population at the beginning of the response.
- You are particularly interested in:
    - What types of aid are the most useful (and that this can vary by person)?
    - Who were the most importance sources of assistance?
    - Who benefited most from the assistance?
- There are no right or wrong answers, only differing views.
- Role of the moderator is to guide the discussion and that participants should address their comments to each other.
- That whatever is said will not be attributed to the person directly.

TOPIC 1: THE USEFULNESS OF TYPES OF AID

Explain that almost all aid is of some use, you may be able to sell it if you don't need it. The interest here is to find out which sorts of aid are the most useful so that in future agencies can provide more of them and less of the less useful sort. This can be a difficult topic as disaster affected populations are often reluctant to identify any type of aid as being less useful.

Q1: Round Robin (asking each group member in turn). Of all the different assistance that you received from any source, (for example being rescued, water, food, shelter, jerricans, blankets, cash, etc - examples deliberately give to explain what sort of aid and services are being considered) which item or service was the most useful?

> *Note: if this question meets the old "everything was useful" answer ask the member if they have a coin or a note, then offer them a smaller denomination coin or note in exchange for it. If they demur ask them why, as from what they are saying they don't see any difference between them. (If they don't demur - ask them for all the money they have in exchange for the smallest coin or note that you have). The point is that while both denominations are useful, one is more useful than the other.*

Main probes:

1. *What made it the most useful assistance?*
2. *How could it have been more useful than it was?*
3. *What sort of assistance would you like to have had more of?*
4. *What was the least useful aid that you got?*

5. *How well did the aid giver know your needs?*

6. *What factors do you think led to the aid giver provided that aid?*

Closing round robin: If you wanted aid agencies to take away one way about how aid could be more useful for you, what would it be?

Summarise the discussion and ask if it is an adequate summary.

Ask if you have missed anything.

TOPIC 2: THE SOURCES OF AID

Q1: Round Robin As I have been travelling around, I see that there are lots of organisations providing assistance. When you think about national and international aid organisations, about national and local government, national and international military, your neighbour, your family, local traders, churches and others, which of them has been the most important sources of assistance to you?

> *Note: If you meet reticence because people think that ranking aid providers will be insulting to the lower ranked give the example that most of us have two parents, who help us through life. However our mother and our father help us differently at different points in our lives - our fathers cannot breastfeed us and our mothers cannot teach us hunting, but this does not mean that they are any less important because of this.*

Main probes:

1. *What factors made them the most important source of assistance for you?*

2. *Was it the same for everybody else, or was it only some people that they were the most useful for? Why?*

3. *Did the most useful provider change over time? How?*

4. *What were the factors that made particular providers the most useful at different times?*

5. *Given what you know now, would you have approached different aid providers in a different way? How?*

Closing round robin: How has your opinion of the sources of aid changed over the response?

Summarise the discussion and ask if it is an adequate summary.

Ask if you have missed anything.

TOPIC 3: COVERAGE?

Q1: Round robin: No matter what happens some people benefit more than others. If it rains, farmers might be happy, but the woman who has just done her washing may be annoyed. If the sun shines, the man who is doing casual work will be happy, but the woman who is drying fish to take to market will be even happier. So the question is- which group do you think benefited most from the aid?

Main probes:

1. *How do you think that they benefited more than others?*

2. *Was it fair that they benefited more than other people? What are their relative positions like now when compared with before?*

3. *Who benefited most: men or women? Why do you say that? What are their relative positions like now when compared with before?*

4. *Who benefited most: rich or poor? Why do you say that? Can you give an example?*

5. *How much redistribution of aid was there within the community? What motivated this?*

Closing round robin: How do you think that aid could have been distributed more fairly?

Summarise the discussion and ask if it is an adequate summary.

Ask if you have missed anything.

Section five:
Approaches to analysis

Moving from data collection to analysis can be one of the more challenging steps and time consuming steps in an evaluation. This section provides some useful steps, including building the 'chain of evidence'.

DATA ANALYSIS

'The purpose of analysis is to transform the data into credible evidence about the development of intervention and its performance. Typically, the analytical process involves three steps:

1. Organising the data for analysis, i.e. data preparation

2. Describing the data, e.g. generating findings of fact

3. Interpreting the data, e.g. assessing the findings against criteria'

(Extract from: Ministry of Foreign Affairs, Denmark, Danida (2006) 'Evaluation Guidelines')

BIAS

All data collection strategies are subject to the problem of bias. Bias leads to misinterpretation or mistaken analysis that draws its conclusions from information which is not correct, not complete or not representative of the affected population. Anecdotes abound of the effects of bias on evaluation results. To summarise, at a minimum, evaluation resources may be wasted if they are affected by bias. In the worst cases, biased evaluations can cause harm to populations they were meant to help, while also affecting agency reputations. Some forms of bias are the following:

Spatial - Issues of comfort and ease determine the assessment site

Project - The assessor is drawn toward sites where contacts and information is readily available and may have been assessed before by many others

Person - Key informants tend to be those who are in a high position and have the ability to communicate

Season - Assessments are conducted during periods of pleasant weather, or areas cut off by bad weather go unassessed, thus many typical problems go unnoticed

Diplomatic - Selectivity in projects shown to the assessor for diplomatic reasons

Professional - Assessors are too specialised and miss linkages between processes (preceding biases, Chambers, 1983)

Battle - Assessors go only to areas of cease-fire and relative safety. (Barakat and Ellis, 1996)

Political - Informants present information that is skewed toward their political agenda; assessors look for information that fits their political agenda.

Cultural - Incorrect assumptions are based on one's own cultural norms; Assessors do not understand the cultural practices of the affected populations.

Class/ethnic - Needs and resources of different groups are not included in the assessment

Interviewer or Investigator - Tendency to concentrate on information that confirms preconceived notions and hypotheses, causing one to seek consistency too early and overlook evidence inconsistent with earlier findings; Partiality to the opinions of elite key informants.

Key informant - Biases of key informants carried into assessment results

Gender - Assessors only speak to men, or male interviewers survey women, or vice versa.

Mandate or speciality - Agencies assess areas of their competency without an inter-disciplinary or inter-agency approach.

Time of day or schedule bias - The assessment is conducted at a time of day when certain segments of the population may be over- or under-represented.

Sampling - Respondents are not representative of the population.

(Source: *Monitoring and Evaluation Guidelines*. A training module prepared for the University of Wisconsin-Madison Disaster Management Center by InterWorks.)

THE CHAIN OF EVIDENCE

The following idealised diagram shows the way in which recommendations should be based on one or more conclusions, and that conclusion should themselves be based on findings.

Findings should be based on different pieces of evidences, ideally even on pieces of evidence from different evaluation methods.
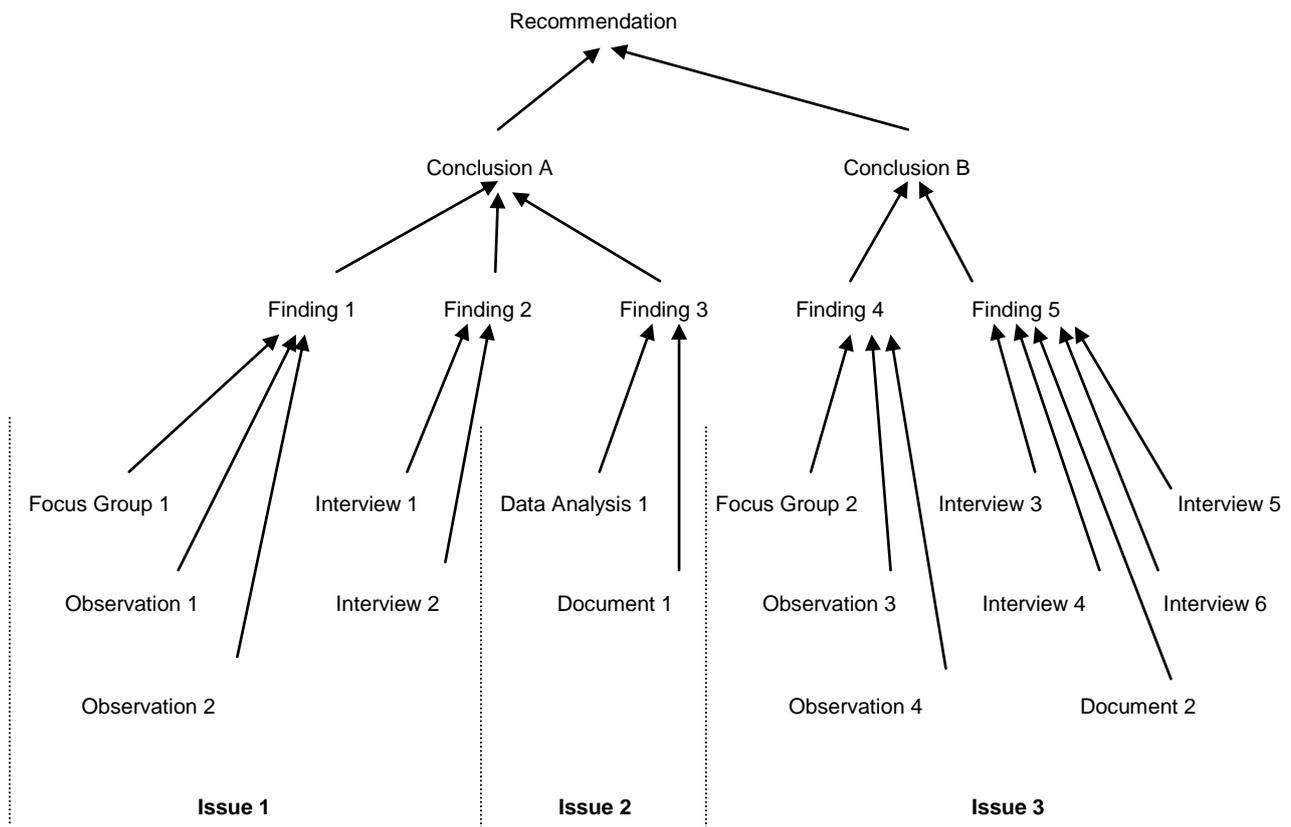
USING A TOOL TO LINK EVIDENCE TO RECOMMENDATIONS

One approach is to use a simple tool to tabulate issues, findings, evidence, as the work proceeds, and then to discuss this as a team to develop conclusions and recommendations. The example below is part of such a tool that was used in an evaluation of a Disaster Risk Reduction programme in India.

The tool is normally done on a spreadsheet with each piece of evidence on a different row. Each finding may be supported by several different pieces of evidence, each conclusion by several different findings, and each recommendation by a number of conclusions.

Issues are the "foreshadowed problems" that occur to the team following reading of the background documents, briefings, or fieldwork. The list of issues will grow during the evaluation as more of them raise their heads.

Findings are what you found out about the issue.

Evidence is the evidence that findings are based on. Evidence can take the form of interviews, direct observation, documents or other data. As the evidence may take the form of individual interviews made under the Chatham House Rule, this tool is not included in the outputs but is only shared within the team.

Conclusions are the result of the analysis of findings.

Recommendations are the actions that are recommended in response to the

Part of an issues/recommendations sheet for an evaluation of a DRR programme in India.

Issue	Evidence	Finding	Conclusion	Recommendation
Volunteer skill level	1. Trainer at mock drill five did not push head back far enough to clear tongue for airway in mouth to mouth demonstration. 2. One third of rope joining knots shows in mock drills were granny knots (and therefore dangerous). 3. Most common resuscitation method	Volunteer and trainer skills need polishing	Training needs a quality control mechanism.	The agency should within six months set up a quality control mechanism to ensure that training is of good quality and that skills are kept up to date.

conclusions.

A tool like the one shown is only useful if it is used throughout the evaluation, with team members adding in issues, findings, and evidence as they come across them. Conclusions should ideally be established through analysis of the findings.

This spreads the analysis of the data throughout the evaluation. One purpose of the tool it to help you identify when your data collection is reaching theoretical saturation, that is, when no new issues or evidence is emerging from the interviews or other data collection methods.

Emerging issues may even lead to a change in the evaluation approach, changes in tools such as interview guides, or more attentions to particular aspects.

Section six:
Evaluation products

This section provides a guide to the most common written products of an evaluation: the inception report and the final report. It provides sound advice and guidance on writing the final report, as well as good practice tips for designing and formulating recommendations.

THE INCEPTION REPORT

The inception report is a key document in the evaluation. Although inception reports take time they are particularly useful because they help to highlight at any early stage any gaps between the evaluation manager's understanding of the evaluation and the evaluation team's understanding.

- Provide an opportunity for the evaluation manager to assess how the team understand and plan to approach the evaluation.

- Enable the evaluation team to turn the terms of reference into a doable task that is agreed with the evaluation manager.

- Allow the evaluation team to plan their work in a coherent way.

- Provide other stakeholders with a clear statement of intent by the evaluation team so that they can quickly flag up any issues that they have with the proposed approach.

Inception reports provide the evaluation manager with an opportunity to address problems with the teams understanding and approach before they become major issues. The inception report should demonstrate the teams' understanding of:

- The context of the humanitarian crisis.

- The context of the response and of the actions to be evaluated.

- The purpose and intent of the evaluation.

- The concern of stakeholders.

If the context section of the inception report is well written, it can be adapted to be used as the introduction to the context in the main report. The same may apply to other parts of the inception report.

The inception report must also present a clear plan of work which takes the constraints into account, is realistic, and specific about what will happen when. The work plan should include the allocation of roles and responsibilities within the team, any deadlines for intra-team reporting, as well as detailed travel plans.

The inception report should present proposed methodologies, including an initial priority interview plan for further interviews. It should acknowledge and where possible specify the role of the advisory group.

Finally the report should annexe the interview guide for the interviews, and possibly a focus group topic list if the team are planning to conduct any focus group interviews.

INCEPTION REPORT CHECKLIST

Inception report contents

❑ Does the inception report demonstrate a clear brief understanding of the context?

❑ Does the inception report demonstrate a clear understanding of the intervention?

❑ Does the inception report demonstrate a clear understanding of the purposes and intent of the evaluation?

❑ Is there a clear plan of work?

❑ Is it realistic?

❑ Will the plan conflict with any planned programme activities?

❑ Does the plan of work identify the key dates for all activities and outputs?

❑ Are roles within the team clearly set out and responsibilities assigned?

❑ Does the plan demonstrate a clear logical progression?

❑ Does the plan clearly allow for the likely constraints?

❑ Are all planned travel itineraries and dates given?

❑ Is the proposed methodology clear?

❑ Will the methodology proposed and the fieldwork allows the team to answer the evaluation questions?

❑ Have the team given adequate consideration to the role of the advisory group? Have the advisory group accepted the report?

❑ Do the team demonstrate an awareness of the possible concerns of stakeholders?

❑ Have the team annexed their interview guide

PROVIDING INITIAL FEEDBACK

Partly based on the guidelines for oral presentation from (Drever, 2003)

Evaluation teams often need to provide initial feedback to the country team at the end of the fieldwork phase. Typically such feedback is through a presentation with a set of overhead slides or via an aide-memoire.

A feedback with a set of overheads is more interactive than simply presenting a document, and the presentation can often be reused in a modified format for debriefing at headquarters. The main points about such feedback are that:

- It should build on what the audience already knows or believes.

- It should follow the same structure as the eventual report. This encourages you to think about the issues using the same structure that you will use for the report, and helps to prevent any confusion about what the main issues are.

- It should raise critical issues now, rather than just leaving them for the report (which can result in the country team feeling that they have been 'ambushed' and unfairly dealt with).

The main components of the feedback should be:

- The introduction. This should present the aims and objectives of the evaluation. It should also summarise the audience's perception of the context as expressed by them during the field work. This should provide an initial platform on which the whole audience can agree.

- The methodology used, including the selection of interviewees, cases, and other sources of information. It is particularly important here to make any limitations in the methodology clear (but be ready to explain in what way your findings may be valid even in spite of these limitations). This section should also indicate the approximate numbers of people and documents consulted so that the audience understand the strength (or weakness) of the evidence for the results.

- The results of your research, normally presented by what will be the main chapters of the final report including.

  o Your findings and the evidence for them (in summary format).

  o Your conclusions. In some cases you may also have to present your preliminary recommendations at this state.

- The planned process for completing the evaluation and the evaluation products.

If you are also going to look for input to the report, or on next steps etc. from the feedback meeting you need to clearly mark the end of the feedback session and the start of the input session by changing the mode. You can do this by switching from a presentation to a series of small groups discussions, or a facilitated group discussion with someone else doing the facilitation.

REFERENCES AND RESOURCES

Drever, E. (2003). *Using semi-structured interviews in small-scale research: a teacher's guide* (revised ed.). Glasgow: Scottish Council for Research in Education

RECOMMENDATIONS

TOO MANY RECOMMENDATIONS

The fewer recommendations that are made, the more usable the report is, However there is an ethical question here, in that if the evaluators have noticed problems, they are duty bound to flag them up and to recommend solutions if they can. One way of address this issue is to put minor recommendations in a technical annexe, as in this example from a Danida report on humanitarian

Appendix 8 Detailed Technical Comments

The following comments relate to specific technical points noted by the evaluator during the field visits. Not all of the points raised refer to Danida funded activities. While some are raised in the general text, other are too specific for general comment.

Location: Renovated school at Kisange.

Problem: One of the Roof trusses of the renovated school has cracks in the lower member. This may eventually lead to failure and collapse.

Possible solution: Fixing of timber reinforcement to truss after temporary propping of the truss.

Agencies should only undertake technical work when they have the competence to supervise it.

Location: Various

Problem: The latrines that have been built at the water points are not being used.

Possible solution: Drop the construction of latrines at water points unless there is a community demand for these. The construction of laundry basins and washing areas is good and should be continued.

Location: Renovated school at Kisange.

Problem: The school desks have no diagonal bracing. Within a year of use the desks will sway from side to side, despite the dovetail joint at the front.

Possible solution: Change the design to include a diagonal cross brace dovetailed into the rear of the seats or just under the seat. Alternatively use triangular gussets or a deeper cross member at the back.

assistance in Angola.

Good candidates for the annexe are very technical points, or relatively minor recommendations, or very specific recommendations.

Another approach is to use omnibus recommendations, where there is a general recommendation followed by a list of specific steps that could achieve this. The following shows an omnibus recommendation in the context where a community falsely believed (after project funding cuts) that project staff were diverting funds from the project to fill their own pockets.

> Recommendation: The agency should, within two months, improve communication with the affected population on what its plans are, and on how funds are being used.
>
> *This could be achieved by:*
>
> > *Providing a report on progress and plans to the community every month.*
> >
> > *Putting signboard at all project sites with the details of the intervention, the donor, including a description, the expected completing data, and an outline budget.*
> >
> > *Posting simplified accounts at each project site every month, showing the expenditure that month against the budget on labour, materials, and other costs.*
> >
> > *Setting up a complaints mechanism so that community members can raise questions about the problems that they have with the implementation of specific project elements.*

USABLE RECOMMENDATIONS

Recommendations are really only usable if they reach the right person. The more recommendations that are made, the harder it is for them to reach the right person. Some recommendations can be delivered verbally in the field or headquarters to the concerned staff. They may even not be couched as recommendations, but as observations on how a similar problem was dealt with

> *"They faced the same problem in Malawi of community leaders adding family members to the relief role. They got around it by having three groups leaders, other man, and other women, draw up separate lists and only including those that appeared on all three, with a discussion on those that only appeared on only one or two lists."*

elsewhere.

Of course, the team may wish to include this observation in the report just to demonstrate that they are providing value for money. Or they may include it in an omnibus recommendation about improving targeting coverage.

For recommendations to be useful they should be:

1. Specific - it must be clear exactly what is being recommended.

2. Measurable - it should be possible to tell whether the recommendation has been implemented or not.

3. Accountable - the person or entity responsible for implementing the recommendation should be identified (responsibility may be further clarified in a management response to the report).

4. Realistic - recommendations need to fall within the range of the possible. This does not mean that they cannot be "outside the box" but they should bear resources and other constraints in mind.

5. Time-bound - the timetable for implementing the recommendation should also be given

6. Coherent - Recommendations should be coherent with each other and should neither contradict nor seem to contradict each other.

7. Ordered by priority - so that it is clear which recommendations are of primary concern and which ones are secondary.

8. Limited in number - if a large number of recommendations are made, they can cloud of recommendations may be so demanding that not one gets implemented. See above for some suggestions on limiting the number of recommendations.

9. Economic - the recommended actions should clearly deliver benefits in proportion to their costs.

WRITING THE REPORT

Writing the report is can be quite challenging as there is there is never as much time as you would like to fully digest the information, sort it, analyse it and develop the report. Evaluation reports are typically of 12,000 to 25,000 words plus annexes.

WRITING STRATEGY

Some team leaders prefer to write the whole report, others prefer to farm out chapters between the team. The first arrangement provides a more coherent document but places a large workload on the team leader. It is also necessary to have some mechanism to agree conclusions and recommendations within in the team if the team leader is to do the bulk of the writing.

The team leader cannot wait until the field work is finished to begin writing the report. A good strategy is to being the report in the early stages of the research, starting with the report structure and an introductory chapter describing the context. This should be shared with the team members for comment so the team have a common understanding of the context.

REPORT STRUCTURE

Setting out the report structure early on is essential to allow team members to take organise their notes in line with the structure to allow easier melding of the different information into one final report. The report layout may be specified in the Term of Reference, but commissioning agencies are usually flexible on the structure of the main report provide that the information needs are met.

The chapters in the report are often organised in a matrix of chapter headings and sub headings. One approach is to use the evaluation criteria as chapter titles with sub-chapters for themes. Another is to reverse the order with themes as the chapter titles and have the evaluation criteria as sub-chapters.

Choices for your themes include:

- Categories from the evaluation terms of reference.

- Sectors or clusters. Having separate chapters for each sector.

- Themes from disaster literature. For example: preparedness; disaster; rescue; relief; and recovery.

- Themes that emergency from the research. For example you might use categories used by the affected community themselves to describe what happened as themes.

- Chronology of the response. For example you might have separate chapters on Day 0, Day 1-3, Day 4-10 etc.

- Trajectory of processes in the response. For an evaluation concentrating on logistics support you might look at the different stages in procurement. Your chapters might be assessment, specification, request supply, request

offers, bid analysis, purchase, delivery, receipt, storage, and distribution or consumption. An evaluation looking at a grant making organisation might divide the trajectory of a grant application into its different stages.

- Comparison with an ideal type. For example, create sections for each of the ten principles of the Red Cross/NGO Code of Conduct or the elements of the Sphere standards or some other criteria.

Example of themes from a humanitarian evaluation drawn from a terms of reference

1. *Human resources and management systems scaling up.*
2. *Collaboration between administration and programme offices*
3. *Partnership*
4. *Community capacities and needs*
5. *Gender*
6. *Other groups with special needs*
7. *Programming and delivery.*
8. *Logistics including procurement and delivery mechanisms*

REVIEWING THE REPORT

The evaluation report is still the main output from evaluations. Whoever is responsible for quality control of the report should check (as should the evaluation manager) that the report:

- Broadly meets the terms of reference, and if not, whether the reasons why it does not are clearly explained.

- Meets any specific agency layout formatting requirements.

- Provides evidence to support the conclusions and recommendations in the report.

- Is coherent and free from internal contradictions.

- Answers the evaluation questions.

- Presents clear recommendations and identifies who is being asked to implement them.

- Contains sufficient information about process and methods to enable readers to judge how much reliance can be put on it.

Readers of the evaluation report will make their own assessment of the quality of the evaluation and of the authority of the recommendations. A good report reflects well on the evaluation manager and the evaluation team, a bad report does otherwise.

EVALUATION REPORT CONTENT CHECKLIST

This checklist can be used when writing the evaluation report (or when reviewing it). It might also be used when developing the terms of reference. The following is loosely based on the checklist presented in (ALNAP et al., 2005, p. 192).

	Front matter elements	Notes
☐	Title page with date	
☐	Data page	Optional. But this may include the administrative details of the evaluation, the report number etc, a caption for any cover photograph, the overall evaluation cost, and the contract number. It is good practice to include a suggested citation for the evaluation report on this page.
☐	Executive summary	The executive summary should not be included with the first draft of the report. This is to ensure that comments are based on the report and not just the executive summary. See a later section for suggestions on the executive summary.
☐	Table of contents with page numbers	Any modern word processing software can quickly produced tables of contents as long as styles as used for the headings.
☐	Map or Maps	Be careful with copyright here.
☐	Acknowledgements	This is not a requirement of the ALNAP proforma, but every evaluator should be careful to give thanks to those who have facilitated the field work and other aspects.

	Main text of the report	Notes
☐	Introduction	This should include the purpose and objective of the evaluation, scope, and team composition.

	Main text of the report	*Notes*
❏	Methodology	This may be part of the introduction chapter, or a separate chapter or even an appendix. The choice of location depends on the intended audience. Even, so, the introduction should include a very short summary of the methodology so that the general reader knows how the information was collected. See the methodology chapter checklist for details of what a full descriptions of the methodology contain.
❏	Context chapter	Setting out both the context of the disaster and the context of the response.
❏	Findings	These can be organised along the line of whatever framework is specified in the Terms of Reference or whatever analytical framework has been used. This chapter or chapters should set out the evaluation findings and the basis for the findings.
❏	Conclusions	Some prefer to present separate chapters for conclusion and recommendations. However, presenting the conclusions for each chapter at the end of the chapter helps to ensure that the conclusions are grounded in the findings and evidence present in the chapter.
❏	Recommendations	Again, it is probably best to present the recommendations directly after the conclusions to which they relate, i.e. at the end of the chapter on a particular topic.

	Annexes	*Notes*
❏	Terms of Reference	By convention, this is generally the first annexe to an evaluation report.
❏	Disaster chronology	This is particularly important in humanitarian evaluations because the speed of response if often critical in saving lives or reducing suffering.
❏	List of interviewees	
❏	Team itinerary	

	Annexes	*Notes*
❑	Short team biographies	Together with the interview list and the itinerary, the team biographies allow readers to judge how much weight to give the evaluation findings. It is important here to present any potential biases or conflicts of interest that the evaluators may be subject to, including previous work history with the organisation concerned, and what steps the team has taken to minimise these.
❑	Interview guides, discussion guides, survey forms etc	
❑	Other annexes	
❑	Bibliography	
❑	Management response	Some organisations may place the management response at the beginning of the report.

REFERENCES AND RESOURCES

ALNAP, Behrman, N., Mitchell, J., Christoplos, I., Minear, L., & Wiles, P. (2005). *ALNAP Review of Humanitarian Action in 2004: Capacity Building*. London: Active Learning Network on Accountability and Performance in Humanitarian Action. Last viewed on 8 June, 2008. URL: http://www.alnap.org/publications/rha.htm

WHAT SHOULD BE IN THE METHODOLOGY CHAPTER?

The chapter on methodology in your report should leave the reader clear as to what methods and data sources were used by the evaluators, how they selected their data sources and what measures they took to ensure the reliability and accuracy of the data they collected. In particular the methodology chapter should address the points laid out in the following checklist.

	Item
❑	A clear statement of what data sources and methods were used.
❑	Description of data collection methods and analysis including:
❑	Any question guides used.
❑	Levels of precision for any quantitative analysis.
❑	Value scales or coding for any quantitative analysis.
❑	Description of sampling including:
❑	Criteria for selection of interviewees.
❑	Numbers of persons interviewed.
❑	Itinerary of sites visited.
❑	List of persons interviewed.
❑	Any limitations on the sampling.
❑	Any standards or benchmarks (e.g. Sphere) used by the evaluators.
❑	Any deviations from the evaluation plan.
❑	A description of what types of triangulation were used.
❑	A statement of the ethical standard adopted by the evaluation team (i.e. use of the Chatham House Rule, treatment of notes etc).

❑   Any particular limitations that the research faced.

Note: Loosely based on (World Bank, 2007, p.107)

An additional resource here is the ALNAP Quality Proforma (ALNAP et al., 2005, pp. 181-192).

REFERENCES AND RESOURCES

ALNAP, Behrman, N., Mitchell, J., Christoplos, I., Minear, L., & Wiles, P. (2005). *ALNAP Review of Humanitarian Action in 2004: Capacity Building*. London: Active Learning Network on Accountability and Performance in Humanitarian Action. Last viewed on 8 June, 2008. URL: http://www.alnap.org/publications/rha.htm

World Bank. (2007). *Sourcebook for Evaluating Global and Regional Partnership Programs: Indicative Principles and Standards*. Washington: The Independent Evaluation Group of the World Bank. Last viewed on 28 June 2008. URL: http://siteresources.worldbank.org/EXTGLOREGPARPRO/Resources/sourcebook.pdf

EXECUTIVE SUMMARY CHECKLIST

The executive summary is the last part of the report to be written. It should not be circulated prior to the penultimate draft, so that comments on these drafts are based on the full text. The shorter the executive summary the more people are likely to read it. The following is a checklist for executive summaries - loosely based on (World Bank, 2007, p. 106).

Item

❑ The name of the commissioning agency.

❑ The names of the evaluation if they are not credited as the authors on the title page.

❑ A brief statement of the context.

❑ A brief description of the response being evaluated, including financial parameters and main activities

❑ The purpose of the evaluation including:

❑ The intended audience of the evaluation report.

❑ The expected use of the evaluation report.

❑ Why it was decided to do an evaluation now.

❑ The objectives of the evaluation and key evaluation questions.

❑ A short description of the methodology including:

❑ The reason for the choice of data sources and methods and which sources were used.

❑ Any major limitations implicit in the methodology

❑ The most important findings and conclusions

❑ Main recommendations and lessons learned

REFERENCES AND RESOURCES

World Bank. (2007). *Sourcebook for Evaluating Global and Regional Partnership Programs: Indicative Principles and Standards*. Washington: The

Independent Evaluation Group of the World Bank. Last viewed on 28 June 2008. URL: http://siteresources.worldbank.org/EXTGLOREGPARPRO/Resources/sourcebook.pdf

Section seven:
Dissemination and utilisation

Ensuring that an evaluation is used, and influences behaviour and performance in the future is not guaranteed and has been a particular challenge to EHA, threatening to undermine the evaluation endeavour.

This section explores why it has been a challenge and potential pitfalls, and also presents some frameworks to encourage a utilisation-focus.

PLANNING FOR DISSEMINATION AND USE

Loosely based on (The World Bank Group et al., 2002, pp. 9-2)

Dissemination is essential to achieve utilisation. The dissemination plan needs to be built in from the start. Dissemination is much wider than just the evaluation

Audience	Product	Who to do	By when
Country team	Mid evaluation briefing	Whole team	Agreed date
Programme staff	Final briefing	Whole team	Agreed date
Agency	First draft report	Team leader	At end of fieldwork
UN HC	Findings and Conclusions briefing	Team leader	At end of fieldwork
Ministry officials	Oral briefing	Team member C	On completion of final report
NGO Forum	Oral briefing	Team member C	On completion of final report
Regional HQ	Briefing on findings, conclusions, and recommendations	Team member B	On exit from field
Global HQ	Briefing on findings, conclusions, and recommendations	Team leader	Within three days of departure from field
Agency	Final report	Team leader	With seven days of receipt of comments
Development education community	Journal article on the evaluation	Whole team	2 months after fieldwork
Humanitarian evaluation community	Presentation at ALNAP Biannual	Team Leader	Next Biannual

report. There are also staff briefings, briefing of other NGOs, Government etc.

The issue is to ensure that dissemination is thought about in advance and that any necessary days for briefings at headquarters and any other dissemination costs including translation are through about in advance.

Dissemination does not guarantee utilisation, but utilisation is impossible without dissemination. It is good to agree a dissemination plan in advance, as there are

those who will be opposed to the broader dissemination of any evaluation that contains even a single word that might be interpreted as a criticism of the agency.

REFERENCES AND RESOURCES

The World Bank Group, Carleton University, & IOB/Netherlands Ministry of Foreign Affairs. (2002). *International Program for Development Evaluation Training (IPDET): Building Skills to Evaluate Development Interventions: Module 9. Presenting Results*. Washington: Operations Evaluation Department of the World Bank and Operations Evaluation Group of the International Finance Corporation. Last viewed on 28 June 2008. URL: http://www.insp.mx/bidimasp/documentos/1/IPDET%20modulo%209.pdf

EVALUATION USE AND DISSEMINATION:

What do we mean by utilisation? Sandison (2006) describes it as follows:

'An evaluation has been utilised if users with the intention and potential to act have given serious, active consideration to its findings, identifying meaningful uses according to their own interests and needs'

Quinn Patton (1997) identifies three primary uses of evaluation findings:

1) Judging the merit or worth of a programme (eg accountability to stakeholders, to inform funding decisions)

2) Improving a programme (eg ongoing learning & development)

3) Generating knowledge

The first two uses are often referred to as 'instrumental', and are the most common expected uses of an evaluation (Sandison, 2006). The third use, generating knowledge, is sometimes referred to as the 'conceptual use', and may be particularly associated with synthesis reports and with some joint evaluations. Where the potential evaluation users are also involved in the evaluation process, this can contribute to individual learning, but also to strengthening an organisation's learning culture (ibid.).

The record of utilisation of humanitarian evaluations does not appear to be a very encouraging one. Consider the following:'although evaluations are generally successful in identifying lessons and building institutional memory, only a minority of evaluations are effective at introducing evident changes or improvements in performance. If this continues, there is a danger that continued poor utilisation will undermine the credibility of evaluation as a tool for accountability and learning in the humanitarian sector' (from Sandison, P., 'The Utilisation of Evaluations' in ALNAP Review of Humanitarian Action, 2006, London: ODI, p 90)

But it is important also to capture the indirect uses and influences of evaluations as well as the anticipated direct uses. A number of factors influence the extent to which an evaluation is used, and how, to do with the quality of the evaluation itself eg the design of the evaluation, the participation of key stakeholders; with relational issues eg the quality of relationship between the evaluators and key users; and with organisational issues eg whether or not there is a strong learning culture within the organisation. This enters the territory of organisational change and organisational learning.

Four key points emerge, about how to ensure greater use of evaluations:

1) Be clear about the purpose of the evaluation and its users at the outset

2) Design a dissemination strategy at the outset

3) Ensure that the key stakeholders (ie intended users of the evaluation) are actively involved in the evaluation throughout

4) Design, and fund, a follow-up process

FUNDAMENTAL PREMISES OF THE UTILIZATION-FOCUS

Michael Quinn Patton (2008, p570-573) presents fifteen fundamental premises for utilisation focused evaluation:

1. Commitment to intended use by intended users should be the driving force in an evaluation

2. Strategizing about use is ongoing and continuous from the very beginning of an evaluation

3. The personal factor contributes significantly to use

4. Careful and thoughtful stakeholder analysis should inform identification of primary intended users

5. Evaluations must be focussed in some way; focusing on intended use by intended users is the most useful way

6. Focusing on intended use requires making deliberate and thoughtful choices

7. Useful evaluations must be designed and adapted situationally

8. Intended users' commitment to use can be nurtured and enhanced by actively involving them in making significant decisions about the evaluation

9. High-quality participation is the goal, not high-quantity participation

10. High-quality involvement of intended users will result in high-quality evaluations

11. Evaluators have a rightful stake in that their credibility and integrity are always at risk, thus the mandate to be active-reactive-interactive-adaptive

12. Evaluators committed to enhancing use have both and opportunity and a responsibility to train users

13. Use is different from reporting and dissemination

14. Serious attention to use involves financial and time costs that are far from trivial

15. Commitment to improving practice means following up evaluations to find out how they have been used.

REFERENCES AND RESOURCES:

Patton, M. Q. (2002). *Utilization-Focused Evaluation Checklist.* Kalamazoo: The Evaluation Center, Western Michigan University. Last viewed on 28 June 2008. URL: http://www.wmich.edu/evalctr/checklists/ufe.pdf

Patton, M. Q. (2008). *Utilization-focused evaluation (4th ed.).* Thousand Oaks: Sage

THE MANAGEMENT RESPONSE

The management response should respond to the conclusions and recommendations in the report stating:

- If the conclusions are broadly accepted or not. If a conclusion is not accepted the evidence against the conclusion should be stated. This may be given as a general commentary on the report specifically identifying which conclusions are not accepted.

- For each recommendation

    o Whether the recommendation is accepted. If recommendations are not accepted the reason for not accepting them should be clearly stated.

    o What action has already been taken and by whom, to address the recommendation, or the underlying problem.

    o What action is planned to address the recommendation, or the underlying problem. Who is going to take this action, and what the timetable is.

- What mechanism the agency is going to use to monitor the planned actions.

Extract from a management response to an FAO evaluation (FAO, 2008).

Recommendations	Further donor funding required	Acceptance by Management			Comment on the Recommendation	Action to be taken		
		Accept	Partially Accept	Reject		Action	Timing	Unit responsible
Recommendation 8: Specific consideration should be given in the national priority framework to cross-cutting issues such as environment, gender equity and HIV/AIDS. Together with MAFS, the National Farmers Association, and civil society agriculture-led HIV/AIDS-mitigation strategies at community level should be developed in the context of expansion of FFS in Sierra Leone	Yes	X			The recommendation will be incorporated in the NMTPF and NADP planning which should consider activities to improve rural living conditions such as better nutrition, HIV/AIDS, malaria, schistosomiasis prevention literacy and other essential livelihood areas.	Liaise with UN agencies, line ministries of the GoSL and NGOs working in these areas.	Immediate	FAOR, SFW (with other Units), ESW

Getting a management response is more difficult in agencies which do not already have a system for this. However, this is something for the evaluation manager to push to ensure that the agency gets value for money from evaluation. If there is no automatic system for generating a management response, it is worth considering which stakeholders might be co-opted to support the need for one.

FAO. (2008). *Management Response to the Evaluation of FAO Cooperation in Sierra Leone 2001-2006*. Rome: Food and Agricultural Organisation of the

United Nations. Last viewed on 21 September 2008. URL:
http://www.fao.org/pbe/pbee/common/ecg/351/en/ManagementResponseSi
erraLeone.doc

Section eight:
Resources and tools

This section presents a series of resources that may be of use to evaluators. There is a companion annotated bibliography of resources for humanitarian evaluators that may also be of use.

The section begins with two statements of good practice for evaluators, the AEA's Guiding Principles, and the Program Evaluation Standards. This is followed by as set of practical tools for administrative tasks in an evaluation, such as keeping a record of persons met etc.

There is section on resources on the web that may be of interest. The section ends with the ALNAP quality proforma.

GUIDING PRINCIPLES FOR EVALUATORS

These are the guiding principles for evaluators developed by the Ethics Committee of the American Evaluation Association (AEA). This document can be found on-line at http://www.eval.org/Publications/GuidingPrinciplesPrintable.asp.

The version below includes the revision ratified by the AEA membership in July 2004.

PREFACE: ASSUMPTIONS CONCERNING DEVELOPMENT OF PRINCIPLES

A. Evaluation is a profession composed of persons with varying interests, potentially encompassing but not limited to the evaluation of programs, products, personnel, policy, performance, proposals, technology, research, theory, and even of evaluation itself. These principles are broadly intended to cover all kinds of evaluation. For external evaluations of public programs, they nearly always apply. However, it is impossible to write guiding principles that neatly fit every context in which evaluators work, and some evaluators will work in contexts in which following a guideline cannot be done for good reason. The Guiding Principles are not intended to constrain such evaluators when this is the case. However, such exceptions should be made for good reason (e.g., legal prohibitions against releasing information to stakeholders), and evaluators who find themselves in such contexts should consult colleagues about how to proceed.

B. Based on differences in training, experience, and work settings, the profession of evaluation encompasses diverse perceptions about the primary purpose of evaluation. These include but are not limited to the following: bettering products, personnel, programs, organizations, governments, consumers and the public interest; contributing to informed decision making and more enlightened change; precipitating needed change; empowering all stakeholders by collecting data from them and engaging them in the evaluation process; and experiencing the excitement of new insights. Despite that diversity, the common ground is that evaluators aspire to construct and provide the best possible information that might bear on the value of whatever is being evaluated. The principles are intended to foster that primary aim.

C. The principles are intended to guide the professional practice of evaluators, and to inform evaluation clients and the general public about the principles they can expect to be upheld by professional evaluators. Of course, no statement of principles can anticipate all situations that arise in the practice of evaluation. However, principles are not just guidelines for reaction when something goes wrong or when a dilemma is found. Rather, principles should proactively guide the behaviors of professionals in everyday practice.

D. The purpose of documenting guiding principles is to foster continuing development of the profession of evaluation, and the socialization of its members. The principles are meant to stimulate discussion about the proper practice and use of evaluation among members of the profession, sponsors of evaluation, and others interested in evaluation.

E. The five principles proposed in this document are not independent, but overlap in many ways. Conversely, sometimes these principles will conflict, so that evaluators will have to choose among them. At such times evaluators must use their own values and knowledge of the setting to determine the appropriate response. Whenever a course of action is unclear, evaluators should solicit the advice of fellow evaluators about how to resolve the problem before deciding how to proceed.

F. These principles are intended to supercede any previous work on standards, principles, or ethics adopted by AEA or its two predecessor organizations, the Evaluation Research Society and the Evaluation Network. These principles are the official position of AEA on these matters.

G. These principles are not intended to replace standards supported by evaluators or by the other disciplines in which evaluators participate.

H. Each principle is illustrated by a number of statements to amplify the meaning of the overarching principle, and to provide guidance for its application. These illustrations are not meant to include all possible applications of that principle, nor to be viewed as rules that provide the basis for sanctioning violators.

I. These principles were developed in the context of Western cultures, particularly the United States, and so may reflect the experiences of that context. The relevance of these principles may vary across other cultures, and across subcultures within the United States.

J. These principles are part of an evolving process of self-examination by the profession, and should be revisited on a regular basis. Mechanisms might include officially-sponsored reviews of principles at annual meetings, and other forums for harvesting experience with the principles and their application. On a regular basis, but at least every five years, these principles ought to be examined for possible review and revision. In order to maintain association-wide awareness and relevance, all AEA members are encouraged to participate in this process.

THE PRINCIPLES

A. SYSTEMATIC INQUIRY

Evaluators conduct systematic, data-based inquiries.

1. To ensure the accuracy and credibility of the evaluative information they produce, evaluators should adhere to the highest technical standards appropriate to the methods they use.

2. Evaluators should explore with the client the shortcomings and strengths both of the various evaluation questions and the various approaches that might be used for answering those questions.

3. Evaluators should communicate their methods and approaches accurately and in sufficient detail to allow others to understand, interpret and critique their work. They should make clear the limitations of an evaluation and its results. Evaluators should discuss in a contextually appropriate way those values, assumptions, theories, methods, results, and analyses significantly affecting the

interpretation of the evaluative findings. These statements apply to all aspects of the evaluation, from its initial conceptualization to the eventual use of findings.

B. COMPETENCE

Evaluators provide competent performance to stakeholders.

1. Evaluators should possess (or ensure that the evaluation team possesses) the education, abilities, skills and experience appropriate to undertake the tasks proposed in the evaluation.

2. To ensure recognition, accurate interpretation and respect for diversity, evaluators should ensure that the members of the evaluation team collectively demonstrate cultural competence. Cultural competence would be reflected in evaluators seeking awareness of their own culturally-based assumptions, their understanding of the worldviews of culturally-different participants and stakeholders in the evaluation, and the use of appropriate evaluation strategies and skills in working with culturally different groups. Diversity may be in terms of race, ethnicity, gender, religion, socio-economics, or other factors pertinent to the evaluation context.

3. Evaluators should practice within the limits of their professional training and competence, and should decline to conduct evaluations that fall substantially outside those limits. When declining the commission or request is not feasible or appropriate, evaluators should make clear any significant limitations on the evaluation that might result. Evaluators should make every effort to gain the competence directly or through the assistance of others who possess the required expertise.

4. Evaluators should continually seek to maintain and improve their competencies, in order to provide the highest level of performance in their evaluations. This continuing professional development might include formal coursework and workshops, self-study, evaluations of one's own practice, and working with other evaluators to learn from their skills and expertise.

C. INTEGRITY/HONESTY

Evaluators display honesty and integrity in their own behavior, and attempt to ensure the honesty and integrity of the entire evaluation process.

1. Evaluators should negotiate honestly with clients and relevant stakeholders concerning the costs, tasks to be undertaken, limitations of methodology, scope of results likely to be obtained, and uses of data resulting from a specific evaluation. It is primarily the evaluator's responsibility to initiate discussion and clarification of these matters, not the client's.

2. Before accepting an evaluation assignment, evaluators should disclose any roles or relationships they have that might pose a conflict of interest (or appearance of a conflict) with their role as an evaluator. If they proceed with the evaluation, the conflict(s) should be clearly articulated in reports of the evaluation results.

3. Evaluators should record all changes made in the originally negotiated project plans, and the reasons why the changes were made. If those changes would significantly affect the scope and likely results of the evaluation, the evaluator should inform the client and other important stakeholders in a timely fashion (barring good reason to the contrary, before proceeding with further work) of the changes and their likely impact.

4. Evaluators should be explicit about their own, their clients', and other stakeholders' interests and values concerning the conduct and outcomes of an evaluation.

5. Evaluators should not misrepresent their procedures, data or findings. Within reasonable limits, they should attempt to prevent or correct misuse of their work by others.

6. If evaluators determine that certain procedures or activities are likely to produce misleading evaluative information or conclusions, they have the responsibility to communicate their concerns and the reasons for them. If discussions with the client do not resolve these concerns, the evaluator should decline to conduct the evaluation. If declining the assignment is unfeasible or inappropriate, the evaluator should consult colleagues or relevant stakeholders about other proper ways to proceed. (Options might include discussions at a higher level, a dissenting cover letter or appendix, or refusal to sign the final document.)

7. Evaluators should disclose all sources of financial support for an evaluation, and the source of the request for the evaluation.

D. RESPECT FOR PEOPLE

Evaluators respect the security, dignity and self-worth of respondents, program participants, clients, and other evaluation stakeholders.

1. Evaluators should seek a comprehensive understanding of the important contextual elements of the evaluation. Contextual factors that may influence the results of a study include geographic location, timing, political and social climate, economic conditions, and other relevant activities in progress at the same time.

2. Evaluators should abide by current professional ethics, standards, and regulations regarding risks, harms, and burdens that might befall those participating in the evaluation; regarding informed consent for participation in evaluation; and regarding informing participants and clients about the scope and limits of confidentiality.

3. Because justified negative or critical conclusions from an evaluation must be explicitly stated, evaluations sometimes produce results that harm client or stakeholder interests. Under this circumstance, evaluators should seek to maximize the benefits and reduce any unnecessary harms that might occur, provided this will not compromise the integrity of the evaluation findings. Evaluators should carefully judge when the benefits from doing the evaluation or in performing certain evaluation procedures should be foregone because of the risks or harms. To the extent possible, these issues should be anticipated during the negotiation of the evaluation.

4. Knowing that evaluations may negatively affect the interests of some stakeholders, evaluators should conduct the evaluation and communicate its results in a way that clearly respects the stakeholders' dignity and self-worth.

5. Where feasible, evaluators should attempt to foster social equity in evaluation, so that those who give to the evaluation may benefit in return. For example, evaluators should seek to ensure that those who bear the burdens of contributing data and incurring any risks do so willingly, and that they have full knowledge of and opportunity to obtain any benefits of the evaluation. Program participants should be informed that their eligibility to receive services does not hinge on their participation in the evaluation.

6. Evaluators have the responsibility to understand and respect differences among participants, such as differences in their culture, religion, gender, disability, age, sexual orientation and ethnicity, and to account for potential implications of these differences when planning, conducting, analyzing, and reporting evaluations.

E. RESPONSIBILITIES FOR GENERAL AND PUBLIC WELFARE

Evaluators articulate and take into account the diversity of general and public interests and values that may be related to the evaluation.

1. When planning and reporting evaluations, evaluators should include relevant perspectives and interests of the full range of stakeholders.

2. Evaluators should consider not only the immediate operations and outcomes of whatever is being evaluated, but also its broad assumptions, implications and potential side effects.

3. Freedom of information is essential in a democracy. Evaluators should allow all relevant stakeholders access to evaluative information in forms that respect people and honor promises of confidentiality. Evaluators should actively disseminate information to stakeholders as resources allow. Communications that are tailored to a given stakeholder should include all results that may bear on interests of that stakeholder and refer to any other tailored communications to other stakeholders. In all cases, evaluators should strive to present results clearly and simply so that clients and other stakeholders can easily understand the evaluation process and results.

4. Evaluators should maintain a balance between client needs and other needs. Evaluators necessarily have a special relationship with the client who funds or requests the evaluation. By virtue of that relationship, evaluators must strive to meet legitimate client needs whenever it is feasible and appropriate to do so. However, that relationship can also place evaluators in difficult dilemmas when client interests conflict with other interests, or when client interests conflict with the obligation of evaluators for systematic inquiry, competence, integrity, and respect for people. In these cases, evaluators should explicitly identify and discuss the conflicts with the client and relevant stakeholders, resolve them when possible, determine whether continued work on the evaluation is advisable if the conflicts cannot be resolved, and make clear any significant limitations on the evaluation that might result if the conflict is not resolved.

5. Evaluators have obligations that encompass the public interest and good. These obligations are especially important when evaluators are supported by publicly-generated funds; but clear threats to the public good should never be ignored in any evaluation. Because the public interest and good are rarely the same as the interests of any particular group (including those of the client or funder), evaluators will usually have to go beyond analysis of particular stakeholder interests and consider the welfare of society as a whole.

BACKGROUND

In 1986, the Evaluation Network (ENet) and the Evaluation Research Society (ERS) merged to create the American Evaluation Association. ERS had previously adopted a set of standards for program evaluation (published in New Directions for Program Evaluation in 1982); and both organizations had lent support to work of other organizations about evaluation guidelines. However, none of these standards or guidelines were officially adopted by AEA, nor were any other ethics, standards, or guiding principles put into place. Over the ensuing years, the need for such guiding principles was discussed by both the AEA Board and the AEA membership. Under the presidency of David Cordray in 1992, the AEA Board appointed a temporary committee chaired by Peter Rossi to examine whether AEA should address this matter in more detail. That committee issued a report to the AEA Board on November 4, 1992, recommending that AEA should pursue this matter further. The Board followed that recommendation, and on that date created a Task Force to develop a draft of guiding principles for evaluators.   The task force members were:

- William Shadish, Memphis State University (Chair)
- Dianna Newman, University of Albany/SUNY
- Mary Ann Scheirer, Private Practice
- Chris Wye, National Academy of Public Administration

The AEA Board specifically instructed the Task Force to develop general guiding principles rather than specific standards of practice. Their report, issued in 1994, summarized the Task Force's response to the charge.

PROCESS OF DEVELOPMENT

Task Force members reviewed relevant documents from other professional societies, and then independently prepared and circulated drafts of material for use in this report. Initial and subsequent drafts (compiled by the Task Force chair) were discussed during conference calls, with revisions occurring after each call. Progress reports were presented at every AEA board meeting during 1993. In addition, a draft of the guidelines was mailed to all AEA members in September 1993 requesting feedback; and three symposia at the 1993 AEA annual conference were used to discuss and obtain further feedback. The Task Force considered all this feedback in a December 1993 conference call, and prepared a final draft in January 1994. This draft was presented and approved for membership vote at the January 1994 AEA board meeting.

RESULTING PRINCIPLES

Given the diversity of interests and employment settings represented on the Task Force, it is noteworthy that Task Force members reached substantial agreement about the following five principles. The order of these principles does not imply priority among them; priority will vary by situation and evaluator role.

A. Systematic Inquiry: Evaluators conduct systematic, data-based inquiries about whatever is being evaluated.

B. Competence: Evaluators provide competent performance to stakeholders.

C. Integrity/Honesty: Evaluators ensure the honesty and integrity of the entire evaluation process.

D. Respect for People: Evaluators respect the security, dignity and self-worth of the respondents, program participants, clients, and other stakeholders with whom they interact.

E. Responsibilities for General and Public Welfare: Evaluators articulate and take into account the diversity of interests and values that may be related to the general and public welfare.

RECOMMENDATION FOR CONTINUED WORK

The Task Force also recommended that the AEA Board establish and support a mechanism for the continued development and dissemination of the Guiding Principles, to include formal reviews at least every five years. The Principles were reviewed in 1999 through an EvalTalk survey, a panel review, and a comparison to the ethical principles of the Canadian and Australasian Evaluation Societies. The 2000 Board affirmed this work and expanded dissemination of the Principles; however, the document was left unchanged.

PROCESS OF THE 2002-2003 REVIEW AND REVISION

In January 2002 the AEA Board charged its standing Ethics Committee with developing and implementing a process for reviewing the Guiding Principles that would give AEA's full membership multiple opportunities for comment. At its Spring 2002 meeting, the AEA Board approved the process, carried out during the ensuing months. It consisted of an online survey of the membership that drew 413 responses, a "Town Meeting" attended by approximately 40 members at the Evaluation 2002 Conference, and a compilation of stories about evaluators' experiences relative to ethical concerns told by AEA members and drawn from the *American Journal of Evaluation*. Detailed findings of all three sources of input were reported to the AEA Board in *A Review of AEA's Guiding Principles for Evaluators*, submitted January 18, 2003.

In 2003 the Ethics Committee continued to welcome input and specifically solicited it from AEA's Diversity Committee, Building Diversity Initiative, and Multi-Ethnic Issues Topical Interest Group. The first revision reflected the Committee's consensus response to the sum of member input throughout 2002 and 2003. It was submitted to AEA's past presidents, current board members, and the original framers of the Guiding Principles for comment. Twelve reviews were received and incorporated into a second revision, presented at the 2003 annual

conference. Consensus opinions of approximately 25 members attending a Town Meeting are reflected in this, the third and final revision that was approved by the Board in February 2004 for submission to the membership for ratification. The revisions were ratified by the membership in July of 2004.

The 2002 Ethics Committee members were:

- Doris Redfield, Appalachia Educational Laboratory (Chair)
- Deborah Bonnet, Lumina Foundation for Education
- Katherine Ryan, University of Illinois at Urbana-Champaign
- Anna Madison, University of Massachusetts, Boston

In 2003 the membership was expanded for the duration of the revision process:

- Deborah Bonnet, Lumina Foundation for Education (Chair)
- Doris Redfield, Appalachia Educational Laboratory
- Katherine Ryan, University of Illinois at Urbana-Champaign
- Gail Barrington, Barrington Research Group, Inc.
- Elmima Johnson, National Science Foundation

THE PROGRAM EVALUATION STANDARDS

The Joint Committee on Standards for Educational Evaluation was formed in 1975 and published a set of standards in 1981 with a revised 2nd edition in 1994 (Joint Committee on Standards for Educational Evaluation and Sanders, 1994). The summary of the standards can be found online at http://www.wmich.edu/evalctr/jc/. The OECD/DAC is currently developing a set of evaluation quality standards. These are in draft form at present (OECD/DAC NDE, 2006).

Although specifically developed for the evaluation of educational programmes, the standards can be applied to evaluations in other sectors as was done for Sida evaluations in the late 90s (Carlsson et al., 1997).

The thirty program evaluation standards are grouped into four areas:

- Utility (7 standards)
- Feasibility (3 standards)
- Propriety (8 standards)
- Accuracy (12 standards)

UTILITY STANDARDS

The utility standards are intended to ensure that an evaluation will serve the information needs of intended users.

U1 Stakeholder Identification: Persons involved in or affected by the evaluation should be identified, so that their needs can be addressed.

U3 Information Scope and Selection: Information collected should be broadly selected to address pertinent questions about the program and be responsive to the needs and interests of clients and other specified stakeholders

U4 Values Identification: The perspectives, procedures, and rationale used to interpret the findings should be carefully described, so that the bases for value judgments are clear.

U5 Report Clarity: Evaluation reports should clearly describe the program being evaluated, including its context, and the purposes, procedures, and findings of the evaluation, so that essential information is provided and easily understood.

U6 Report Timeliness and Dissemination: Significant interim findings and evaluation reports should be disseminated to intended users, so that they can be used in a timely fashion.

U7 Evaluation Impact: Evaluations should be planned, conducted, and reported in ways that encourage follow-through by stakeholders, so that the likelihood that the evaluation will be used is increased.

FEASIBILITY STANDARDS

The feasibility standards are intended to ensure that an evaluation will be realistic, prudent, diplomatic, and frugal.

F1 Practical Procedures: The evaluation procedures should be practical, to keep disruption to a minimum while needed information is obtained.

F2 Political Viability: The evaluation should be planned and conducted with anticipation of the different positions of various interest groups, so that their cooperation may be obtained, and so that possible attempts by any of these groups to curtail evaluation operations or to bias or misapply the results can be averted or counteracted.

F3 Cost Effectiveness: The evaluation should be efficient and produce information of sufficient value, so that the resources expended can be justified

PROPRIETY STANDARDS

The propriety standards are intended to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results.

P1 Service Orientation: Evaluations should be designed to assist organizations to address and effectively serve the needs of the full range of targeted participants.

P2 Formal Agreements: Obligations of the formal parties to an evaluation (what is to be done, how, by whom, when) should be agreed to in writing, so that these parties are obligated to adhere to all conditions of the agreement or formally to renegotiate it.

P3 Rights of Human Subjects: Evaluations should be designed and conducted to respect and protect the rights and welfare of human subjects.

P4 Human Interactions: Evaluators should respect human dignity and worth in their interactions with other persons associated with an evaluation, so that participants are not threatened or harmed.

P5 Complete and Fair Assessment: The evaluation should be complete and fair in its examination and recording of strengths and weaknesses of the program being evaluated, so that strengths can be built upon and problem areas addressed.

P6 Disclosure of Findings: The formal parties to an evaluation should ensure that the full set of evaluation findings along with pertinent limitations are made accessible to the persons affected by the evaluation and any others with expressed legal rights to receive the results.

P7 Conflict of Interest: Conflict of interest should be dealt with openly and honestly, so that it does not compromise the evaluation processes and results.

P8 Fiscal Responsibility: The evaluator's allocation and expenditure of resources should reflect sound accountability procedures and otherwise be prudent and ethically responsible, so that expenditures are accounted for and appropriate.

ACCURACY STANDARDS

The accuracy standards are intended to ensure that an evaluation will reveal and convey technically adequate information about the features that determine worth or merit of the program being evaluated.

A1 Program Documentation: The program being evaluated should be described and documented clearly and accurately, so that the program is clearly identified.

A2 Context Analysis: The context in which the program exists should be examined in enough detail, so that its likely influences on the program can be identified.

A3 Described Purposes and Procedures: The purposes and procedures of the evaluation should be monitored and described in enough detail, so that they can be identified and assessed.

A4 Defensible Information Sources: The sources of information used in a program evaluation should be described in enough detail, so that the adequacy of the information can be assessed.

A5 Valid Information: The information-gathering procedures should be chosen or developed and then implemented so that they will assure that the interpretation arrived at is valid for the intended use.

A6 Reliable Information: The information-gathering procedures should be chosen or developed and then implemented so that they will assure that the information obtained is sufficiently reliable for the intended use.

A7 Systematic Information: The information collected, processed, and reported in an evaluation should be systematically reviewed, and any errors found should be corrected.

A8 Analysis of Quantitative Information: Quantitative information in an evaluation should be appropriately and systematically analyzed so that evaluation questions are effectively answered.

A9 Analysis of Qualitative Information: Qualitative information in an evaluation should be appropriately and systematically analyzed so that evaluation questions are effectively answered.

A10 Justified Conclusions: The conclusions reached in an evaluation should be explicitly justified, so that stakeholders can assess them.

A11 Impartial Reporting: Reporting procedures should guard against distortion caused by personal feelings and biases of any party to the evaluation, so that evaluation reports fairly reflect the evaluation findings.

A12 Metaevaluation: The evaluation itself should be formatively and summatively evaluated against these and other pertinent standards, so that its conduct is appropriately guided and, on completion, stakeholders can closely examine its strengths and weaknesses.

REFERENCES AND SOURCES

Carlsson, J., Forss, K., Metell, K., Segnestam, L., & Strömberg, T. (1997). *Using the evaluation tool: A survey of conventional wisdom and common practice at Sida* (Sida Studies in Evaluation 97/01). Stockholm: Swedish International Development Agency. Last viewed on 09 July 2009. URL: http://www2.sida.se/shared/jsp/download.jsp?f=STUD97-1.pdf&a=2366

Joint Committee on Standards for Educational Evaluation, & Sanders, J. (1994). *The program evaluation standards* (2nd ed.). Thousand Oaks: Sage Publications.

OECD/DAC NDE. (2006). *DAC Evaluation Quality Standards (for test phase application)*. Paris: Network on Development Evaluation, Development Assistance Committee of the Organisation for Economic Cooperation and Development. URL: http://www.oecd.org/dataoecd/30/62/36596604.pdf

TOOL FOR RECORDING DETAILS OF PERSONS MET

It is important to establish how much reliance readers can place on the evaluation by demonstrating the breadth of the underlying research.

You need to record details of the persons met during your fieldworks. If you only do this afterwards you will have great difficulty afterwards remembering whom you talked to.

This tool consists of a spreadsheet with the following column headings:

Column heading	Example	Notes
Surname, and name	Smith, James	Putting surnames first allows sorting by surname for your final list. This makes it easier for people to find their details and check them.
Agency and function	MSF, Water Engineer	Again, putting the agency name first allows you to sort the names by agency if this is what you want to do.
Gender	♂	You should always present an analysis of the gender of interviewees. You can either use the Venus/Mars symbols or simply code f/m.
Interview method	ssi - Semi structured individual interviewee	You can use a list of codes for this and restrict entries to these codes. This helps to provide readers with an idea of what methods you used.
Category of interviewee	N (code for NGO)	You can use a list of codes for this and restrict entries to these codes.
Place Interviewed	Lilongwe	
Country	MW (ISO country code for Malawi)	You can use a list of codes for this and restrict entries to these codes. Country may not be appropriate for your evaluation.

Date interviewed	Mon 2 Mar 09	Format your dates to give the day as well as the date (custom format *"ddd d mmm yy"* in Excel). You only need to enter the date in your normal manner ("2/3" in this case), and your spreadsheet does the rest.
Interviewer	tl (initials of the team member present during the interview)	You may keep this information for team use rather than publishing it in the report.
Most detailed	Yes - indicating that this was the most detailed of all interviews with this person	This is useful to present double counting of persons who are interviewed more than once.
Telephone - direct	Phone number for any follow-up questions or clarification	This will not be included in your report, but is useful for managing your contacts during the evaluation
Telephone - mobile	Phone number for any follow-up questions or clarification	
Email	Email address for sending a copy of the draft report or for any follow-up questions or clarification	

The advantage of entering this data in a spreadsheet during the evaluation is that the information can then be used to check the balance of key informants (for example, if you have interviewed relatively few government or donor representatives). It also allows you to quickly generate summary tables for the evaluation report

Example of an interview summary table:

Function:	T	No	% of total	Of which ♀	
Agency (a)	a	69	30%	23	33%
Beneficiary (b)	b	113	48%	61	54%
Donor (d)	d	17	7%	7	41%
Government (g)	c	34	15%	6	18%
Total		233		97	42%

Key informant interviews are the main method in humanitarian evaluations. Including a full list of those interviewed, along with their agency affiliation and role allows readers to judge how well grounded your findings and conclusions are.

In some cases the ethical need to protect interviewees may demand that you do not give their names in the final report. In such cases you can either give these interviewees names such as 'Villager 1' etc, or avoid giving a detailed list, but just present the summary table.

TOOL FOR RECORDING DETAILS OF GROUP MEETINGS

Many evaluation interviews may take the form of meetings with large groups of people. It is not appropriate in that case to take everyone's name (as relatively few people have an opportunity to speak) but to record the details of the meeting.

Date	Attendees	Number of ♂	Number of ♀	Location	Chair	Duration	Topics	Team

Providing the details of the general meetings that you have had with beneficiaries and others allows readers to judge how much reliance can be placed on your findings and conclusions.

Extract from a meeting record sheet.

Date	Attendees	Number of ♂	Number of ♀	Location	Chair	Topics	Team
Mon 21 Jan	DC and staff	13	4	Ri Bhoi, Meghalaya	DC	DRM, planning, EOC, training, effectiveness	jc
Tue 22 Jan	Village DMC, volunteers, DMT	30	25	Urksen Wahpathaw, Meghalaya	Block Nodal Officer	Mock Drill, training, DRM effectiveness	jc
Tue 22 Jan	Block Officials	5	4	Urksew, Meghalaya	Block Nodal Officer	DRM, planning, training, impact	jc
Tue 22 Jan	State Steering Committee Members	14	6	Shillong, Meghalaya	Chief Minister	DRM, planning, budgets, plans	jc
Wed 23 Jan	Stake holders in Assam	42	6	Guhawati, Assam	State Secretary	DRM, EUVR, training, DRF, bye–lays, microzonation.	jc, ra, ms
Fri 25 Jan	Village DMC, volunteers, DMT	60	10	Guit, WB	ADM	DRM, Planning, effectiveness, training	jc, ra, ms
Fri 25 Jan	GP Members, NGOs, Volunteers	15	50	Rainai, WB	ADM	DRM Planning, impact, needs, effectiveness	jc, ra, ms
Fri 25 Jan	Block Officials	15	2	Khandachosh	BDO	DRM, training, impact, effectiveness	jc, ra, ms

Clearly, you should not include people in the individual interview and in the group interview list to avoid overstating the amount of consultation on which your evaluation is based.

Evaluation quality is a key factor in utilisation, and you need to demonstrate that your research is broadly based to persuade readers to apply the results of your evaluation.

TOOL FOR RECORDING THE TEAM ITINERARY

This simple tool records the itinerary for the evaluation team. This shows the location and activity undertaken by each team member during the evaluation. It is essential to keep this up to date during the evaluation or you will spend valuable hours of thinking time trying to remember who when to which village on which day during which week.

The itinerary is important as it establishes the trajectory of the teams evaluation, and allows readers to judge how well founded the opinions offered in the evaluation report are. An itinerary that shows extensive fieldwork at a number of sites increased the authority of the evaluation recommendations.

Date	Team leader	Team member 2	Team member 3	Team member 4

Excerpt from an evaluation team itinerary.

Mon 09 Apr	Team Meeting. Presentation of initial findings to HCT. Interview with World Vision	Team Meeting. Presentation of initial findings to HCT	Presentation of initial findings to the HCT. Interviews with World Vision and UN-Habitat representative	Presentation of initial findings to the HCT. Interview with UN-Habitat representative	Team Meeting. Presentation of initial findings to HCT
Tue 10 Apr	Interviews with the Resident Coordinator, WFP, Unicef, and the Mozambican Red Cross.	Travel to Caia via Beira. Field work in accommodation centers in Caia and Sena and travel to Mutarara	Travel to Caia, fieldwork in Mopeia area camps	Travel to Caia via Beira. Field work in accommodation centers in Caia and Sena and travel to Mutarara	Travel to Caia, fieldwork in Mopeia area camps

TOOL FOR KEEPING TRACK OF REFERENCES

An evaluation may make use of hundreds or even thousands of documents. If you are going to refer to more than a few handfuls of documents it makes sense to use specialist reference management software (or the bibliographic features offered by some office software suites). However, unless all team members have such software and know how to use it, you may find that one team member has to collate the references provided by other team members.

Collating references is much easier if they are in a consistent format. One approach is to provide team members with a spreadsheet template for recording references. This makes it much easier to collate the references.

You will need separate templates for:

- Books and reports
- Web pages
- Journal articles

Authors	Book or report title	Published by	In City	or date	URL if available on the web	viewed if on web	Notes	Comments/Abstract
Actionaid	The evolving UN cluster approach in the aftermath of the Pakistan earthquake : an NGO perspective.	ActionAid	London	2006	www.actionaid.org.uk/doc_lib/234_1_u n_cluster_approach.pdf	04-Sep-08	English version	Presents a critical NGO perspective of the cluster system. Points out areas where the system needs to improve, including the quality of leadership

Authors or institutional author	Title of web page	Published by	Last update date	Date accessed	URL	Notes
Norwegian Mission to the UN	Norway gives 57 millions USD to UN Fund	Norwegian Mission to the UN	08-Dec-06	02-Jul-07	http://www.norway-un.org/NorwegianStatements/OtherStat ements/donor+cerf.htm	Despite its teething troubles, Norway believes that the CERF has been a success. It is clear, however, that it requires ongoing development.

Authors	Article Title	Journal Title	Vol	Iss	t pag	End page	Year or Date	URL if available	Date of access is URL given	Notes
Dhar, S., Manzoor A. Halwai, Mohammed R. Mir, Zaid A. Wani, M. F. Butt, Masood I. Bhat and Arshiya Hamid	The Kashmir Earthquake Experience	European Journal of Trauma and Emergency Surgery	33	2	74	80	2007	http://www.springerlink.com/cont ent/v3173714q117n07j/	02-Aug-08	(URL to abstract only)

Tip: If you have the necessary software skills you can even have your spreadsheet generated a list that your reference manager programme can import, or even use mail-merge to create a list of formatted references.

RESOURCES FOR EHA

This is partially based on a list originally developed by John Borton in 2002.

PROFESSIONAL SOCIETIES

All evaluation societies organise conferences and meetings and keep their members informed of relevant developments, trends, publications and upcoming events. Some evaluation societies sponsor evaluation journals and members are offered reduced subscription rates. Increased membership of such societies by those involved in the evaluation of humanitarian action will help encourage the wider evaluation community to engage with the issues faced in the evaluation of humanitarian action. 'Southern' evaluation societies also offer a useful way for 'northern' managers of evaluation processes to establish contact with national and local evaluators and consultants in the country or the region where an evaluation is planned.

AFRICAN EVALUATION ASSOCIATION (AFrEA)

Umbrella association of 20 plus national evaluation networks and associations. UNICEF has been very active in supporting the formation and development of these networks and the African Evaluation Association Contact: Mahesh Patel mpatel@unicef.org

http://www.afrea.org/index.htm

AMERICAN EVALUATION ASSOCIATION

By far the largest professional evaluation society, the AEA provides a focus for evaluation managers, evaluators and researchers around the world. Members receive both *New Directions in Evaluation* and the *American Journal of Evaluation* as well as getting on-line access to other evaluation journals.

http://www.eval.org/

ASSOCIAZONE ITALIANA DI VALUTAZIONE

http://www.valutazione.it/

AUSTRALASIAN EVALUATION SOCIETY

Members of the Australasian Evaluation Society have free access to the journal *Evaluation and Program Planning.* The Society also publishes the *Evaluation Journal of Australasia* (free access to all but the current edition)

http://www.aes.asn.au/

CANADIAN EVALUATION SOCIETY

http://www.evaluationcanada.ca/

EUROPEAN EVALUATION SOCIETY

http://www.europeanevaluation.org/

DEUTSCHE GESELLSCHAFT FÜR EVALUATION

http://www.degeval.de/

INTER-AMERICAN ROUNDTABLE ON EVALUATION AND PERFORMANCE MEASUREMENT

A network of government departments, universities and professional associations in Latin America and the Caribbean involved in evaluations. Members include a Central American Evaluation Association

http://www.iadb.org/evo/roundtable/about.htm

INTERNATIONAL DEVELOPMENT EVALUATION ASSOCIATION

The International Development Evaluation Association (IDEAS) was inaugurated in September 2002 to help fill a gap in the international evaluation architecture. IDEAS' mission is 'to advance and extend the practice of development evaluation by refining methods, strengthening capacity and expanding ownership', with a particular focus on developing and transitional economies.

http://www.ideas-int.org/

MALAYSIAN EVALUATION SOCIETY

http://www.angelfire.com/ab/mes/

SOCIETE FRANÇAISE DE L'ÉVALUATION

http://www.sfe.asso.fr/

SRI LANKA EVALUATION ASSOCIATION (SLEvA)

http://www.naresa.ac.lk/sleva/profile.htm

SOCIETE SUISSE DE L'ÉVALUATION

http://www.seval.ch/

UK EVALUATION SOCIETY (UKES)

http://www.evaluation.org.uk/

JOURNALS

AMERICAN JOURNAL OF EVALUATION

(Formerly *Evaluation Practice*) published four times a year by Elsevier Science and sponsored by the American Evaluation Association

http://www.elsevier.com/inca/publications/store/6/2/0/1/8/2/index.htt

CANADIAN JOURNAL OF PROGRAM EVALUATION

Published twice a year by the University of Calgary Press for the Canadian Evaluation Society. It seeks to promote the theory and practice of program evaluation in Canada.

http://www.evaluationcanada.ca/site.cgi?s=4&ss=2&_lang=EN

EVALUATION JOURNAL OF AUSTRALASIA

The *Evaluation Journal of Australasia* is published twice a year. Issues of the Journal are available for downloading six months after publication.

http://www.aes.asn.au/publications/

EVALUATION REVIEW: A JOURNAL OF APPLIED SOCIAL RESEARCH

Published quarterly by Sage Publications. Members of the AEA have online access to this journal.

http://www.sagepub.com/journalsProdDesc.nav?prodId=Journal200935

EVALUATION: THE INTERNATIONAL JOURNAL OF THEORY, RESEARCH AND PRACTICE

SAGE publications (London, Thousand Oaks CA and Delhi)

www.sagepub.co.uk/journals/

EVALUATION AND PROGRAM PLANNING

Published quarterly by Elsevier Science. Member of the Australasian Evaluation have free access to this journal.

http://www.elsevier.com/inca/publications/store/5/9/3/

JOURNAL OF MULTI-DISCIPLINARY EVALUATION

This is a peer-reviewed open-access online journal. It focuses on education (which is a very large segment of evaluation effort in the US) but also deals with international evaluation.

 http://survey.ate.wmich.edu/jmde/index.php/jmde_1/issue/current

NEW DIRECTIONS FOR EVALUATION

(Titled *New Directions for Program Evaluation* until the end of 1997*)* is a themed journal that is published four times a year by Jossey-Bass for the American Evaluation Association.

http://www3.interscience.wiley.com/journal/117946202/grouphome/home.html

PRACTICAL ASSESSMENT, RESEARCH & EVALUATION

This is an peer-reviewed open access on-line journal supported entirely by volunteer efforts. Its purpose is to provide access to refereed articles that can have a positive impact on assessment, research, evaluation, and teaching practice. While some of the articles have a very technical bent, it also features some items of general evaluation interest.

http://pareonline.net/Home.htm

WEBSITES

ALNAP

http://www.alnap.org

As well as information on the ALNAP membership, ALNAP activities and the results of studies commissioned by ALNAP the site contains the Evaluative Reports Database.

The Evaluative Reports Database contains some 750 reports (April 2007), most of which are evaluations commissioned by ALNAP Member agencies.

The website formerly included a Useful Resources Database containing details of various books, reports, etc., identified as useful by ALNAP members and Secretariat. This facility was withdrawn as visit statistics showed that the usage did not justify the work needed to maintain it. Some of the documents previously held may be available on application to the ALNAP Secretariat in London.

EVALUATION PORTAL

http://www.evaluation.lars-balzer.name/

Lars Balzer is a Senior researcher with responsibility for evaluation at the Swiss Federal Institute for Vocational Education and Training. His Evaluation Portal presents hand-picked, human-edited, categorized information about the topic "evaluation" (and a bit about social science methods) that he found to be useful during my work in this field.

This is a very good site for links to lots of evaluation resources. Recommended.

OECD-DAC WORKING PARTY ON AID EVALUATION

http://www.oecd.org then click on "Development" and then on "Evaluation"

Contains WPAE publications, a database of evaluations undertaken by DAC member organisations and links.

CHARITIES EVALUATION SERVICE (CES)

http://www.ces-vol.org.uk

Offers evaluation services for NGOs and publishes tools and guidance

IPDET

http://www.ipdet.org

The International Program for Development Evaluation Training (IPDET) is an executive training program in development evaluation. IPDET is a collaboration of the Independent Evaluation Group of the World Bank and Carleton University. It is best known for its flagship program, an intensive and unique training opportunity offered once each year in Ottawa on the Carleton University campus in June and July. In 2008, 251 participants from over 80 countries attended the program.

M & E NEWS

http://www.mande.co.uk/news.htm

A news service focusing on developments in monitoring and evaluation methods relevant to development projects and programmes with social development objectives. Short summaries of planned and ongoing work, News update facility and good links to other Monitoring and Evaluation sites. The M & E News site is

supported by OXFAM (GB), Save the Children Fund (UK), Action Aid (UK), Water Aid, CAFOD and Christian Aid

PARC (PERFORMANCE ASSESSMENT RESOURCE CENTRE)

http://www.parcinfo.org/

Resource centre established in Birmingham, UK with support from DFID Evaluation Department

FREE RESOURCES FOR PROGRAM EVALUATION AND SOCIAL RESEARCH METHODS

http://gsociology.icaap.org/methods/

This page lists free resources for program evaluation and social research methods. The focus is on "how-to" do evaluation research and the methods used: surveys, focus groups, sampling, interviews, and other methods. Most of these links are to resources that can be read over the web.

DISCUSSION GROUPS

AMERICAN EVALUATION ASSOCIATION TIGS

http://www.eval.org/TIGs/tig.html

The AEA and its members maintain more than 30 Topical Interest Groups (TIGs) covering areas such as: Collaborative, Participatory and Empowerment Evaluation; Evaluation Managers and Supervisors; Human Services Evaluation, International and Cross Cultural Evaluation; Human Services Evaluation. Most TIGs have their own officers, means of communicating with members, and special events. All TIGs co-ordinate their efforts through the AEA and participate actively in AEA's annual conference. Each TIG receives conference paper proposals in their area of interest and sets up a series of paper sessions and panels for the conference. Members of AEA may join up to five Topical Interest Groups.

ALNAP QUALITY PROFORMA

(as published in ALNAP Review of Humanitarian Action in 2004)

To assess the quality of the Evaluation of Humanitarian Action (EHA) process through EHA reports

ASSESSING THE QUALITY OF HUMANITARIAN EVALUATIONS

THE ALNAP QUALITY PROFORMA 2005 (v. 02/03/05)

1. Background

ALNAP developed this Quality Proforma in 2000/2001 as a way of assessing humanitarian evaluation reports drawing on current thinking and good practice in the evaluation of humanitarian action[6]

The overall aim of the Quality Proforma is to improve the quality of humanitarian evaluation practice. It does this by:

1. Providing an assessment tool for ALNAP's annual meta-evaluation of humanitarian evaluation reports as part of its *Review of Humanitarian Action*[7] series. The meta-evaluation seeks to identify trends in the quality of humanitarian evaluations, identifying both good and weak practices.[8]

2. Providing a checklist for evaluation managers and evaluators.

The Quality Proforma has undergone refinements during its application in four ALNAP *Reviews* between 2001 and 2003/4, in order to strengthen consistency in interpretation and usage and reflect developments in current thinking in the evaluation of humanitarian action. This version of the Proforma has undergone a process of simplification and reordering for the *Review of Humanitarian Action in 2004* in order to make it more accessible.

2. Meta-evaluation process

Each evaluation report included in ALNAP's meta-evaluation is rated against the Quality Proforma by two assessors working independently. For each report, every area of the criteria is given a comment and a rating. The ratings are then used to assess strengths and weaknesses of the set as a whole.

Since 2003/4, the draft findings of the Quality Proforma assessments have been discussed with a selection of the commissioning agencies in order to better understand the background to the evaluation process, gather information that may not show up in the written report and stimulate agency involvement and interest. The outcome of these discussions may lead to revisions of the final assessments. In 2005 for the first time, a selection of evaluators will also be consulted on the evaluation processes.

3. Using the ALNAP Quality Proforma

The development of the Proforma is linked to ALNAP's definition of the Evaluation of Humanitarian

---

[6] Sources used in the development of the Proforma are listed at the end of this document.

[7] The *Annual Review* series was renamed *Review of Humanitarian Action* series in 2004.

[8] Two assessors are used for the meta-evaluation exercise to mitigate potential assessor bias

Action (EHA) given in the box below.

The Proforma is intended to be used for reports dealing with natural disasters and complex political emergencies. It should also be of value for other types of evaluative exercises in the humanitarian context. Although originally designed with programme evaluations in mind, the Proforma can also be used to review evaluations of such activities as humanitarian management processes, funding partnerships and sectoral approaches. In these cases, some questions in the Proforma may be noted as not relevant.

ALNAP Definition of Evaluation of Humanitarian Action (EHA)

"A systematic and impartial examination of humanitarian action intended to draw lessons to improve policy and practice, and enhance accountability. It has the following characteristics: i). it is commissioned by or in cooperation with the organisation(s) whose performance is being evaluated; ii). it is undertaken either by a team of non-employees (external) or by a mixed team of non-employees (external) and employees (internal) from the commissioning organisation and/or the organisation being evaluated; iii). it assesses policy and/or practice against recognised criteria (e.g., efficiency, effectiveness/timeliness/coordination, impact, connectedness, relevance/appropriateness, coverage, coherence and as appropriate, protection); and, iv). it articulates findings, draws conclusions and makes recommendations." ALNAP 2001, *Humanitarian Action: Learning from evaluation*, ALNAP Annual Review 2001. London: ALNAP/ODI.

The Quality Proforma is divided into six sections:

1. Assessing the Terms of Reference;
2. Assessing Evaluation Methods, Practice and Constraints;
3. Assessing Contextual Analysis;
4. Assessing the Intervention;
5. Assessing the Report;
6. Overall Comments.

Each section has four column headings:

- **Area of Enquiry** (subject matter)
- **Guidance Notes** (guidance as to what is deemed 'satisfactory' to ensure a degree of consistency of interpretation)
- **Comments** (to include a brief reason for the rating given)
- The **Rating**.

The rating system used for the meta-evaluation is as follows:

A = Good
B = Satisfactory
C = Unsatisfactory
D = Poor
Z = Not applicable. (Where an area of enquiry is deemed not applicable, reasons should be given in the 'Comments' column. The proforma user's judgement remains a central factor in the rating exercise.)

Where the Guidance Note lists a number of areas that should be covered for an Area of Enquiry, a 'B' (Satisfactory) rating will normally only be given if the report is judged to be satisfactory in all those areas.

In some cases, the assessors may note in the Comments section that the rating is borderline, indicating that it is a matter of fine judgement as to whether the rating falls into one category or another. This most

often happens when the assessors are deciding between B or C ratings.

The Glossary defines many of the terms used in this Proforma.

The ALNAP proforma

EVALUATION TITLE	
COMMISSIONING AGENCY	
DATE OF REPORT	
NAME AND POSITION OF ASSESSOR	
REASON FOR ASSESSMENT	
DATE OF ASSESSMENT	
DATE OF AGENCY INTERVIEW (if held)	

Section 1. Assessing the Terms of Reference (ToR)

Area of enquiry	Guidance Notes	Comments	Rating
1.1 The Terms of Reference	The ToR should clearly describe: The work to be evaluated including its objectives and key stakeholders. The purpose, objectives and focus of the evaluation *(Purpose might be donor requirement, accountability, lesson learning, community empowerment. Focus might be on partner performance, programme, project, policy, institutional analysis, sector, coordination).* The intended use and users of the evaluation outputs and the individual or department responsible for follow-up. The desired report framework. (*A sample framework is outlined in Annex 2*). The rationale for the timing of the evaluation. The evaluator selection process (e.g., competitive bidding, standing offer).		
1.2 Expectation of good evaluation practice	The TOR should clarify the commissioning agency's expectation of good humanitarian evaluation practice. *(e.g., application of DAC criteria;[9] reference to international standards including international law; multi-method approach i.e., quantitative and qualitative; consultation with key stakeholders to inform findings, conclusions and recommendations; and gender analysis).*		

Section 2. Assessing Evaluation Methods, Practice and Constraints

Area of enquiry	Guidance Notes	Comments	Rating
2.1 Nature, make up and appropriateness and biases of the evaluation team	The report should outline the nature (e.g., external or mixed) and make up of the team (e.g., sectoral expertise, local knowledge, gender balance) and its appropriateness for the evaluation. The evaluation report should outline the evaluator(s)' biases that might have affected the evaluation and how these have been counteracted.		

---

[9] See Section 5.3 for criteria definitions drawn from OECD/DAC (1999) *Guidance for Evaluating Humanitarian Assistance in Complex Emergencies*, Paris.

2.2 Clarification process	The evaluation report should outline any clarification process between the commissioning agency and the evaluation team about the scope and methods of the evaluation that resulted in modifications to the ToR.		
2.3 Appropriateness of the overall evaluation methods	The evaluation methods should be clearly outlined in the report and their appropriateness, relative to the evaluation's primary purpose, focus and users, should be explained pointing out the strengths and weaknesses of the methods.		
2.4 Consultation with and participation by primary stakeholders	The evaluation report should outline the nature and scope of consultation with, and participation by, beneficiaries and non-beneficiaries within the affected population in the evaluation process. (A satisfactory or higher rating should only be given where evidence is presented of adequate consultation and participation of primary stakeholders in the evaluation process, or where, in the assessor's view, it has been successfully argued as inappropriate due to security or other reasons.) The evaluation report should outline the nature and scope of consultation with other key stakeholders in the evaluation process. The report should include a list of the other key stakeholders who were consulted or who participated in the evaluation process.		
2.5 The use of and adherence to international standards	The evaluation report should assess the intervention against appropriate international standards (e.g., international humanitarian and human rights law; the Red Cross/ NGO Code of Conduct, Sphere).		
2.6 Evaluation constraints	The evaluation report should outline key constraints to carrying out the evaluation (e.g., lack of time, difficult travelling conditions, lack of baseline data, poor agency monitoring systems, lack of access to key information sources, difficulties setting up control groups, use of translators), and the effect of these constraints.		

Section 3. Assessing Contextual Analysis

Area of enquiry	Guidance Notes	Comments	Rating
3.1 Analysis of context and of the crisis to which the intervention is responding	The evaluation report should provide analysis of the affected area and population *(including relevant historical, social, economic, political and cultural factors)* to inform the evaluation and draw on this information in the text to support the analysis of the intervention. The evaluation report should provide a clear analysis of the crisis, including key events (and a chronology where appropriate).		
3.2 Past involvement of the agency and its local partners	The evaluation report should provide analysis of the implementing agency's and its local partners' past involvement and main areas of work, so that the influence of the agency's past involvement on the intervention, including its geographical and sectoral focus, can be understood.		

Section 4. Assessing the Intervention

4.1 Institutional Considerations			
Area of enquiry	**Guidance Notes**	**Comments**	**Rating**
4.1.i The agency's guiding policies and principles	The evaluation report should provide an analysis of the extent to which agency policies and principles were applied, and their relevance to and effect on the intervention.		
4.1.ii The agency's management and human resources	The evaluation report should provide an analysis of the agency's management and human resource procedures and practices as applied and their effect on the intervention. *(This might include: level of experience/expertise of field staff; use of national and expatriate staff; staff turnover; field/HQ communications & relations; briefing and debriefing procedures; training and learning practices; security)*		

4.2 Needs Assessment, Objectives, Planning and Implementation			
Area of enquiry	**Guidance Notes**	**Comments**	**Rating**
4.2.i The needs and livelihoods assessments that informed the intervention	The evaluation report should provide analysis of the needs and livelihoods assessment practices that informed the intervention and their effect on the intervention.		
4.2.ii Intervention objectives	The evaluation report should assess the relevance of the intervention objectives to the contextual analysis and needs/livelihoods assessments assessed in 3.1 and 4.2.i above.		
4.2.iii Programme cycle processes.	The evaluation report should provide analysis of the following processes and their effect on the intervention: planning implementation monitoring and/or real-time evaluative mechanisms intervention expenditure. *(Consideration in this analysis should be given to local capacities; primary stakeholder consultation and participation; local and national partnerships)*		

4.3 Application of EHA Criteria			
Area of enquiry	**Guidance Notes**	**Comments**	**Rating**
	The evaluation report should provide evidence of an adequate application of standard evaluation of humanitarian action criteria as per the OECD/DAC definitions given below:[10]		
4.3.i Efficiency (including cost-effectiveness)	Efficiency measures the outputs - qualitative and quantitative - in relation to the inputs. This generally requires comparing alternative approaches to achieving the same outputs, to see whether the most efficient process has been used. Cost-effectiveness looks beyond how inputs were converted into outputs, to whether different outputs could have been produced that would have had a greater impact in achieving the project purpose.		
4.3.ii Effectiveness (including timeliness)	Effectiveness measures the extent to which the activity achieves its purpose, or whether this can be expected to happen on the basis of the outputs. Implicit within the criteria of effectiveness is **timeliness** of the response. (*Although **coordination** is not a formal criterion, the OECD/DAC Guidance suggests that given its importance, it should be considered under this criterion*).		
4.3.iii Impact	Impact looks at the wider effects of the project - social, economic, technical, environmental - on individuals, gender, age-groups, communities, and institutions.		
4.3.iv Relevance/ appropriateness	Relevance is concerned with assessing whether the project is in line with local needs and priorities (as well as donor policy). It refers to the overall goal and purpose of a programme. Appropriateness - the need to tailor humanitarian activities to local needs, increasing ownership, accountability, and cost-effectiveness accordingly … is more focused on the activities and inputs. (Minear, 1994)		

---

[10] from OECD/DAC (1999) *Guidance for Evaluating Humanitarian Assistance in Complex Emergencies*. Paris, pp 30-32.

4.3.v Sustainability/ connectedness	Sustainability is concerned with measuring whether an activity or an impact is likely to continue after donor funding has been withdrawn. … many humanitarian interventions, in contrast to development projects, are not designed to be sustainable. They still need assessing, however, in regard to whether, in responding to acute and immediate needs, they take the longer-term into account. (*Minear (1994) has referred to this as connectedness, the need… to assure that activities of a short-term emergency nature are carried out in a context which takes longer-term and inter-connected problems into account.*)		
4.3.vi Coverage	The need to reach major population groups facing life-threatening suffering wherever they are, providing them with assistance and protection proportionate to their need and devoid of extraneous political agendas.		
4.3.vii Coherence	Coherence refers to policy coherence, and the need to assess security, developmental, trade and military policies as well as humanitarian policies, to ensure that there is consistency and, in particular, that all policies take into account humanitarian and human rights considerations.		

4.4 Consideration given to Cross-cutting Issues			
Area of enquiry	**Guidance Notes**	**Comments**	**Rating**
4.4.i The use of and adherence to international standards	The evaluation report should assess the extent to which relevant international standards were used in the planning, implementation and monitoring of the intervention *(e.g., international humanitarian and human rights law; the Red Cross/ NGO Code of Conduct and developing standards - e.g., Sphere)*		
4.4.ii Gender Equality	The evaluation report should analyse consideration given to gender equality throughout the intervention and the effect on the intervention. *(i.e. was gender equality taken into consideration in all relevant areas? Did the intervention conform to the implementing organisation's gender equality policy? It should be noted if there is no gender equality policy).*		
4.4.iii Protection	The evaluation report should analyse the consideration given to protection throughout the intervention cycle and the effect on the intervention.		
4.4.iv Capacity building	The evaluation report should analyse the consideration given to the capacity building of key and primary stakeholders government and civil society institutions, and the effect of this on the intervention.		
4.4.v Advocacy	The evaluation report should analyse consideration given to advocacy and the effect on the intervention. *(e.g., attempts to influence donors, partners, government, concerning their policies or actions).*		
4.4.vi Vulnerable and marginalised groups	The evaluation report should provide an analysis of consideration given to vulnerable and marginalised groups (e.g., elderly, disabled, children, HIV/AIDS sufferers) and to other groups that suffer discrimination and disadvantage.		

Section 5. Assessing the Report

5.1 Findings, Conclusions and Recommendations			
Area of enquiry	**Guidance Notes**	**Comments**	**Rating**

5.1.i Secondary sources	The evaluation report should use and refer to relevant secondary sources to support its findings, conclusions and recommendations (*a satisfactory or higher rating should only be given where a reference list of secondary sources is included as part of the report*).		
5.1.ii Conclusions	The report's conclusions should flow logically from, and reflect, the report's central findings. The report should provide a clear and defensible basis for value judgements in each case.		
5.1.iii Recommendations	Recommendations should be clear, relevant and implementable, reflecting any constraints to follow up. Recommendations should follow on from the main conclusions and reflect consultation with key stakeholders. The evaluation report should suggest a prioritisation of recommendations, timeframe for implementation and suggest where responsibility for follow-up should lie if that is not indicated in the ToR.		

5.2 Report Coverage, Legibility and Accessibility			
Area of enquiry	**Guidance Notes**	**Comments**	**Rating**
5.2.i Coverage of the evaluation report	The evaluation report should adequately cover all areas specified in the ToR and additional factors that affected the performance of the intervention.		
5.2.ii Format of the report	The evaluation report format should follow that outlined in the ToR (*if the ToR did not propose a format for the report, this area should be assessed on the basis of the good practice suggested in Annex 2*).		
5.2.iii Accessibility of the report	The evaluation report should cater for the intended readership and users (*In general reports should use language clearly; be succinct; be clearly laid out e.g. with different information levels and appropriate visual aids. Some organisations have their own style guides*).		
5.2.iv Executive Summary	The executive summary should reflect the format of the main text, and clearly outline key evaluation conclusions and recommendations.		

Section 6. Overall Comments (*for information purposes and not rated*)

Area of enquiry	**Guidance Notes**	**Comments**
6.i Comments on issues not covered above.	This is an opportunity for comment on any issues not covered by the areas of enquiry.	
6.ii Overall comments on the report.	This is an opportunity to make an overall comment on the report, including its strengths and weaknesses.	

ALNAP proforma: glossary

Accountability

Accountability is the means by which individuals and organisations report to a recognised authority, or authorities, and are held responsible for their actions. (Edwards & Hulme, 1995).

Advocacy

Advocacy refers in a broad sense to efforts to promote, in the domain of humanitarian aid, respect for humanitarian principles and law with a view to influencing the relevant political authorities, whether recognised governments, insurgent groups or other non-

state actors. (SDC, 2004).[11] One could add "international, national and local assistance agencies".

Appropriateness

The need to "tailor humanitarian activities to local needs, increasing ownership, accountability, and cost-effectiveness accordingly" (Minear 1994) is more focused on the activities and inputs. (ALNAP Annual Review Glossary 2003)

Coherence

Refers to the policy coherence and the need to assess security, developmental, trade and military policies to ensure that there is consistency and, in particular, that all policies take into account humanitarian and human rights considerations. (DAC Evaluation Criteria)

Complex political emergency

A situation with complex social, political and economic origins which involves the breakdown of state structures, the disputed legitimacy of host authorities, the abuse of human rights and possibly armed conflict, that creates humanitarian needs. The term is generally used to differentiate humanitarian needs arising from conflict and instability from those that arise from natural disasters. (ALNAP Annual Review Glossary 2003)

Conclusions

Conclusions point out the factors of success and failure of the evaluated intervention, with special attention paid to the intended and unintended results and impacts, and more generally to any other strength or weakness. A conclusion draws on data collection and analyses undertaken through a transparent chain of arguments. (OECD/DAC Glossary of Key Terms in Evaluation and Results-Based Management, 2002)

Context (of an evaluation)

The combination of factors accompanying the study that may have influenced its results, including geographic location, timing, political and social climate, economic conditions, and other relevant professional activities in progress at the same time. (Programme Policy and Procedures Manual, UNICEF, May 2003)

Cost Effectiveness Analysis (see also 4.3.i above)

Cost-effectiveness analysis entails comparing costs across different strategies for achieving a given outcome, with a view to determining the lowest cost approach. For example, cost-effectiveness analysis might explore three different approaches to getting girls working in the informal sector back into school. As compared to cost-efficiency analysis, it is wider in scope, looking beyond outputs to outcomes. (M&E Training Resources, UNICEF, 2004)

Coverage

The need to reach major population groups facing life-threatening suffering wherever they are, providing them with assistance and protection proportionate to their need and devoid of extraneous political agenda. (DAC Evaluation Criteria)

Effectiveness

Effectiveness measures the extent to which the activity achieves its purpose, or whether this can be expected to happen on the basis of the outputs. Implicit within the criteria of effectiveness is timeliness of the response. Although coordination is not a formal criterion, the OECD/DAC Guidance suggests that given its importance, it should be considered under this criterion. (DAC Evaluation Criteria)beffectiveness

---

[11] Definitions of advocacy within the humanitarian sector appear to be very limited (SDC, 2004).

Humanitarian action

Assistance, protection and advocacy actions undertaken on an impartial basis in response to human needs resulting from complex political emergencies and natural hazards. (ALNAP Annual Review Glossary 2003)

Impact

Impact looks at the wider effects of the project - social, economic, technical, environmental - on individuals, gender, age-groups, communities, and institutions. (DAC Evaluation Criteria)

Impartiality

An approach to the provision of humanitarian assistance and services which is non-discriminatory, proportionate to needs and free of subjective distinction. A guiding principle of organisations claiming to be humanitarian. (ALNAP Annual Review Glossary 2003)

Input

The financial, human, material, technological and information resources used for the intervention. (OECD/DAC Glossary of Key Terms in Evaluation and Results-Based Management Proposed Harmonized Terminology, 2002)

Lesson learned

Conclusions that can be generalized beyond the specific case. This could include lessons that are of relevance more broadly within the country situation or globally, to an organization or the broader international community. (Programme Policy and Procedures Manual, UNICEF, May 2003)

Lesson-learning study

A study initiated by an organisation with the explicit objective of lesson-learning within that organisation, but that falls outside the full evaluation definition. A process that may be facilitated by external consultants but is generally an internal process.(ALNAP Annual Review Glossary 2003)

Meta-evaluation

Simply stated, meta-evaluation is the evaluation of an evaluation, evaluation system or evaluation device (Hummel 2003). A process of delineating, obtaining, and applying descriptive information and judgmental information – about the utility, feasibility, propriety and accuracy of an evaluation and its systematic nature, competent conduct, integrity/honesty, respectfulness and social responsibility – to guide the evaluation and/or report its strengths and weaknesses (Stufflebeam)

Outcome

The intended or achieved short-term and medium-term effects of an intervention's outputs, usually requiring the collective effort of partners. Outcomes represent changes in conditions which occur between the completion of outputs and the achievement of impact. (OECD/DAC Glossary of Key Terms in Evaluation and Results-Based Management Proposed Harmonized Terminology, 2002)

Output

The products and services which result from the completion of activities within an intervention. (OECD/DAC Glossary of Key Terms in Evaluation and Results-Based Management Proposed Harmonized Terminology, 2002)

Protection

Activities aimed at obtaining full respect for the rights of the individual in accordance with the letter and the spirit of the relevant bodies of law (i.e. human rights,

humanitarian and refugee law) which are conducted impartially and not on the basis of race, national or ethnic origin, language or gender. (ALNAP Annual Review Glossary, 2003)

Relevance

Relevance is concerned with assessing whether the project is in line with local needs and priorities (as well as donor policy) … refers to the overall goal and purpose of a programme. (DAC Evaluation Criteria)

Retrospectively, the question of relevance often becomes a question as to whether the objectives of an intervention or its design are still appropriate given changed circumstances. (OECD/DAC Glossary of Key Terms in Evaluation and Results-Based Management, 2002)

Stakeholder

All those – from agencies to individuals – who have a direct or indirect interest in the humanitarian intervention, or who affect or are affected by the implementation and outcome of it. (ALNAP Annual Review Glossary 2003). Within the context of the Quality Proforma **'primary stakeholders'** refers to both beneficiaries and non-beneficiaries within the affected population.

Sustainability

Sustainability 'is concerned with measuring whether an activity or an impact is likely to continue after donor funding has been withdrawn … many humanitarian interventions, in contrast to development projects, are not designed to be sustainable. They still need assessing, however, in regard to whether, in responding to acute and immediate needs, they take the longer term into account. (DAC Evaluation Criteria). Minear has referred to this as *Connectedness. Connectedness*, the need "to assure that activities of a short term emergency nature are carried out in a context which takes longer-term and inter-connected problems into account" (Minear, 1994).

Terms of Reference

Terms of reference define the requirements and paramters for conducting an evaluation. (ALNAP Annual Review Glossary 2003)

ALNAP proforma: evaluation report format - check list

Preliminaries	
	Title page (should include date of report)
	List of contents with page numbers
	Acronyms
	Map(s)
	Executive Summary
Main text	
	Introduction (including motivation for commissioning evaluation, purpose of study, scope, approach, methods, composition of team, constraints)
	Context in which humanitarian action took place, humanitarian context and response
	Findings
	Conclusions
	Recommendations
Annexes	
	Sources/bibliography

	ToR
	Timetable
	Evaluation team profiles
	List of Interviewees
	Timeline
	Evaluation Material (questionnaires etc)
	Collated stakeholder feedback on findings, conclusions and recommendations
	Other appendices/annexes

ALNAP Proforma: References

ALNAP (2003) *Humanitarian Action: Improving Monitoring to Enhance Accountability and Learning. ALNAP Annual Review 2003.* London: ALNAP.

Apthorpe, R. (2000) K*osovo Humanitarian Programme Evaluations: Towards Synthesis, Meta-Analysis and Sixteen Propositions for Discussion.* Background Paper for the ALNAP Symposium '***Learning*** from Evaluation: Humanitarian Assistance and Protection in Kosovo' 17th & 18th October 2000, Geneva (London: ALNAP)

Minear, L. (1994) *The International Relief System: a critical review.* Paper presented to the Parallel National Intelligence Estimate on Global National Emergencies, Meridian International Centre, Washington DC, September 2002.

OECD/DAC (1999) *Guidance for Evaluation Humanitarian Action in Complex Emergencies*. Paris: OECD/DEC

OECD/DAC (2002) *Glossary of Key Terms in Evaluation and Results-based Management Proposed Harmonized Terminology.* Paris: OECD/DAC.

Patton, M. (1997) *Utilization-Focused Evaluation: The New Century Text* (Thousand Oaks: Sage)

Raynard, P. (2000) *Mapping Accountability in Humanitarian Assistance* (London: ALNAP)

SPHERE (2000) *Humanitarian Charter and Minimum Standards in Disaster Response* (Geneva: The Sphere Project)

SDC (Swiss Agency for Development and Cooperation) 2004 *Advocacy Guidelines: Humanitarian Aid of the Swiss Confederation.* Berne: SDC, March.

UNICEF (2003) *Programme Policy and Procedures Manual.* UNICEF, May

Valadez, J. and M. Bamberger (1994) *Monitoring and Evaluating Social Programs in Developing Countries: A Handbook for Policymakers, Managers and Researchers* (Washington DC: World Bank EDI Development Series)

Wood, A., Borton, J. and R. Apthorpe (eds.) (2001) Evaluating *International Humanitarian Action: Reflections from Practitioners* (London: Zed Press)

Refs