

## **ALNAP**

### Active Learning Network for Accountability and Performance in Humanitarian Action

ALNAP is an international interagency forum working to improve learning and accountability across the humanitarian system. Established in 1997, it has some 50 Full Member Organisations and 280 Observers. It encompasses all the types of organisation that make up the international humanitarian system Bilateral and Multilateral donors: UN Agencies and Departments; NGOs and NGO umbrella organisations; the International Red Cross and Red Crescent Movement; and independent consultants, academics and research institutes

To encourage self-criticism and learning, ALNAP ensures that no one type of organisation is able to dominate discussions or set the agenda.

ALNAP is governed by a Steering Committee of 8 Full Member Representatives selected to maintain the balance of representation. It is serviced by a small Secretariat located at the Overseas Development Institute, in London.

The Annual Review Series was conceived as part of ALNAP's 2000-02 Workplan constructed around the following 3 themes:

- I Making the Evaluation Process More Effective;
- II Strengthening Accountability Frameworks;
- III Improving Field-level Learning Mechanisms.

The Annual Review draws extensively on ALNAP's Evaluative Reports Database (ERD).

Workplan activities are funded voluntarily by Full Members organisations.

#### **ALNAP Full Members:**

Prof. Apthorpe, AusAID, BADC, BRCS, CAFOD/Caritas, CARE International, CIDA, DANIDA, DFA Ireland, DFID, DMI, ECHO, Emma Ltd., ETC UK, FAO, Groupe URD, HAP, ICVA, ICRC, IFRC, InterAction, JICA, MFA France, MFA Germany, MFA Netherlands, MFA Norway, MSF-H, OXFAM, ODI, PIA, SCF UK, SCHR, SDC, Prof. Shaw, SIDA, SPHERE, Tufts University, UNDP, UNHCR, UNHCHR, UNICEF, UNOCHA, Uppsala University, USAID/OFDA, VOICE, WFP, WHO, World Bank, World Vision.

**ALNAP Annual Review Series 2001**

**Humanitarian Action:**  
Learning from Evaluation

**ALNAP**

Active Learning Network for Accountability and Performance  
in Humanitarian Action

## **Editors**

Rachel Houghton with Kate Robertson

## **Contributors**

**Chapter 1** John Borton with Lucy Carver

**Chapter 2** Tony Beck

**Chapter 3** Based on a paper by Raymond Apthorpe

**Chapter 4** Tony Beck with John Borton

**About Contributors:** Prof. Raymond Apthorpe, Visiting Professor at Australia National University, Canberra; Dr Tony Beck, Evaluation and Poverty Specialist, Vancouver; John Borton and Lucy Carver, ALNAP Secretariat, London.

**ISBN** 0-85003-516-3

© Overseas Development Institute, London, 2001

Photocopies of all or part of this publication may be made providing that the source is acknowledged. ALNAP would appreciate receiving details of any use of this material in training, research or programme design, implementation or evaluation. Requests for commercial reproduction of ALNAP material should be directed to ALNAP.

**For further copies** of this publication please contact:

**ALNAP**

c/o Overseas Development Institute

111 Westminster Bridge Road

London SE1 7JD

United Kingdom

Tel: +44 (0)20 7922 0300

Fax: +44 (0)20 7922 0399

Email: [alnap@odi.org.uk](mailto:alnap@odi.org.uk)

Website: [www.alnap.org](http://www.alnap.org)

**Price per copy** £15 (excluding postage and packaging)

## **Foreword**

Twelve months ago the Full Members of ALNAP approved the launch of the Annual Review Series. They did so in the belief that an annual summary of evaluation findings, coupled with an assessment of the overall quality of evaluation reports, would provide significant added value to their work and efforts to improve the international humanitarian system's performance. This first Annual Review confirms that belief.

It draws together the principal findings of over 49 evaluations of humanitarian action published during 1999 and 2000. The findings highlight areas where the performance of the humanitarian system was judged by evaluations to have been satisfactory and where poor. The fact that findings are drawn from a large sample of reports gives them considerable force, and they deserve a role in the reflections and dialogues that set the agenda for improving the system's performance.

Synthesising the findings of so many separate evaluations inevitably raises questions about the standards used for assessing performance, and the quality of the evaluation reports themselves. The Annual Review, therefore, combines synthesis of the evaluations' main findings with a meta-evaluation of the quality of reports, showing us where evaluation practice deserves emulation and where it is in need of strengthening. By applying a quality proforma developed by ALNAP, the Annual Review has established a precedent. It is not the last word on what constitutes quality, and there will be further debate. However, in providing a systematic basis for assessment, it promises to be an important tool for monitoring the progress of efforts to improve the quality and effectiveness of the evaluation of humanitarian action mechanism.

I firmly believe that this annual opportunity for the humanitarian system to hold a mirror to itself and reflect on its performance, and, for those involved in the commissioning, undertaking and use of evaluation to reflect on practice and quality, will prove invaluable. This is a positive addition to the system's growing repertoire of instruments for measuring performance, enabling it to feel proud of its successes and admit to the continuing challenge to improve.

**Wayne Macdonald**

**Chairman of ALNAP**

**(Head of Planning, Monitoring and Evaluation, ICRC)**

**Geneva, April 2001**

## CONTENTS

|                |   |            |
|----------------|---|------------|
|                | <b>Evaluation Reports Database</b>  | <b>6</b>   |
| Chapter 1      | <b>Introduction</b>   | <b>7</b>   |
| 1.1            | Purpose and Scope of the ALNAP Annual Review Series   | 9          |
| 1.2            | A Decade of Rapid Change  | 12         |
| 1.3            | Evaluation: a Key Tool for Accountability and Learning  | 15         |
| 1.4            | Defining 'Evaluation of Humanitarian Action'  | 18         |
| 1.5            | Evaluation of Humanitarian Action: Current Issues   | 20         |
| Chapter 2      | <b>Summary of Findings and Assessment of Non-Kosovo Evaluation Reports</b>                                  | <b>27</b>  |
| 2.1            | Introduction  | 29         |
| 2.2            | The Sample Assessed   | 29         |
| 2.3            | Synthesis of Main Findings  | 31         |
| 2.4            | Meta-Evaluation of Non-Kosovo Reports   | 52         |
| Chapter 3      | <b>Evaluations of Humanitarian Aid in Response to the 1999 Kosovo Crisis: Synthesis and Meta-Evaluation</b> | <b>65</b>  |
| 3.1            | Overview of this Chapter  | 67         |
| 3.2            | The Kosovo Conflict: a Set of Interlocking Crises   | 72         |
| 3.3            | Agency Response Issues  | 82         |
| 3.4            | Meta-Evaluation: Kosovo Evaluations as Learning Tools   | 97         |
| Chapter 4      | <b>Overview of the Analysis and Next Steps</b>  | <b>109</b> |
| 4.1            | Introduction  | 111        |
| 4.2            | Principal Findings and Issues Raised by the Reports   | 112        |
| 4.3            | The Quality of Evaluation Reports   | 125        |
| 4.4            | Implications of the LRRD Debate for Establishing Evaluation Standards                                       | 132        |
| 4.5            | Next Steps in Evaluation of Humanitarian Action   | 133        |
| 4.6            | Concluding Remarks  | 135        |
|                | <b>Endnotes</b>   | <b>137</b> |
| <b>Annex 1</b> | Quality Assessment Proforma   | 143        |
| <b>Annex 2</b> | Acronyms  | 147        |
| <b>Annex 3</b> | Glossary  | 151        |
| <b>Annex 4</b> | References for Source Evaluations   | 155        |
| <b>Annex 5</b> | References for Other Source Material  | 171        |

## **THE EVALUATIVE REPORTS DATABASE**

Annual Review Principal Data Source

ALNAP's Evaluative Reports Database (ERD) was set up in 1997 to facilitate access to evaluative reports of humanitarian action and improve interagency and collective learning.

It contains over 288 evaluative reports of humanitarian action classified under 8 categories 'Evaluations', 'Reviews', 'Lesson Learning Studies', 'Evaluation Syntheses and Meta-Evaluations', 'Conference/Seminar and Workshop Reports', 'Evaluation Follow-Up Studies', 'Audits' and 'Good Practice Reviews'.

At the time of publication, 130 of the reports were categorised under 'Evaluations', the key data source for ALNAP's Annual Reviews

Although an expanding collection, the majority of the reports held are provided by ALNAP Full Members and relate directly to humanitarian actions they have either funded or implemented. It is hoped that the Annual Reviews will raise awareness of the ERD resulting in an increased availability of evaluative reports from non-ALNAP sources and a more comprehensive collection.

To encourage the sharing of reports that might otherwise be retained within commissioning organisations, reports may on request be classified 'ALNAP Full Members only' and password protected within the ALNAP website. If you are aware of an evaluative report of humanitarian action not included, please notify the Database Manager <[alnap@odi.org.uk](mailto:alnap@odi.org.uk)>

Non-classified reports are open to all, with key sections (e.g., contents; executive summary; principal findings and recommendations; methodology; Terms of Reference) fully searchable on the ALNAP website <[www.alnap.org](http://www.alnap.org)>.

An additional source is the 'Useful Texts' database. It contains over 100 references on key ALNAP areas of interest, such as Learning Theory and Practice, Accountability Theory and Practice, Evaluation Guidance, Evaluation Training Materials, Evaluation Theory and Practice and Humanitarian Codes and Standards. Abstracts and a table of contents will be accessible on the ALNAP website by mid-2001.

# **CHAPTER 1**

## **INTRODUCTION**

# 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<sup>1</sup>. 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 the results of 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, enabling organisations and individuals to draw learning from experience through a systematic and impartial scrutiny of performance

When considered and analysed *en masse* – through synthesis – the potential of evaluation of humanitarian action to facilitate system-wide scrutiny and learning becomes all the more



powerful. This first Annual Review synthesises 49 such evaluation reports and 5 existing syntheses, published between 1999–2000 and held on the ALNAP Evaluation 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 personnel, alike. 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.

Its author 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 fully exploit 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 between 1991-96 and 1996-2000, and Danida's humanitarian assistance between 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 well over 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<sup>2</sup>.

## 1.1.4 **Chapter 4: Conclusions**

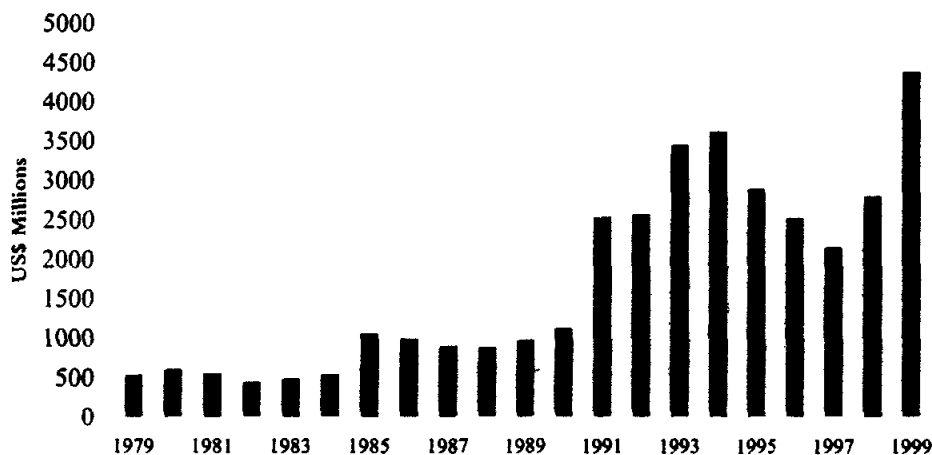
Chapter 4 draws together and reflects on the findings of the preceding chapters. In so doing, it outlines an agenda of issues to be addressed by the international humanitarian system and by those involved in the commissioning, undertaking and use of evaluations of humanitarian action. Based as it is on 49 evaluations published over a two-year period covering operations in 36 countries in four 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'<sup>3</sup>, 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

**Box 1 1 Principal Accountability Initiatives in the Humanitarian System over the Last Decade**

**1994** Publication of the Code of Conduct for the International Red Cross and Red Crescent Movement and NGOs in Disaster Relief. <[www.ifrc.org](http://www.ifrc.org)>

**1996** Publication of the Joint Evaluation of Emergency Assistance to Rwanda. <[www.um.dk](http://www.um.dk)>

Start of the Sphere Project by a coalition of European and US NGOs to develop minimum standards for five sectors of assistance which resulted in the publication of a final version of a Humanitarian Charter and Minimum Standards in Disaster Response in 2000.

<[www.sphereproject.org](http://www.sphereproject.org)>

Formation of People in Aid by a group of UK organisations to focus on issues in the management and support of aid personnel and which in 1997 produced the People in Aid Code of Best Practice.

<[www.peopleinaid.org.uk](http://www.peopleinaid.org.uk)>

**1997** Formation of the Active Learning Network for Accountability and Performance in Humanitarian Action (ALNAP) to provide an interagency forum to discuss and link these initiatives and undertake complementary activities on learning and accountability issues.

<[www.alnap.org](http://www.alnap.org)>

Group of UK organisations form to develop the concept of a Humanitarian Ombudsman (renamed in 2000 as the Humanitarian Accountability Project) as recommended by the Rwanda evaluation – in 2001 this initiative was raised to an international action-research plane.

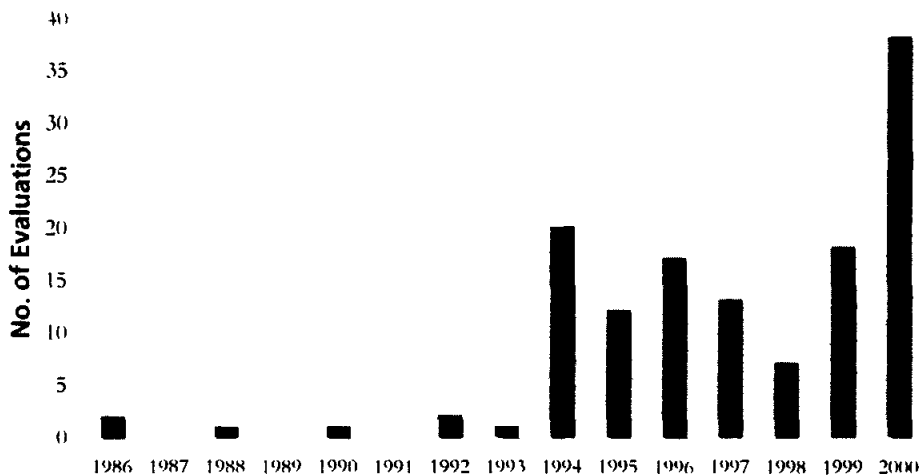
<[www.oneworld.org/ombudsman](http://www.oneworld.org/ombudsman)>

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

Fig 1.2 Evaluations in ERD by year of publication



organisations, UN agencies and non-governmental organisations (NGOs).

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 by outsiders looking in, but also 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 experiences 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 intelligence that greater accountability and transparency was 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 being tailored to reflect the particularities of those activities. The definition of evaluation of humanitarian action is introduced below.

Compared with the application of evaluation to development cooperation (see Box 1.2), its application to humanitarian action has been slower. This was partly the result of initial resistance but also due to 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

#### **Box 1 2 The Development of Evaluation**

As a field of professional practice, evaluation began in the United States of America during the 1960s when the 'Great Society' legislation poured Federal expenditures into poverty alleviation, housing, welfare and education programmes. Many of these programmes were required by legislation to undertake evaluations of their effectiveness to guide use of available funding. Evaluation then spread beyond the Federal programmes and was applied to other areas of US public expenditure and then adopted as a practice in other countries. As the use of evaluation grew and spread, so professional evaluation societies were formed with the first International Evaluation Conference held in 1995.

Evaluation was applied to international development assistance at a comparatively early stage. The 1950s and 60s de-colonisation process led to new departments and ministries for the administration of development cooperation programmes. Evaluation units were gradually introduced and by the end of the 1970s most aid administering organisations had one in one form or another. Under the auspices of the Organisation for Economic Cooperation and Development (OECD), representatives of these units formed a Group on Aid Effectiveness that subsequently evolved into the DAC Working Party on Aid Evaluation. It meets regularly to share experience to improve evaluation practice and strengthen its use as an instrument of development cooperation policy.

**<[www.oecd.org/dac/evaluation](http://www.oecd.org/dac/evaluation)>**

the evaluation of humanitarian action ‘took off’ (see Figure 1.2). Since then, however, a veritable ‘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 apparent ‘boom’ in the evaluation of humanitarian action 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<sup>4</sup> 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 clearly demonstrate 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<sup>5</sup> 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 p22)

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



## 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 broader definition as a starting point, ALNAP undertook a process to identify 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 normal accountability mechanisms such as 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 sensitivities, 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 multiplicity 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 creates considerable difficulties for a process trying to attribute

- relative roles and evaluate shortcomings and impacts.
- 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<sup>6</sup> is often not appropriate for interventions with short-term objectives. Given the need for policy coherence in respect of security, diplomatic, humanitarian and development actions, as well as the ‘many actors’ problem, criteria such as coherence and coordination are, or should be, of central importance to evaluations of humanitarian action.
  - 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**

Without wishing to anticipate or dilute the impact of points made in the following chapters, 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' category reports, a significant proportion of these tend to look 'out' at partner/ implementing organisation's performances rather than 'in' at their own performance – e.g., the provision of funds in a predictable and timely manner and at levels that enable partners to meet international standards in the provision of assistance and protection. Thus while bilateral and multilateral donor organisations commission a significant proportion of the total number of evaluations, their focus is usually on the performance of NGOs and to a lesser extent that of UN and Red Cross agencies.

Despite the need for caution, evidence suggests that not all parts of the international humanitarian system are contributing to the evaluation of humanitarian action 'boom' to the same degree. 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, it 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 or DEC). The available data does not allow judgement to be made as to 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.

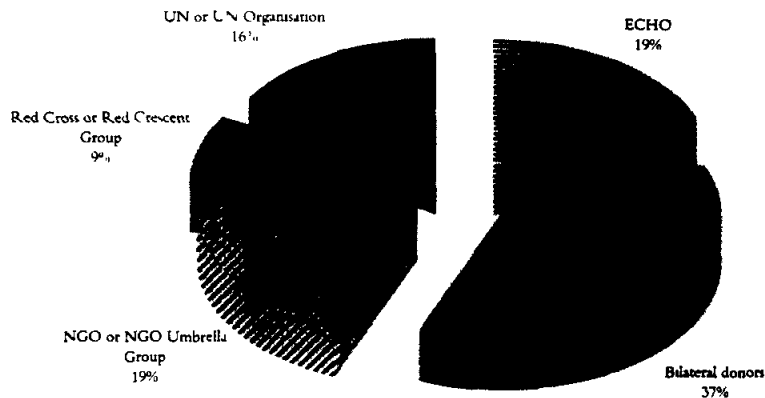
The dependence of most NGOs on private funding sources that are particularly susceptible to influence by adverse publicity is probably a significant factor contributing to their possible caution and sensitivity in the commissioning and/or sharing of evaluations. It is not however in the interests of 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 wider donating public.

While the extent to which the evaluation mechanism has become part of the culture of the international humanitarian system is

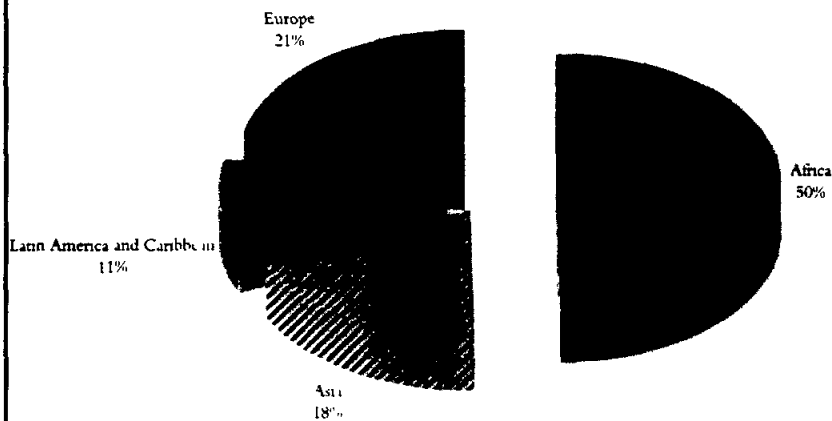
### Box 1 3 The Evaluative Reports Database: Trends

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

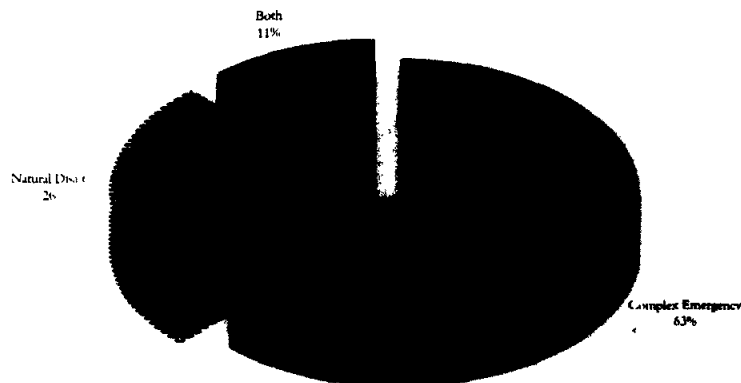
#### Commissioning Organisations.



#### Geographical Coverage.



#### Type of Emergency.



important in and of 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 at least are 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)<sup>7</sup> 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' (Wood et al., 2001)

### 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' (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.

## **CHAPTER 2**

### **SUMMARY OF FINDINGS AND ASSESSMENT OF NON-KOSOVO EVALUATION REPORTS**

# CHAPTER 2

## SUMMARY OF FINDINGS AND ASSESSMENT OF NON-KOSOVO EVALUATION REPORTS

### 2.1 Introduction

This chapter has two interlinked foci:

1. synthesis of the main findings, conclusions and recommendations of non-Kosovo 1999–2000 evaluation reports of humanitarian action;
2. meta-evaluation of non-Kosovo 1999–2000 evaluation reports.

The purpose of the synthesis is to provide an overview of the performance of the humanitarian system for the period under review. The topics covered mirror the main topics in the evaluation reports themselves, they also represent the main issues in humanitarian action as defined by OECD-DAC (1999) and the general evaluation and humanitarian literature. The meta-evaluation assesses the quality of the evaluation reports. The preliminary proforma was developed to assist in this objective. This attempts to measure the quality of reports against generally acknowledged good practice criteria. The proforma was introduced in Chapter 1 and further details are provided in the meta-evaluation section of this chapter.

### 2.2 The Sample Assessed

This chapter analyses the findings of 37 evaluation reports of humanitarian action completed during the 1999–2000 period that did not cover Kosovo. These constitute most of the major evaluation reports for the period. All the reports are classified as 'evaluations of humanitarian action' in the ALNAP Evaluative Reports Database. The set comprises 33 individual evaluation reports and four synthesis reports which themselves cover a further 237 evaluations<sup>1</sup>. Consequently this chapter represents one of the most comprehensive surveys of such evaluations to date.

The four synthesis reports are.

- DANIDA (1999a), which covers 7 evaluations (also assessed individually in this chapter).
- ECHO (1999), which covers 140 evaluations.
- ECHO (2000, 2000a), which covers 93 evaluations.
- USAID (2000a), which covers 4 evaluations.

The 33 individual evaluation reports are listed together in Annex 4, where they are broken down by salient characteristics. Of these, 14 relate to complex emergencies, 11 relate mainly to rehabilitation after conflict or complex emergencies, and nine relate to natural disasters. The total disbursement evaluated by the individual evaluation reports amounts to some US\$1.7bn.<sup>2</sup>

About half of the evaluation reports explicitly note that they have a joint lesson learning and accountability focus; also five have a mixed focus on lesson learning and planning. (All conform to the ALNAP definition of the evaluation of humanitarian action and are judged to have lesson learning and accountability objectives, even where not explicitly stated). The majority of the reports evaluate both relief and rehabilitation, and many cover the connection between the two. The set covers 30 countries and 4 regions, including Afghanistan, Central America, the Great Lakes and the Horn of Africa regions, Angola, Bangladesh, China, East Timor, India, North Korea, and the Caucasus. It also covers a wide range of interventions, the most common of which were related to nutrition, health, water and shelter (see Annex 4). The humanitarian actors assessed are for the most part the main UN agencies (WFP, UNHCR, ICRC, IFRC, IOM, OCHA and UNICEF) and the major INGOs and their in-country partners. Due to the wide variety of implementation channels, individual reports mainly cover both UN agencies and INGOs.

It is difficult to generalise when dealing with such a diverse set of evaluation reports and comparing across a range of emergencies and countries. However, a number of common areas were found and these are highlighted below. Findings are also compared to those of earlier synthesis studies. Unless otherwise noted, findings are consistent between the individual and synthesis reports.

All the evaluation reports held by ALNAP in English for the 1999–2000 period have been included, representativeness was not an issue in terms of selecting which ones to review. ECHO, as one of the largest agencies, is adequately represented by eight studies. However, the findings of this chapter should be read in