

Chapter 1

Introduction

Purpose and Scope of the ALNAP Annual Review

1.1

The *ALNAP Annual Review* series aims to advance understanding and practice in terms of how the quality of humanitarian action can be improved. It provides a snapshot of current trends in humanitarian action through a synthesis of evaluation findings, as well as an opportunity for more critical reflection on an area of particular concern – this year, the topic of monitoring is addressed. It also provides a platform for sharing lessons, identifying common approaches, and building consensus on directions for improving learning and accountability.

In addition to those working within the humanitarian sector, the *ALNAP Annual Review* series offers valuable insights to those involved in observing and commenting on the sector, for example, journalists, researchers, educationalists, and parliamentary and congressional committee members. Drawing on ALNAP's Evaluative Resources Database, the three years of the *Annual Review* have now covered 145 independent evaluation reports and 20 synthesis reports, perhaps the most exhaustive analysis of evaluation of humanitarian action (EHA) yet accomplished.

Box 1.1 Purpose and Scope of the *ALNAP Annual Review*

The *ALNAP Annual Review* has three main objectives:

- 1 To provide the humanitarian sector with the means to reflect annually on its performance, and identify generic strengths and weaknesses through a synthesis of the principal findings, conclusions and recommendations of EHA made available to ALNAP during the preceding year.
- 2 To address each year a central theme of common and current concern to those within the sector.
- 3 To monitor and assess the quality of EHA by highlighting good and poor practice through a meta-evaluation of evaluations received the preceding year – a key learning and accountability tool for the humanitarian sector.

The *ALNAP Annual Review* series complements other annual publications focusing on the humanitarian sector, such as the *World Disasters Report* (IFRC) and the *World Vulnerability Report* (UNDP).

The Emergence of ALNAP and the Accountability Agenda

1.2

As many readers will know, ALNAP emerged from what was perceived as a crisis of accountability in humanitarian action in the mid-1990s. Its mandate was to provide a sectorwide forum, owned by all and dominated by none, as a means of discussing concerns about learning, accountability and quality. At the time ALNAP found itself as part of an emerging group of initiatives that intended to place accountability firmly on the humanitarian agenda (see Table 1.1). However, while a consensus emerged that ‘something needed to be done’ – as humanitarian agencies realised the need to demonstrate they were serious about acting in an accountable manner to beneficiaries, donors and other stakeholders – it was less clear how to do this.

Table 1.1 Accountability Initiatives in the Humanitarian Sector

Project Name How to find out more	Project Objective	Participants
The Red Cross/ Crescent Code of Conduct (1994) www.ifrc.org	A voluntary code seeking to guard standards of behaviour in humanitarian action.	207 agency signatories
People in Aid (1996) www.peopleinaid.org.uk	To improve support and management of field staff and volunteers through adherence to a code of best practice.	11 NGOs/networks
The Sphere Project (1997) www.sphereproject.org	To develop minimum technical standards and a humanitarian charter for disaster response.	A coalition of European and US NGOs
Humanitarian Accountability Project (HAP) (2000) www.hapgeneva.org	To find a means of strengthening accountability towards those directly affected by disasters.	13 humanitarian organisations and networks

Adapted from Hilhorst (2002).

ALNAP initially found a niche that stemmed directly from the discussions in the mid-1990s about how to strengthen evaluation practice. The Joint Evaluation of Emergency Assistance to Rwanda (JEEAR) had already demonstrated how evaluation could have a radical affect on the sector, and later the Organisation for Economic Development Assistance (OECD) and the Development Assistance Committee (DAC) Working Group on Evaluation looked at ways to more effectively institutionalise EHA (OECD-DAC, 2001). Since then, ALNAP has been involved in the search for ways in which to combine accountability and learning in evaluation design and practice. This is one of the themes of this *Annual Review*. In these early years more managerial interventions, such as monitoring, received less attention, as it was assumed that an understanding of the problems of past operations would provide the basis and impetus for systemic reform.

Learning and Monitoring and Evaluation 1.3

The assumption that systemic reform would occur did not come to pass, and over the years it has been found that evaluation needs to be complemented by other approaches to raise accountability and learning during humanitarian operations. Monitoring is gradually being recognised as a key activity in achieving the joint objectives of learning and accountability, and it is partly as a result of this process, in addition to reflection on the findings from the two previous *Annual Reviews*, that ALNAP decided to focus on monitoring this year. Indeed, the topic of monitoring is in many respects a natural progression of the dialogue that ALNAP has developed in the humanitarian community over the past three years.

The first *Annual Review* in 2001 traced the development of humanitarian evaluation and demonstrated that evaluation has become an integral part of how the sector operates. In addition to documenting the considerable achievements that had occurred in the past 10 years, it also analysed the weaknesses. On the plus side, evaluations were found to expose organisations, teams and individuals to critical appraisal and provided a means of assessing when and where problems were addressed. However, there was a lack of clarity as to whether the objective of evaluation was to provide a practical tool for supporting organisational learning or whether EHA was merely intended to create greater upward accountability to

donors. There was a tendency to attempt to combine the two objectives, which in many instances led to lack of attention to the learning agenda (ALNAP, 2001).

Moreover, in 2001 ALNAP commissioned a review of how agencies followed up on the recommendations made in humanitarian evaluations (van de Putte, 2001). The findings showed that recommendations were rarely linked to learning processes and that practical follow-up to evaluations – especially when evaluation was designed as a stand-alone activity – was poor. This of course is not unique to EHA but is endemic in the evaluation field as a whole (Patton, 1997). Too many evaluations are left to collect dust on agency shelves. The need to better integrate evaluation into overall project management cycles and internal policy debates could not have been clearer. This, in turn, required backing from senior management and integration of the planning of evaluation into ongoing management frameworks, such as monitoring.

Last year's *Annual Review* put this in a broader context by providing an analysis of some of the key constraints to learning in the humanitarian sector. This focus on learning was in line with the movement over the last few years from a focus on evaluation methods to a focus on how knowledge produced during evaluation processes can be used to effect change (Patton, 2001). Some worrying observations were made as to the limited impact that evaluation findings and recommendations were having on field practice itself; this corroborated findings from the follow-up study. In other words, why invest so much in evaluation if lessons from past experience are not being learned? One of the main conclusions from last year's *Review* was that, in spite of some improvements in the quality of evaluations, there had been little impact on learning and a corresponding absence of any significant change in humanitarian action (ALNAP, 2002).

Constraints identified in the humanitarian sector included the following:

- Incentive structures in agencies that promote defensive behaviour and a culture of blame.
- Short-term funding mechanisms that militate against a learning environment for field staff.
- Very high rates of staff turnover within ongoing programmes and between programmes.
- Lack of clarity as to intervention objectives and desired outcomes.
- Training provision not properly linked to learning processes.
- Poorly developed mechanisms for cross-organisational learning.

As a result of this analysis an agenda for change and action was developed and presented in the concluding chapter of *Annual Review 2002*. The importance of this agenda has not diminished. Key constraints facing this reform will require much concerted work over a considerable period of time. In this and next year's *Annual Reviews*, ALNAP will take a closer look at the prevailing structures that impede agencies' abilities to learn and will explore how different tools can be used to address these dysfunctions. Specifically, it will ask whether monitoring can become more than a data collection exercise (as is often the case) to instead become a valuable vehicle for 'learning while doing'.

Monitoring and Learning in Real-time

1.3.1

The experiences of the past few years both within the humanitarian sector and the wider evaluation field have thus contributed to a desire in ALNAP and the humanitarian community to look beyond evaluation to see how other approaches may contribute to learning. Partly as a consequence we are now seeing an increase in new kinds of activities such as Real-time Evaluations (RTE), strategic review and self-evaluation (see Chapter 2). These aim to make up for deficiencies in traditional ex-post evaluation and are intended to provide a timely, rapid review of a particular response so that findings can be used to feed into ongoing decision making processes.

As demonstrated in Chapter 2, UNHCR started the RTE trend, closely followed by WFP, ALNAP's Learning Support Office (Malawi LSO) and the Humanitarian Accountability Project (HAP). The UK Disaster Emergency Committee (DEC) has also carried out some 'monitoring missions' that closely resemble RTE.

Alongside RTE other 'real-time' monitoring practices have increasingly been acknowledged as potentially important though often neglected topics. This is true for both of the main types of monitoring identified in Chapter 2 – situation/context monitoring, and performance/impact monitoring. The renewed interest in these activities is motivated by the concern that ex-post evaluations often come too late to affect the operations they assess and, given the weak institutional memories in many organisations, too early to influence the next operation. Ex-post evaluations will in some cases support learning, in particular where there are mechanisms for the integration of evaluation findings. However, the structure of many agencies militates against such integration: in the UN system and donors, for example,

evaluation offices are often kept separate from programming branches and report directly to the Executive Director's Office – in order to ensure 'independence'. This practice may, ironically, ensure lack of integration of findings into the programming branches.

Monitoring, on the other hand, is an ongoing agency function that is generally integrated into everyday programming. And as opposed to ex-post evaluation, monitoring may be able to offer a 'short action learning loop' (ALNAP, 2002). The importance of this cannot be underestimated given that information from monitoring can potentially enable mid-course corrections of programmes. It could also provide information for humanitarian field workers which could potentially empower them to make better judgements during the course of their immediate work. In this respect one area where monitoring has a distinct advantage is its potential to examine social process, such as why interventions are or are not working, who is benefiting, and why.

But monitoring will require more attention if this is going to happen. Last year's *Annual Review* indicated that the potential benefit of monitoring for learning is not being fully exploited, and conclusions from analysis of the 165 evaluation reports over the three years of the *Annual Review* strongly support the notion that the current quality of monitoring is poor. For these reasons it was decided that the themed chapter in this year's *Review* would directly address the question of whether robust monitoring systems are really able to make up for the deficiencies identified in the two previous *Reviews*. Can monitoring 'fill the gap' by playing a crucial role in both broad sectoral learning and in helping field workers make critical judgements in the midst of humanitarian crises?

Learning and Downward Accountability: Is Monitoring a Missing Link? 1.4

While evaluation remains an important tool for upward accountability – to donors, headquarters and auditors – it is clear that there are still prevailing learning and downward accountability gaps, the latter especially to the primary stakeholders. Edwards and Hulme (1995) comment that accountability is generally interpreted as

‘the means by which ... organisations report (upwards) to a recognised authority ... and are held responsible for their actions, with insufficient attention directed “downwards” to the views of the intended beneficiaries.’¹ Yet downward accountability has both a practical and an ethical dimension. Practically, improved downward accountability will support closer consultation and participation of affected people in the design and implementation of interventions. This means interventions are more likely to reflect genuine needs and priorities, and achieve optimum impact. Ethically, downward accountability is embedded in the values and principles central to humanitarian action, notably the Red Cross/Crescent Code of Conduct and the Humanitarian Charter. Thus a commitment to downward accountability is part of the living value system that underpins humanitarian action itself.

Initial findings from the ALNAP-commissioned Global Study on the Consultation with and Participation by Affected Populations in Humanitarian Action, as well as the findings of the *Annual Reviews*, suggests that there is a dearth of good practice in involving the affected population in the delivery of humanitarian aid (ALNAP 2003: Draft Practitioner’s Handbook and Global Study Monographs). Similarly, findings from the three HAP field trials in Sierra Leone, Afghanistan and Cambodia describe the challenge of finding and institutionalising adequate methods of effectively listening and responding to the needs and concerns of affected populations during the course of humanitarian operations (HAP, 2003).

Moreover, based on the content of recent reports from the ALNAP Evaluative Reports Database, it would appear that a significant proportion of field workers in emergencies are still not aware of the existence of the Red Cross/Crescent NGO Code of Conduct – which contains a commitment to downward accountability – even though it has been in existence for nearly 10 years (see RRN, 1994). The synthesis of evaluations in Chapter 3 reinforces this conclusion, noting: ‘In terms of consultation with and participation of primary stakeholders, this year’s reports echo the disturbing story of the past two years: the limited ability of agencies to promote participation beyond implementation activities.’ It therefore appears that good ideas and intentions about accountability to affected populations have yet to result in widespread good practice. This in turn is mirrored by poor evaluation practice, as evidenced in the cumulative three year findings of the *Annual Reviews* that 86 per cent of reports were rated as unsatisfactory in terms of consultation with, and participation of, primary stakeholders in the evaluation process. As the meta-

evaluation section notes ‘Despite some good practice, EHA could rightfully be accused of systematically ignoring the views and perspectives of primary stakeholders ... This undermines its credibility and continues in the vein of treating primary stakeholders as passive recipients of aid rather than active participants in their own recovery.’

A fundamental issue in terms of changing mind-sets within ‘traditional’ humanitarian action relates to the extent to which those involved in monitoring, strategic review and RTE are willing to enter into dialogue with primary stakeholders. Usually rushed evaluation exercises allow limited time for such dialogue, while ongoing monitoring over a period of several months offers the opportunity for longer term contact, with repeat visits to the same site, household or individual which should lead to both improved situation analysis and understanding of impact.

A key question is thus whether monitoring systems can be designed to support downward accountability given agency structures that are usually hierarchical in addition to the current extractive function of monitoring with information normally flowing from the field to country offices and thence to HQ. This is not necessarily about replacing specific monitoring activities but rather about being clearer about what monitoring is intended to achieve and how, and changing mind-sets about what monitoring can and should accomplish. Currently, monitoring activities tend to be a repository for a wide assortment of implicit and explicit aims and objectives related to learning and accountability. When problems arise, there is a tendency to call for ‘better monitoring’, without analysing the nature and content of current monitoring and the resources required to shoulder additional tasks, and/or the necessary resources to do what we already do, but well.

Greater downward accountability may mean that monitoring moves further in the direction of RTE and strategic review, as introduced earlier. Of course there are unanswered questions as to how RTE differs from strategic review, and how they both differ from monitoring. This will depend on how an agency defines the scope of monitoring itself. Since monitoring is usually associated with direct management tasks, especially data collection, the difference between these approaches may depend on how real-time evaluators and those involved in strategic review are able to take a step back from day-to-day reporting and administration and use their time for analysis, making sound judgements, and involving primary stakeholders.

A System under Strain?

1.5

In order to consider the idea of 'monitoring as a missing link' it is necessary to look at the wider picture of how monitoring fits into broader humanitarian structures as well as into the task environment of field staff. As this *Annual Review* illustrates, the big picture is one of humanitarian systems under great strain. Two key points can be identified. First, there is an increasing internal reporting burden on field staff and agencies due to multiple reporting demands, increasing earmarking from donors, and a proliferation of cross-cutting themes (e.g. human rights, gender equality, environment), all of which are important but all of which bring their own reporting requirements. It is probably the case that the 'accountability lobby' has also added to the burden. Anecdotal evidence suggests that multiple accountability initiatives are viewed with trepidation by field staff – not necessarily because they will reveal malpractice but because they could lead to a time-consuming round of additional workshops and reporting. In the end these initiatives may be more about repackaging field-level knowledge as 'quality assurance' information for donors and HQ rather than as useable support for addressing genuine dilemmas of practice.

Second, there are significant problems in enhancing monitoring in relation to human resource practice, especially in relation to excessively long hours of work and high levels of staff turnover, both of which mitigate against providing the time needed for staff to be able to contribute to, and learn from, the information produced by monitoring. Illustrative of the extremes that can be reached is the case of Oxfam in Angola, where over 32 international staff filled a total of 11 posts – an average of three incumbents per post per year. Many other agencies have had similar experiences as detailed in the themed chapter in last year's *Annual Review*, as well as, for example, a total of 12 evaluative reports this year that note the negative effects of staff turnover on performance (see Chapter 3). In such situations monitoring is less about learning than it is about damage control, as it may be the only way to ensure that incoming staff have some way of understanding what is going on.

The failure of good ideas about accountability and learning to result in widespread good practice may be indicative of blindness to the kind of pressurised work environments that humanitarian aid workers actually experience. It is within this context that genuine incentives to learn – through monitoring and other means – must be put into place. Instead of coming up with additional tasks there is a need to look at how people try to solve problems and make sense of their situation within

prevailing duties and responsibilities. The overwhelming quantity of information and reports that many offices have to produce may reflect a lack of awareness of the actual pressures of fieldwork in terms of poor communication between field and HQ. This theme is also highlighted in Chapter 3

There is also a danger that increased investment in information flows is not sufficiently related to how that information will provide the knowledge needed to deal with the complex and dynamically changing situations that define the working environment of the humanitarian field worker. Operational staff are frequently held accountable based on their capacity to provide more, but not necessarily more useful, information. The concern is that increased information flows may be at the expense of efforts to help transform information into knowledge, especially to embed efforts in an awareness of the context in which affected people struggle to survive and the potential impact of aid on their very survival. It is such wisdom – that is, the ability to apply knowledge to practical action – that the humanitarian sector needs to foster.

In this respect there seems little point in developing yet more methods and/or toolboxes if they are not preceded by a concomitant effort to streamline existing data collection responsibilities. While information needs will vary according to different organisational cultures, contexts in which monitoring is happening and the balance each agency requires between reporting for upward and downward accountability, there are positive signs. Chapter 2 notes an emerging awareness among donors and operational agencies that harmonisation of reporting requirements are needed to free up resources for other tasks. This must be a first step. Learning can only be improved if field staff are given the time to do it. The key message is one of sector reform.

Linking Relief, Rehabilitation and Development (LRRD)

1.6

Improvements in monitoring for both learning and downward accountability depend to a large degree on the expectations as to what humanitarian assistance sets out to achieve in the first place. Humanitarian programming today, for example, in

Southern Africa, Afghanistan, Sierra Leone, and Ethiopia, is increasingly intermingled with rehabilitation. This implies that humanitarian assistance is no longer only about saving lives but must also relate to the root causes of conflict and poverty. The sector is being called upon to deliver aid in a way that can ultimately reduce violence and promote recovery, development and peace. This ambitious agenda has major consequences for learning, downward accountability and performance. One hesitates to talk about paradigm shifts, an already overused term, but if these aims are genuinely to be realised, monitoring needs to become more than just a vehicle for upward accountability and must be redirected towards providing contextual information to fill information gaps that currently exist between the relief, rehabilitation and development phases of response.

ALNAP Annual Reviews 2002 and 2003 highlight that a significant proportion of humanitarian funding is actually being spent on activities normally associated with rehabilitation and development and, as Chapter 3 points out, evaluators are increasingly paying attention to this shift. Yet while the sector is beginning to ask the right questions as to LRRD, there are major deficiencies in finding the right answer. The synthesis in Chapter 3 focuses on this theme. It reveals an overall picture of short-term success in most direct emergency interventions, especially in health, water and sanitation, and food aid, but a failure to link short-term objectives with any real lasting benefit. For example, food aid may feed the hungry and save life in the short term but food-for-work schemes appear to provide little lasting benefit. Likewise, clean water is provided in camps but longer term maintenance by community groups and spare parts for pumps are lacking when it comes to providing water on a more sustainable basis. Furthermore, there is little or no sign of capacity building in humanitarian response, including of government and local institutions. As a consequence the most vulnerable are more likely to slip back into destitution when the initial phase of the relief intervention are over.

Perhaps the basic issue is whether it is realistic to expect longer term and more sustainable impacts from emergency interventions? LRRD policies paired with short funding cycles and so-called sunset clauses have created pressure on agencies to make unrealistic claims about the prospects for recovery in order to 'declare victory' and move onto the next humanitarian crisis. The move to results-based management in the sector has supported this tendency towards results inflation. In reality, the most vulnerable are being left behind as the rhetoric of LRRD moves ahead. When struggling to rebuild their livelihoods, the destitute, the disabled and

the landless often lack the resources to keep up with the project cycle. And, at the same time, agencies tend to ignore the coping strategies and capacity of affected populations.

The failure of LRRD on the ground is mirrored by the failure to find a useful synergy between the relief and development communities. Development actors who have knowledge about the nature of ongoing vulnerability and risk are still not engaging with the humanitarian agencies that ‘parachute in’ when, for example, structural food insecurity turns into acute famine. Recent experience in Southern Africa has taught us that chronic vulnerabilities caused by HIV/AIDS and changing political economies mean that the boundaries that are supposed to separate relief and development are becoming fainter still. Understanding the nature of chronic risk is something that needs to be addressed at various levels – namely research, practice and policy.

It may be postulated that this is due to a continued failure, within the humanitarian sector, to look more closely at the wider impacts of the aid provided on the lives of beneficiaries. Downward accountability means more than caring about whether the food was delivered and the bellies filled. It means caring about livelihoods too. Humanitarians may only have relatively blunt tools at their disposal with which to ensure that people are not hungry tomorrow, but that does not absolve the sector from the need to sit with development actors to discuss what needs to be done. And it may be that monitoring has an important role to play in providing information and feedback at key moments in the LRRD process.

A Summary of Contents

1.7

Chapter 2 begins by providing a brief overview of current monitoring frameworks and practices in the sector, as well as a provisional exploration of the links between monitoring and learning. It then covers the areas discussed in the last three sections of this introduction in relation to potential areas of gap filling, highlighting good practice – improving information flows, simplifying systems, promoting joint activity, focusing on process and impact – and strengthening downward accountability.

Chapter 3 provides the annual synthesis of EHA reports that were made available to the ALNAP Evaluative Reports Database in 2002 (55 reports). The chapter is organised around the main humanitarian action sectors of food aid, water and sanitation, health, and shelter and housing. Supporting sectors are considered in a table. In addition the chapter covers cross-cutting themes, including human resources and management, participation and consultation of affected populations, protection and human rights, gender equality, and results-based management.

The conclusions from this year's *Annual Review* are drawn together in Chapter 4, which stresses the continued importance of the agenda for change and action recommended in *Annual Review 2002*. In addition, some recommendations are made for reviewing and streamlining monitoring systems of humanitarian agencies. The coloured section of the *Annual Review* contains the meta-evaluation, ALNAP's annual assessment of the quality of the previous year's evaluation set. This meta-analysis is achieved through assessment against the ALNAP Quality Proforma (QP) which has been revised this year in light of experience from the 2002 meta-evaluation. In addition to this year's evaluation analysis there is a comparative analysis of the 127 evaluative reports which have been assessed in the meta-evaluation exercise since the *Annual Review* series was launched in 2001, based on comparable questions from the QP used across the three years.

Annual Review 2004

1.8

Next year's *Annual Review* will analyse what happens when humanitarians talk to development actors and primary stakeholders about what needs to be done in humanitarian action. It will focus on field-level learning for improving the understanding of the contexts of humanitarian action. Analysis will draw on ALNAP's experience of testing the LSO concept (Box 2.8), which takes a proactive operational approach to promoting and facilitating opportunities for field-level learning. Additional input will come from the findings of the ALNAP Global Study which has raised important questions about how far agencies are willing to go in operationalising their commitments to learn from affected populations.