CHAPTER 1
INTRODUCTION

1.1   Purpose and Scope of the Annual Review Series

This publication is the first in the ALNAP Annual Review series and forms part of ALNAP's broader efforts to support and improve learning and accountability within the international humanitarian system[1]. The Annual Review series is intended to:

- provide the humanitarian system with a means to reflect annually on its performance;
- increase awareness of the principal findings of evaluations of humanitarian action published over the previous year;
- monitor, analyse and report on trends in the evaluation of humanitarian action;
- encourage improvements in the quality of evaluation work undertaken, and develop and publicise assessments of the quality of evaluation of humanitarian action reports.

It is also intended to complement other publications focusing on particular aspects of, and issues related to, humanitarian action and the international humanitarian system, such as the *World Disasters Report* (IFRC 2000), *Global Humanitarian Assistance 2000* (IASC 2000) and *The State of the World's Refugees* (UNHCR 2000).

The series will make extensive use of evaluations of humanitarian action (defined later in this chapter) in its analysis, and prove of considerable relevance to a broad spectrum of those directly or indirectly engaged with or within the humanitarian system.

At its best, when not undermined by poor practice and inappropriate resourcing, evaluation of humanitarian action provides an essential tool for the system. It enables organisations and individuals to draw learning from experience through a systematic and impartial scrutiny of performance.

The potential of evaluation of humanitarian action to facilitate system-wide scrutiny and learning becomes all the more powerful when considered and analysed *en masse*. This first Annual Review synthesises 49 such evaluation reports and 5 existing syntheses, published between 1999 and 2000 and held on the ALNAP Evaluative Reports Database (see p6) – 17 of these focus on humanitarian action in Kosovo.

The story it tells is pertinent to humanitarian policy, advocacy, knowledge management, evaluation and operations. It will equally be of interest to parliamentary and congressional committees that scrutinise public expenditure on humanitarian action; journalists and researchers that focus on humanitarian issues; and those providing training to humanitarian personnel.

In addition, assessment of the quality of the evaluation reports reviewed will be of particular interest to evaluation managers, the independent consultants who undertake evaluations of humanitarian action, and those involved in training evaluation personnel.

1.1.1   Chapter 1: Setting the scene

This chapter establishes the context for the Annual Review by considering the profound changes that have affected and influenced the international humanitarian system over the past decade, and the subsequent demands for improved mechanisms for accountability and learning. It introduces and defines the 'evaluation of humanitarian action', traces the rapid growth in its use, and considers the differences (as well as the similarities) between it and the evaluation of development co-operation.

The chapter reflects on the degree to which evaluation has become ingrained within the international humanitarian system; on the balance between evaluation's accountability and learning objectives; and on the current gaps in evaluation guidance material. Finally, it introduces ALNAP's development and application of an evaluation reports quality assessment proforma.

The remainder of the Annual Review is dedicated, through the synthesis and meta-evaluation of the 1999–2000 evaluations, to drawing out and highlighting strengths and weaknesses in current practice, successes and failures, and the dilemmas, lessons and challenges (old and new) facing the system and the evaluation of humanitarian action genre.

To exploit fully the opportunity for analysis provided by the Kosovo set, a series of evaluations resulting from and focussing on humanitarian action within the same context, the synthesis and analysis of the evaluation reports has been split over two chapters.

1.1.2 Chapter 2: the non-Kosovo set

Chapter 2 reviews the 37 non-Kosovo evaluations. Amongst these are major evaluation efforts: evaluations of ECHO's operations 1991–1996 and 1996–2000, and Danida's humanitarian assistance 1992–98; evaluations of responses to major humanitarian emergencies such as Central American countries affected by Hurricane Mitch in 1998; flooding in China and Bangladesh in 1998–99; food insecurity in North Korea during the late 1990s; civil conflict in East Timor in 1999; and evaluations of humanitarian action in continuing conflict countries such as Afghanistan, the Democratic Republic of the Congo, Sudan and the Caucasus.

1.1.3 Chapter 3: the Kosovo set

The international response to the humanitarian needs created by the 1999 Kosovo conflict ranks as one of the largest, in terms of the scale of resources. It is also one of the most contentious, with many nations involved in the provision of humanitarian assistance also involved, via their military forces and NATO, as protagonists in the conflict between NATO and the Federal Republic of Yugoslavia. This has been the subject of more than 20 separate evaluations of humanitarian action of which 16 have been published or made available to ALNAP.

Chapter 3 is dedicated to the 'Kosovo' evaluation reports, enriching its analysis through reference to some 50 other Kosovo evaluative reports. It presents a synthesis and comments on the quality of the evaluation reports in terms of their contribution to institutional learning. The analysis is also informed by discussions at the ALNAP *Learning from Evaluation: Humanitarian Assistance and Protection in Kosovo* symposium, in Geneva in October 2000[2].

1.1.4 Chapter 4: Conclusions

Chapter 4 draws together and reflects on the findings of the preceding chapters. In so doing, it outlines an agenda of issues to be addressed by the international humanitarian system and by those involved in the commissioning, undertaking and use of evaluations of humanitarian action. Based as it is on 49 evaluations published over a two-year period covering operations in 36 countries in 4 regions of the world, the agenda it sets out is robust and comprehensive.

1.2 A Decade of Rapid Change

The publication of this first ALNAP Annual Review coincides with the tenth anniversary of Operation Provide Comfort[3], an unprecedented military intervention and humanitarian operation launched in

April 1991 to provide assistance to displaced Kurds trapped in the mountains along the Iraqi/Turkish border. The operation contributed to a sharp increase in the overall annual expenditure on humanitarian assistance (see Figure 1.1) and marked the start of a decade of frenetic activity for the international humanitarian system. High levels of expenditure were sustained throughout and, boosted by the Kosovo operations, in 1999 ended at its highest level ever.

This ten-year period witnessed an unprecedented number of large-scale humanitarian operations (notably Somalia, southern Africa, former Yugoslavia, Rwanda, Central America, Kosovo, and East Timor) many involving armed intervention in the form of peace enforcement or peacekeeping operations. Humanitarian assistance shifted from a focus primarily on refugees in asylum countries to a post-Cold War regime characterised by the provision of assistance to populations within areas of ongoing conflict.

Fig. 1.1 Humanitarian Assistance in Real Terms (1998 prices and exchange rates)
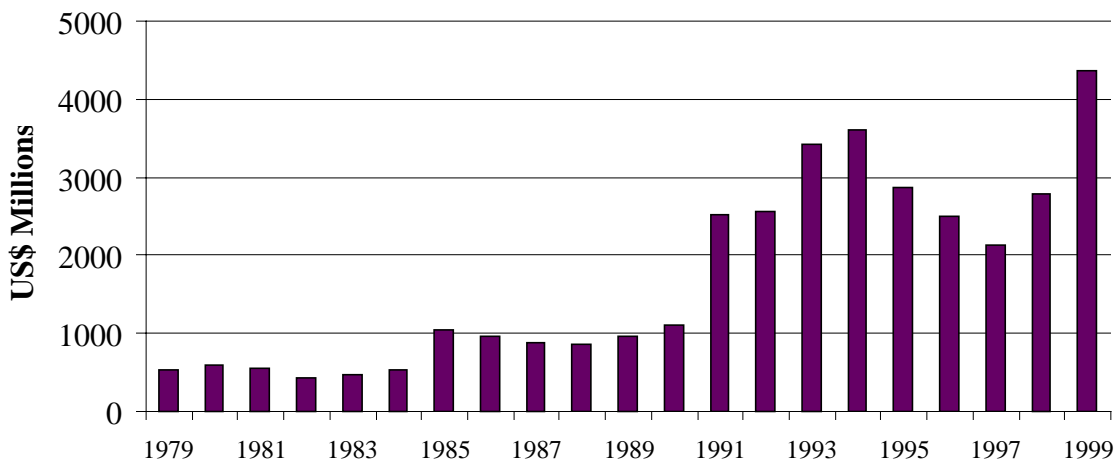
Source: IASC, 2000 (updated from OECD DAC figures 2001)

It also saw an unprecedented debate on humanitarian policies and practices and the development of fundamental critiques of humanitarianism (e.g. de Waal, 1997), debates and critiques that continue today. How should humanitarian assistance be provided in contexts where it is often impossible to differentiate civilians from soldiers and where assistance can contribute to the fighting ability of parties to the conflict? What is the appropriate relationship between humanitarian organisations and military contingents engaged in peacekeeping or peace-enforcement? How can the safety of humanitarian staff working in areas of active conflict be ensured?

It is important to recognise that the developments within the international humanitarian system took place against a backdrop of fundamental change in the management of national and international public sector organisations, resulting from the ideology of neo-liberalism. Behind the political rhetoric of downsizing government, increasing quality and value for money in public services, programmes were introduced under headings such as 'total quality management' or 'results-based management'. These aimed at improving the clarity of objectives and the measurement of processes and results (see Kirkpatrick and Martinez Lucio, 1995; Wholey, 1997). Though contentious, such programmes did

contribute to an increased transparency of processes and, to an extent, of organisational accountability in the public sector.

Similar programmes have permeated through to the international humanitarian system as it has also attempted to take on board the new management approaches. A significant source of change has been the bilateral donor organisations and the requirements attached to their funding contributions to multilateral organisations, UN agencies and non-governmental organisations (NGOs).

Fig. 1.2 Evaluations in ERD by year of publication
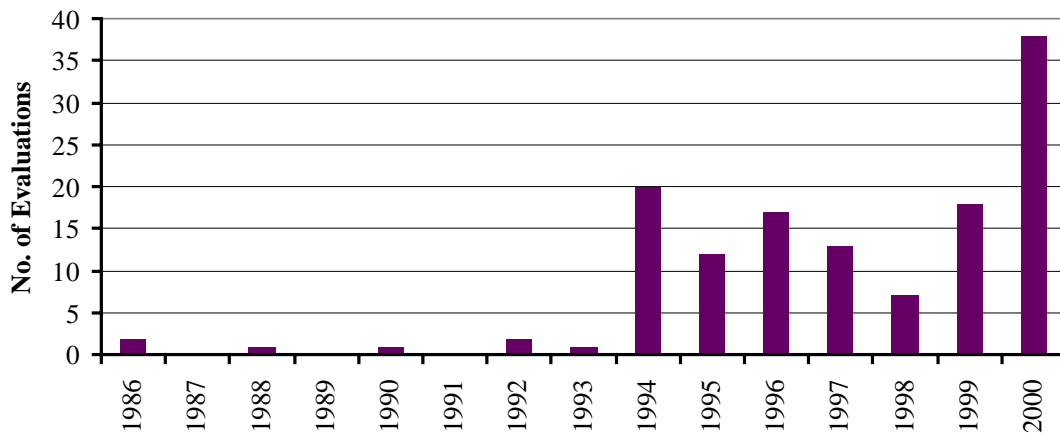
Responding to these new demands and pressures, the system has in the past decade instigated and witnessed a range of initiatives aimed at improving its own transparency and accountability. These include the development of codes of conduct, benchmarks and standards and the exploration of mechanisms for encouraging compliance with these codes and standards (see Box 1.1). Significantly, it was a system-wide evaluation of the international humanitarian response to the genocide in Rwanda (JEEAR, 1996) that provided a final impetus for some of these, including ALNAP.

1.3   Evaluation: a Key Tool for Accountability and Learning

The changes and pressures of the 1990s, combined with wider management changes, resulted in an unprecedented level of scrutiny of the system, both by outsiders, and from within. Bilateral donor organisations began to apply their new management approaches to their funding and therefore to those they funded, but these management strategies were complemented by growing recognition across the system of the need to learn from experience and translate learning into improved practice.

A powerful spur to this new-found commitment to improving learning and performance was the increased competition for funds, coupled with the realisation that greater accountability and transparency were necessary to maintain the support of bilateral and multilateral donors.

A wide range of mechanisms have been involved in both scrutiny and learning, including commissions of enquiry, parliamentary committees, evaluations, reviews, learning workshops, media coverage and research studies. While all play a useful role for the international humanitarian system, evaluation has

emerged as a key tool for assessing how effectively resources have been used and what lessons might be learnt.

1.3.1 Evaluation of humanitarian action

Evaluation has been variously defined, but its central concepts are those described by Scriven in his *Evaluation Thesaurus*: 'the process of determining the merit, worth or value of something, or the product of that process,' (Scriven, 1991: p139). The application of evaluation to particular areas of activity has resulted in more specific definitions, reflecting the particularities of those activities. The definition of evaluation of humanitarian action is introduced in Section 1.4.

Compared with the application of evaluation to development cooperation (see Box 1.2 above), its application to humanitarian action has been slower. This was partly the result of initial resistance but also of technical, practical and methodological difficulties (discussed later). The first evaluations of humanitarian action weren't undertaken until the second half of the 1980s. On the evidence available, 1993 appears to have been the year in which the evaluation of humanitarian action took off (see Figure 1.2). Since then, however, a boom has been underway in the use of evaluation by the humanitarian system. Interestingly the sharp increase in the number of evaluations in 1993 appears to follow on from the 1991 funding increase (see Figure 1.1) with an approximately two-year time lag.

What does this boom represent and how long can it be expected to last? Is it bringing genuine improvements to the accountability and learning processes within the international humanitarian system? These questions are central to this Annual Review.

The boom undoubtedly represents a significant investment by the humanitarian system, and presents a considerable opportunity for critical reflection and learning in humanitarian operations. While information on the overall level of resources devoted to evaluation of humanitarian action is not readily available, evidence from a benchmarking study, being undertaken by the ICRC[4] at time of publication, is indicating that humanitarian agencies devote on average the equivalent of 0.5% of their humanitarian action expenditure to inspection, audit and evaluation activities.

However, if evaluation is to continue to receive its current levels of attention and resourcing, and be embraced by all – whether at policy or operational level – it needs to improve on the quality of the undertaking and its product to demonstrate clearly its contribution to improved performance.

1.3.2 The ALNAP Evaluative Reports Database (ERD)

From the outset, those involved in the formation of ALNAP (in 1997) recognised the actual and potential role of evaluation, as well as the need for a mechanism that would enable the systematic collection and sharing of evaluative reports to facilitate and improve interagency and collective learning. The ERD[5] has been a central ALNAP activity ever since, and the Annual Review takes the initiative one step further. Using the ERD as its primary source the Review provides a baseline analysis against which to track and report on future trends. (see Box 1.3).

For the full accountability and learning potential of evaluation to be realised, a more systematic exploitation of the mechanism will be required.

Box 1.3 The Evaluative Reports Database: Trends

Analysis of the 130 reports categorised as Evaluations of Humanitarian Action in the ERD:

Commissioning Organisations

UN or UN
Organisation
16%

Red Cross or Red
Crescent Group
9%

ECHO
19%

NGO or NGO
Umbrella Group
19%

Bilateral donors
37%

Geographical Coverage

Europe
21%

Latin America
and Caribbean
11%

Asia
18%

Africa
50%

Type of Emergency

Both
11%

Natural
Disaster
26%

Complex
Emergency
63%
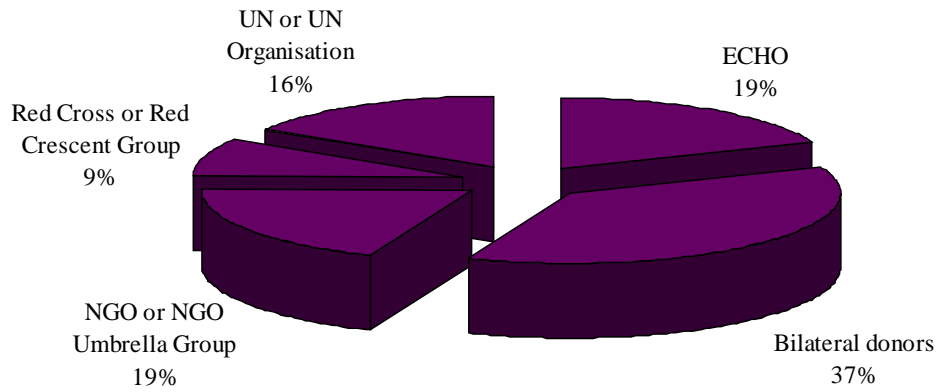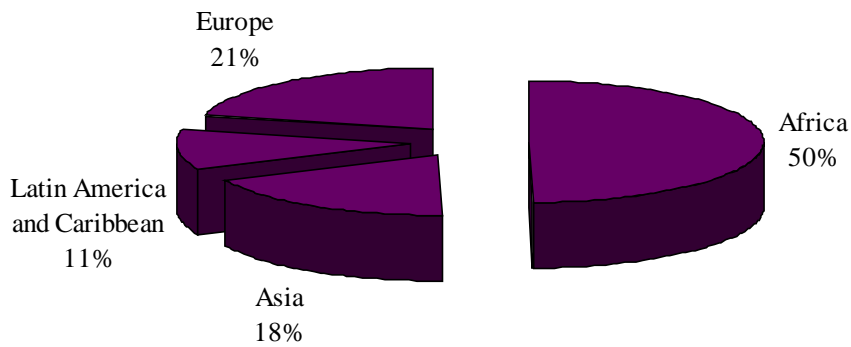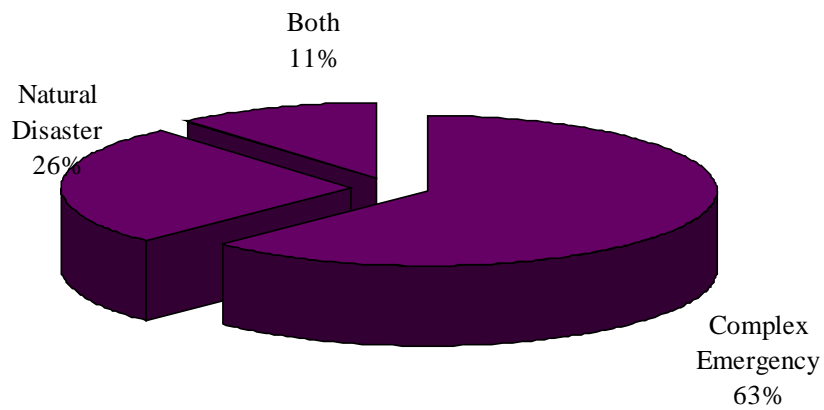
1.4 Defining 'Evaluation of Humanitarian Action'

Defining the key characteristics of 'evaluation of humanitarian action' was of particular importance in determining the reports to be considered in the Annual Review. Using DAC's definition as a starting point, ALNAP identified key characteristics, resulting in the adoption of the following in early 2001:

'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:

- it is commissioned by or in cooperation with the organisation(s) whose performance is being evaluated;
- 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;
- 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 (see OECD-DAC, 1999 for an explanation of these criteria).
- It articulates findings, draws conclusions and makes recommendations.'

Other than in Chapter 3, which also draws on other types of evaluative report looking at Kosovo, the Annual Review considers only those reports that conform to the above definition. (The process of categorising the reports is undertaken only on the basis of information contained in, or that can be inferred from, a report.)

While adherence to the definition has the considerable benefit of providing a coherent set of reports, it inevitably excludes. Such exclusion should not be read as a negative judgement on the quality or value of other types of evaluative report but simply as an acknowledgement of their differences. As the approach taken in Chapter 3 indicates, other types of evaluative report can make a valuable contribution to accountability and lesson learning. The process of comparing and assessing the value and contribution of evaluation and other types of evaluative reports also helps to sharpen analysis of the strengths and weaknesses of each.

1.4.1 Particularities of evaluation of humanitarian action

The evaluation of humanitarian action should be viewed as a particular and recent form of evaluation despite drawing significantly on the techniques and approaches developed in the 1970s for the evaluation of international development co-operation. Both provide reports intended to feed into the policy process and improve future performance, both are centrally concerned with assessment of impact, appropriateness, effectiveness etc, and both are generally undertaken by external consultants or mixed teams. However, the major contextual differences between the humanitarian and development arenas have required considerable adaptation of the development genre and the adoption of particular criteria, techniques and approaches for the former.

The principal contextual differences are indicated below.

- Humanitarian programmes frequently operate under considerable time pressure and in very fluid and fast-changing contexts. Consequently, project cycles are compressed so that initial planning and baseline data gathering, to enable 'before' and 'after' comparisons, are so perfunctory (if done at all) as to be of little use in subsequent evaluations. This lack of baseline data, and in some cases basic project information, requires evaluators of humanitarian action to reconstruct 'histories' and 'pictures'.

- Humanitarian programmes often operate in areas where the police, judiciary and press are overwhelmed, ineffective or no longer functioning. This increases both the pressure on, and the significance of, the evaluations' accountability role.
- Humanitarian programmes, with their focus on saving lives in extreme situations, are often accorded a high profile by the media and politicians. The resulting increase in sensitivity, to direct or implied criticism of programmes under evaluation, impacts both on the process and how results are placed in the public domain.
- Humanitarian operations invariably involve a number of different of agencies, in contrast to the 'dominant and clear leader' usually found in development contexts. This multiplicity is often characterised by a high degree of interdependency, where one agency's performance relies significantly on the performance of others – e.g. for funding or delivery in the delivery chain, or the provision of complementary services. This makes trying to attribute relative roles and evaluate shortcomings and impacts more difficult.
- The overlap in the criteria used by the development co-operation and humanitarian genres of evaluation is acknowledged, but some are less applicable than others. For instance, sustainability[6] is often not appropriate for interventions with short-term objectives. Criteria such as coherence and coordination are of central importance in evaluating humanitarian action, given the need for policy coherence in respect of security, diplomatic, humanitarian and development actions.
- Humanitarian intervention is often highly multidisciplinary – e.g., public health, logistics, food and nutrition, water and sanitation, etc. Evaluation teams are required to reflect such diversity in their composition and expertise.
- Conflicts polarise perspectives so that the same events are often subject to widely differing interpretations, diminishing the space for objective assessment.
- The experience of conflict or a natural disaster may traumatise individuals who would have acted as information sources for the evaluation.
- Finally, teams undertaking evaluations of humanitarian action often experience the same problems (insecurity, logistical difficulties, increased risk of illness, etc.) as the agency personnel whose activities they are assessing.

1.5   Evaluation of Humanitarian Action: Current Issues

It is useful at this stage to introduce four issues currently facing the system. These have emerged from experiences and discussion within ALNAP and an analysis of the ERD:

- the extent to which evaluation is now ingrained in the humanitarian system;
- the balance between the dual aspirations of evaluation – accountability and lesson learning;
- the gaps in existing guidance on the evaluation of humanitarian action;
- the assessment of quality in relation to evaluation reports.

1.5.1   Is evaluation ingrained in the humanitarian system?

Statistics drawn from the ERD (see Box 1.3) need to be interpreted with caution. While bilateral and multilateral donors commissioned 54% of the total number of evaluations held, the focus of these is usually on the performance of NGOs and to a lesser extent that of UN and Red Cross agencies, rather than on their own performance.

Despite the need for caution, evidence suggests that not all parts of the international humanitarian system are contributing to the same degree to the evaluation of humanitarian action 'boom'. Bilateral and multilateral donor organisations appear significantly better at commissioning and sharing their evaluations than other types of organisation. The UN group accounts for just 17% of the total. However, the UN is able to point to the declassification, in 2000, of all UNHCR's evaluation reports as a major achievement.

NGOs were responsible for commissioning 20% of the evaluation reports available on the ERD, however no less than 54% of these reports were commissioned by just two organisations (MSF-Holland and the UK Disasters Emergency Committee [DEC]). The available data does not show conclusively whether these results reflect a reluctance to share evaluation reports, a general lack of commitment to the evaluation process, or a tendency to use internal personnel for evaluative reviews, thus falling outside the ALNAP definition.

The dependence of most NGOs on private funding sources, particularly susceptible to influence by adverse publicity, is probably a significant factor contributing to NGOs' possible caution and sensitivity in the commissioning and/or sharing of evaluations. It is not however in the interests of the international humanitarian system that a group of agencies that plays such a significant role in the delivery of assistance to affected populations does not engage more directly and openly with such an important vehicle for accountability and learning. Greater encouragement may need to be given to NGOs to undertake evaluations of humanitarian action and share the reports not only with other agencies, but more importantly with their supporters and the wider donating public.

While the extent to which the evaluation mechanism has become part of the culture of the international humanitarian system is important in itself, it remains of limited significance if evaluation findings, conclusions and recommendations are not being systematically acted on. The results of work commissioned by ALNAP on the follow-up to evaluation will feature in some detail in the second *Annual Review*, to be published in April 2002.

1.5.2 Balance between accountability and lesson-learning

The tension between the principal dual objectives of evaluation of humanitarian action, accountability and lesson-learning, has long been recognised. For the most part, it is regarded as a creative tension that contributes to the uniqueness and potential value of the evaluation process.

The view held by this author is that lessons are more likely to be learned if the process is open to external scrutiny with organisations held accountable for their actions (or inactions). However, the presence of the accountability objective, enabling 'outsiders' to assess the performance of organisations and their staff (who tend to see themselves as doing their best in adversity), is also acknowledged as a major factor in resistance to evaluation of humanitarian action.

The relative balance accorded these two objectives influences the evaluation approach adopted. An evaluation placing emphasis on accountability tends to be perceived as a threat to those being evaluated. This often results in a process with less engagement between the evaluators and personnel from the agency being evaluated than an evaluation emphasising learning. The latter tends to take, and arguably requires, a more participatory approach, engaging both programme staff and its intended beneficiaries. The choice of emphasis, however, remains with the commissioning organisation and, though some will have strong accountability requirements imposed on them by their statutes and governing bodies, others' accountability requirements will be lower, perhaps by virtue of their funding structures or weak relations with funders. Many NGOs fall into this category, allowing them greater leeway to emphasise the lesson-learning objective in their approach to evaluation.

Given the intensity of humanitarian action over the last decade and the need for the humanitarian system to undertake fundamental changes in response to the sharper critiques, some may be tempted to argue that evaluations of humanitarian action should place greater emphasis on the lesson-learning objective. However, such a general recommendation would not be appropriate, as different organisations and contexts will require different approaches and emphases. Without the

encouragement and degree of pressure afforded by the accountability objective, an agency's commitment to learning the lessons drawn, and integrating them into its policies and practices, may well be insufficient to drive the necessary changes and realise the goal of improved performance.

While there are many genuinely new issues (e.g., protection and security) a striking feature of the analysis in the following chapters is the extent to which the old issues recur. Among them poor interagency coordination (including that between donor organisations); poor linkages between the humanitarian and development system; and, the lack of beneficiary participation in humanitarian action and its evaluation. Perhaps the lesson-learning objective should be given emphasis only in those areas where there are genuinely new lessons to learn whilst accountability should be given emphasis in other areas where lessons are evidently not being learnt, or not being translated into improved practice.

Whatever the chosen emphasis, a fundamental and critical factor in good evaluation practice is that prioritisation of objectives should be shared by key stakeholders and made explicit in the terms of reference and throughout the evaluation process.

The ALNAP book *Evaluating International Humanitarian Action: Perspectives from Practitioners* (Wood et al., 2001) concludes that lack of clarity on the purpose and focus of evaluations of humanitarian action has, together with inadequate time and resourcing, been a significant factor limiting the effectiveness and value of evaluations:

'Ensuring that the evaluation team knows what is wanted and for what purpose, and forcing the commissioning agency, and not just its evaluation unit, to think about this and clarify the usually excessive and even incoherent TOR is a critical step to achieving more effective and useable evaluations.'

1.5.3   Gaps in guidance

Developments in thinking and practice around evaluation of humanitarian action are so rapid that what might be regarded as the 'first generation' of guidance (Hallam, 1998; OECD-DAC, 1999) is already considered to be in need of updating.

ALNAP has identified guidance gaps in a range of areas and prioritised the development of practical 'how to evaluate' guidance on the following:

- Protection
- Policy
- Security
- Human resources
- Information & knowledge management
- Process (how does one evaluate against recognised criteria such as 'coherence', 'connectedness' and 'coverage'?)

Progress will be reported in the second Annual Review to be published in April 2002.

1.5.4   Assessing the quality of evaluation reports

To some extent, the credibility, potential and effectiveness of evaluation as a tool for change, as well as the follow-up to evaluation, is being undermined by the variable quality of both the undertaking and its product. A pressing requirement therefore is to reach general agreement on what constitutes good practice in the evaluation of humanitarian action and to develop quality benchmarks for use by evaluation managers and evaluators.

In October 2000 ALNAP full members triggered a process to identify and agree appropriate criteria for assessing the quality of evaluation reports of humanitarian action. A preliminary 'quality' proforma[8] (see Annex 1) was developed drawing on existing guidance, including the OECD-DAC guidance, and a growing body of what is increasingly acknowledged as good practice in evaluation. It uses the following 11 headings:

- Purpose and focus of the evaluation
- Constraints experienced
- TOR, team composition and time allowed
- Information on the context and the intervention
- Methodology and transparency
- Consultation with beneficiaries and the affected population
- Reference to international standards
- Attention to gender and vulnerable or marginalised groups
- Coverage of factors potentially influencing performance
- Conclusions and recommendations
- Legibility and dissemination of the final report

The first two and their sub-headings allow key contextual information to be established, and the nine other heading areas contain a series of criteria against which quality assessments of the evaluation reports considered in this Annual Review have been undertaken. Assessment findings and analysis thereof appear in aggregate form in Chapter 2, in respect of the non-Kosovo evaluation reports, and in Chapter 4 in respect of all the evaluation reports considered in this Annual Review.

Having set the background, we can move on to consider the synthesis and meta-evaluation of the evaluations of humanitarian action reports published during 1999–2000 for the non-Kosovo and Kosovo sets.

Endnotes

1. As well as the Annual Review series, ALNAP's 2001–02 Workplan activities in relation to the evaluation of humanitarian action include: the publication of a volume of reflections by evaluation practitioners on nine humanitarian action evaluations undertaken between 1993 and 2000 (Wood et al, 2001); the development of short courses and training modules for agency personnel and evaluators; studies on how evaluations of humanitarian action are used and followed up by agencies; and the sharing of evaluation plans between members to reduce duplication and encourage greater collaboration in evaluation. Other accountability and learning activities include a book on accountability frameworks in the international humanitarian system; a major international collaborative study to develop guidance for humanitarian agencies on consultation with and participation of beneficiaries and the affected population; and testing the field level 'Learning Support Office' (see Endnote 5, Chapter 3).
2. The ALNAP symposium 'Learning from Evaluation: Humanitarian Assistance and Protection in Kosovo' was held in Geneva on 17 & 18 October 2000. Some 51 representatives (military/Red Cross movement /NGO/bilateral and multilateral donor organisations/UN plus academics and consultants) from 43 different organisations that had been involved in the Kosovo crisis or the evaluation of humanitarian programmes, participated in the event.
3. Operation Provide Comfort was mounted in the aftermath of the Gulf War between Coalition forces and the Iraqi forces that had occupied Kuwait. Following the March 1991 ceasefire between the Coalition forces and the Iraqi government, a popular uprising took place in the predominantly Kurdish areas of northern and eastern Iraq. The Iraqi military's brutal suppression of the uprising led to the exodus of almost 2 million Kurds towards Iran and Turkey. The Turkish authorities refused them admission and almost 500,000 Kurds were trapped in the mountains along the Iraqi/Turkish border. Following Security Council Resolution 688 of 5 April 1991, US and European airforces began airdropping supplies and US, British and French forces created a 'safe haven' in northern Iraq, enabling the provision of direct

humanitarian assistance by the militaries, the UN, the Red Cross and NGOs and the movement of the Kurds down from the mountains back into Iraq. The operation was a watershed event for the humanitarian system as it set a precedent for military intervention in support of humanitarian objectives on sovereign territory that was subsequently repeated in Somalia, Kosovo and East Timor.

4. ICRC, IFRC, UNICEF, UNHCR and WFP are the five organisations involved in the benchmarking study.

5. The DAC Working Party on Aid Evaluation maintains a database of evaluations of development cooperation undertaken by bilateral and multilateral donor organisations <www.oecd.org/dac/evaluation>. Whilst the database is extremely valuable, evaluations undertaken by organisations other than DAC members, in the international humanitarian sector are not included and the number of evaluations of humanitarian action is limited.

6. See Chapter 2 Endnote 4.

7. The book is scheduled for publication by Zed Press in August 2001.

8. The proforma has been given a blind pre-test by two independent consultants to assess consistency of application across different assessors and help guide the proforma's further development/refinement. The objective is to strengthen its methodological basis to ensure its acceptance.