

special mandate for humanitarian coordination, attracts absolutely no attention in the evaluations.

UNHCR, as designated UN Lead Agency, serves throughout as chief whipping boy, with the lowest marks for achievement (for example, due to its 'missing middle management'), although higher marks for effort and achievement as events unfolded. WFP gets the top mark for effective deployment of human resources and logistics. UNOCHA, confined to arranging military and civil defence packages from Geneva, producing situation reports in Macedonia, and coordinating with the Albanian government's EMG (Emergency Management Group), was virtually absent, so goes unmarked. UNHCR rather than UNOCHA participated in key EU-NATO coordination meetings.

Lack of strategic planning, and not just information sharing (considered by some evaluations to have been reasonable), carries a share of the blame. This lack was highlighted by Dr Bernard Kouchner (UN Secretary General Special Representative) at an August 1999 meeting in Pristina, when he lamented the fact that there was 'simply no strategic intelligence available' to allow him to tell a visiting new donor (Taiwan) how best to support the relief effort.

The imbalance of participation in Kosovo from within the UN system, as well as contributing to the overall coordination crisis, may reflect a coordination crisis within the UN itself.

A comparison of the imbalance of participation of big UN players with non-UN (but still big) players, including ECHO and OSCE, would be an interesting evaluation focus, as would a focus on UNDP (another part of the UN with notable coordination capacity and experience). It is likely, however, that only a system-wide evaluation would attempt this.

### 3.2.8 **A Crisis of Proliferation**

As with Rwanda, so with Kosovo, what all commentators call 'an unusually large number' of NGOs responded to the crisis. One document famously speaks of this conflict having attracted one NGO for every day of the year. Unfortunately no report goes beyond this rather simplistic 'large number' plaint to tell us just what all these small (and smallest) organisations do, where they come from, what resources they have, etc. They only reveal that having them there makes coordination difficult or nigh on

impossible. Their message is of the 'once again' variety where lessons have not been learned.

On the face of it, this is an extraordinarily limited, unconvincing, and trivial assessment. It is most unlikely that fewer of the smaller and smallest would have made that much difference to balanced participation and coordination. It reads less as evaluation than special pleading

### 3 2 9 **A Crisis of Unpredicted and Unpredictable Scale and Speed**

Many evaluations share the 'unpredictable nature of the crisis' view. First hundreds of thousands of displaced people streamed out of Kosovo, then, equally rapidly and unexpectedly – perhaps especially from Macedonia but also Albania – they streamed back again. Unexpectedly to whom?

The evaluations merely report that this was unexpected by the international relief agencies. Both the evaluations and the more research-based policy reviews are completely silent on the speed and scale of movement as seen and managed by the affected populations themselves. This is another instance of the virtual reality of 'foreign-aid land', owing to its social blinkering (and prejudice) about the social, cultural and other institutions of the populations they are meant to be serving

### 3.2.10 **A Host Nation Crisis: Maintaining Domestic Sociopolitical Stability**

In this complex emergency, governments were in place in Macedonia and Albania (as of course in Serbia). The Macedonian government in particular considered that the refugee outflow from Kosovo threatened its stability. At the peak of the crisis refugees from the Kosovo conflict constituted 15% of the Macedonian population.

The threat to host nations' domestic sociopolitical stability is, however, only dimly seen and appreciated by the evaluations, where the emphasis in this respect is always on humanitarian and human rights principles and international refugee law issues. Yet, if the Macedonian government had fallen this would undoubtedly have made the whole humanitarian scene even worse.

Instead of treating issues of regional government stability as integral to the humanitarian context, the tendency is to leave the

former completely outside the latter's framework, benefiting neither in practical nor policy terms.

What programmes and their evaluations do seem to have been aware of, however, was the regional poverty factor: 'The countries that bore the brunt of the Kosovo crisis are among the poorest in Europe' (DEC Vol. I, 2000c: p xvii); 'MCIC [Macedonian Centre for International Co-operation] in Macedonia explicitly voiced the need to "*balance the refugee crisis with the social crisis*"' (DEC Vol. I, 2000c: p 45).

### 3.2.11 **A Crisis of Overbearing Humanitarian Assistance**

Thankfully two or three of the evaluative reports make reference to additional problems created by the humanitarian response process itself: '[An] agency complained that the population was fully employed in reconstruction until the international community intervened' (DEC Vol. II, 2000c: p96). Dispossession is common in foreign-aid land, however well intentioned.

The dangers of negative impact are not unrecognised, but presaging voices clearly went unheeded. Written in April 1999: 'For a decade, ethnic Albanians have been forced to create their own parallel society, involving every domain from kindergartens to health clinics to architecture schools to a vibrant independent media. It is critical that the international community not "colonise" Kosovo with a sudden "invasion" of well-intended but overbearing humanitarian assistance. Such a misguided deluge of aid last autumn, in the wake of the ill-fated October [1998] ceasefire agreements, seriously jeopardised the integrity and growth of Kosovo's home-grown, still fragile, civil society institutions. It is clear from the many mistakes made by the international community in its efforts to rebuild Bosnia that a peremptory, non consultative approach using "in-and-out" commercial contractors can inhibit the efforts of a vulnerable, war-torn, population to get back on its feet' (International Crisis Group, 1999a p27).

### 3.2.12 **A Crisis in Assistance: Assisting a 'Non-Crisis'**

Despite all of the above, a theme that appears in some of the evaluations (particularly with reference to assistance) is that 'no classically defined emergency actually occurred' (UNICEF/DFID, 2000: p13); 'The refugees were relatively better nourished, healthier, and with access to greater resources (including, very significantly, from the Kosovar Albanian Diaspora) compared to

those in many emergencies ... [and] mortality and morbidity rates were generally well within emergency levels' (ibid.). Also: 'The humanitarian intervention as a whole can only take very limited credit for the low mortality and morbidity [noted by many commentators]. The quality of the sanitation and environmental health interventions would probably not have been sufficient to prevent large-scale breakdowns of public health epidemics and, as a result, excess loss of life, had the crisis occurred in a more usual emergency context' (DEC Vol. II, 2000c: p25).

WFP's recognition that 'coping mechanisms available to many of the affected groups ... played a decisive role in avoiding hunger and malnutrition ... [as] refugees, IDPs, and other vulnerable groups had their own contingency plans, emergency food stocks and resources to overcome periods of dislocation from their homes ... [and] themselves assumed responsibility for their survival and well being, particularly during the first weeks of the emergency' (WFP Executive Board, 2000a: p5) should also be carefully noted.

However, no organisation is reported as having considered there was anything but a huge programme of assistance to be undertaken, or that such efforts were inadequate or, in some regards at least, unnecessary.

### 3.2.13 **Calling a Conflict a Crisis: Some Discourse Analysis**

A final important point is that most of the evaluative reports, particularly the evaluations, tend to speak not of the Kosovo 'conflict' but of the Kosovo 'crisis' or 'crises'.

Talk of 'crisis' tends mainly to ask 'what?' questions (for example, What is the crisis? What is the assistance required?). 'Conflict' talk triggers more 'who?' questions (for example, Who is responsible for this mess? Who is fighting whom?). Of course there is no watertight divide between the two, especially when hinged on the use of two ordinary words that have overlapping as well as different dimensions.

The actual word 'conflict' makes only a very rare appearance – generally where some broader background or scene setting chronology is done. While the humanitarian response in Kosovo is not charged with fuelling the fire, the evaluations do say that much of the response missed the fire, in respect of protection and coverage. At best, only a third of those affected (those in refugee camps) received assistance.

The discursive practice of calling a conflict a crisis is not however without effect, intended or not. It focuses problem identification on crisis management rather than conflict resolution, peace and development, de-politicising the analysis. This leads to organisations such as the Mother Teresa Society and the Kosovo Liberation Army (KLA), both prominent local players, being in effect written out of the humanitarian scene and its analysis. They are labelled 'political' or 'military' and hence not 'approved' civil-society organs, which humanitarian agency discourse takes to be apolitical, somehow beyond politics and war. It is a grave gap in evaluation

Not one of the evaluations addresses its or its programme's mode of discourse. Commissioned evaluation as a genre is rarely self-reflective

#### 3.2.14 **Conclusion: Whose Crisis Anyway?**

The chief findings of this section are that the various actors were responding not to one crisis but to many crises, and that differences in perception as to the nature of these crises led to the use of different success/failure indicators.

The question of whether the crises were found, perceived or made was not asked by either the evaluations or the policy reviews, explaining their tendency to dwell on and over-emphasise donor relief supply-driven factors.

This myriad of crises (perceived or not) will of course have implications for whoever may take on the task of commissioning and planning a system-wide evaluation. It will need to look not only at all the players (international or national) but also at all the crises that lie within this single conflict.

### 3.3 Agency Response Issues

#### 3.3.1 Introduction

This section provides some synthesis of the evaluations' principal findings, conclusions and recommendations – with elements of meta-evaluation. They fall into three broad categories:

- 1 humanitarian policy, practice and strategic planning for emergency response;
2. organisation and management (O&M) issues for emergency response;
- 3 technical and standards aspects of emergency operations.

#### 3.3.2 Humanitarian Policy, Practice and Strategic Planning

##### ***To Prepare or Not to Prepare?***

All the evaluations consider preparedness as an issue and most take the view that none of the agencies involved seriously anticipated the scale and speed of the emergency. They find agencies' preparedness capacities greatly in need of improvement, but few make even minimal proposals as to how this might be achieved. There is no examination of how the preparedness lessons should be incorporated into policy, practice and strategic planning.

Given the poor record of early warning systems, it is suggested that investment should be made in the development of rapid reaction rescue capacities: 'The failure of "early warning" in the Kosovo case confirms the historic tendency of such systems to be unreliable or inadequate. Rather than develop its "early warning" capacity, UNHCR should strengthen its mechanisms to react rapidly' (UNHCR, 2000a: p xii).

##### ***Conflicts of Principle: Neutral Versus Impartial Assistance?***

It is important to understand the difference between neutrality and impartiality. The principle of 'neutrality' being the foregoing of taking sides in hostilities, and the principle of 'impartiality' that of non-discrimination, a commitment to deliver relief solely on the basis of priority-of-need, regardless of allegiance.

The evaluations of and for INGOs discuss at length whether the military can pursue military and humanitarian objectives that are both independent and simultaneous in a principled humanitarian way. Most of the evaluations conclude that it cannot, seeing a

fundamental contradiction in principle. Other reports, particularly those from NATO's lesson learning workshops, take an opposing view.

Conflicts of principle between military and civil agencies are exemplified by one of the more policy-oriented studies. The issue it addressed most directly was 'whether the Kosovo crisis exemplified the harnessing of the military for humanitarian tasks or the militarization of humanitarian action, or some combination of both' (Minear, 2000: p viii). Although many of those involved in the Kosovo crisis, both military and humanitarian contingents, saw this as 'the harnessing of the military for humanitarian tasks. Yes, the harness chafed here and there, but on balance the collaboration was productive. ... we as researchers read the Kosovo crisis as reflecting the militarization of humanitarian action, with certain ominous portents for the future particularly in terms of the politicisation of humanitarian access and activities' (Minear, 2000: p viii).

The differing views are characteristic of the corpus of Kosovo reports, and the issues introduced or intensified by NATO's involvement are seen as vitally important and underlined by all. But while some are afraid it presages 'humanitarianism by default' because of military mandates, objectives and conditionalities, others are more content to emphasise the immediate and temporal specificities, hoping that NATO's engagement brings only limited implications.

Areas of possible conflict of principle and image between military and civil agencies are only part of the picture. Another perspective on the 'conflict of principle' in the Kosovo context is the perception held by many within western society, as well as other cultural traditions, of humanitarian action as a new form of western imperialism

Finally, clashes of commercial, along with civil and military, principles and practice need to be considered as well: '[T]he humanitarian marketplace It becomes increasingly competitive, not least within the field of emergency response, and increasingly driven by donors and their priorities' (DRC, 1999: p20).

### ***Practice Ahead of Principle: Humanitarian Evacuation and Humanitarian Transfer***

The particular pressures of the Kosovo conflict, fast and massive outflows of refugees and resistance to hosting by regional govern-

ments combined with media demands, led to the introduction of two new strategies by UNHCR. These were Humanitarian Evacuation Programmes (HEP), involving movement out of the region, and Humanitarian Transfer Programmes (HTP), involving in-region cross-border movement.

These are examined in three or four of the evaluations with, on the whole, similar conclusions. Success in addressing such massive refugee flows would have been unlikely if such innovative practices had not been employed – i.e., principle needed to catch up with practice. ‘The innovative HEP resulted in an operation of unprecedented speed and scale that contributed positively to the protection of refugees by alleviating the burden on a reluctant host state that feared destabilisation’ (UNHCR, 2000a: p102).

Problems associated with both are explored including issues of design, execution, vulnerability, eligibility and fairness in selection. ‘.. HTP did not contribute significantly to protection during the emergency. Part of the limited implementation of HTP relates to confusion over the degree of voluntariness or consent required for these refugee movements ... HEP undermined HTP: UNHCR’s stated preference for protection options within the region becomes difficult to maintain if evacuations outside the region are available and preferred by the refugees’ (UNHCR, 2000a: p102).

### ***The Non-camp Refugee Population***

For those who remained in the region, a substantial proportion of the total refugee caseload found shelter in what the humanitarian community first called ‘host families’ and later ‘private accommodation’. This latter reference is perhaps not only because payment passed hands but also because non kith-and-kin relations were involved. This is not explored in the reports.

Although they represent the majority, little is revealed about the assistance and protection provided to the Kosovar refugees in Macedonia and Albania that did not go to camps. In Macedonia local religious charities and governments, rather than the international community, probably deserve the principal credit for what was made available. Needs assessment of these non-camped affected populations was even more neglected than that of those in camps, though arguably: ‘It was not significantly relevant to the well-being of the majority of the refugee population who were in private accommodation’ (DFID, 2000: p5). However, another



notes: 'For the first two months not a single food distribution reached the 27,000 refugees in private accommodation. ... This failure to reach refugees living with host families will be looked upon (alongside the failure to complete the registration of refugees while they were in Albania) as among the gravest shortcomings of the aid effort in Albania' (Porter, 1999: p22).

The term 'vulnerability' also took on a special sense in this complex emergency. Given the extent to which refugees were in receipt of remittances from relatives living and working, for instance, in Germany, vulnerability referred not, for example, to the elderly or the sick, but to those who weren't in receipt of such remittances. There is some evidence that only the poorest went to the camps and that some 'host families' charged rent for the 'private accommodation' they provided. According to Australian media reports the evacuees airlifted to Australia complained that their temporary (Australian) accommodation was far from consistent with what they had had at home.

### **Coordination**

As mentioned in the previous section, what emerges from the evaluations and other evaluative reports is the complexity of the coordination problematic. Individual evaluations tend not to take a comprehensive view but focus only on one aspect or another, making pronouncements accordingly. Furthermore, different stages in the conflict require different assessments. Overall, accounts agree that in certain areas there was little policy coordination to begin with, though this became marginally better as the emergency wore on. UN focal points such as food (WFP) and aspects of education (UNICEF) were apparently very effective throughout and it is also reported that, within and across some more specifically INGO concerns, cooperation occasionally went well.

What varies from one programme or agency to another, in respect of what was well done, may depend on the extent to which each consciously built on earlier experience and evaluations of that experience. The reports however give greater attention to explanations for lack of achievement: 'For reasons that are well-known, the Kosovo crisis brought in very substantial funds from member states and each wished to maintain a high profile in-country including undertaking their own programmes. This led to considerable confusion and lack of co-ordinated activities' (ECHO, 2000k, 2000l, 2000m: p5). But note also that:

'Within a few weeks as donor liaison mechanisms also gained in effectiveness, a more coherent and consistent needs-driven approach was evident. Coordination, nevertheless, proved particularly difficult to handle in this emergency. ... We should have articulated the many problems and constraints more clearly from the start' (Morris, 1999: p17).

What may come as a greater surprise is the comment that, despite shortcomings, adverse effects can be exaggerated: 'UNHCR shortcomings [in assistance and coordination functions] ... did not have grave consequences for the welfare of the refugees: indeed they were relatively minor in relation to the overall relief response. ... [Nonetheless] areas of demonstrated weakness and inability to rapidly meet its own standards of response affected the credibility of the agency [which therefore could be considered to have suffered the most]' (UNHCR, 2000a: p xi).

While the evaluations mention the constraints of consensual coordination, none make recommendations for change in this area nor examine it in other ways.

### ***What Constitutes an Emergency?***

As with any disaster or humanitarian crisis that occurs in a sovereign state, a state of emergency cannot be acted on until it is declared and recognised as such by the national government concerned. The simple question of whether there is an emergency to meet may be part of the problem of little or no response at the outset. It would also appear that key institutions lack clear rules and guidelines for consideration of what constitutes an emergency status and how it should be acted on.

Where guidelines exist, to what extent are they known and used? 'The underuse in UNICEF (even ignorance) of the emergency handbook throughout the operation is notable' (UNICEF/DFID, 2000: p34). The same evaluation comments: 'software may not be totally emergency friendly' (ibid.: p34).

In terms of whether a situation merits being regarded as having emergency status or not, one evaluation reads as follows: 'Between April and May 1999, morbidity and mortality rates in Albania and Macedonia remained below the threshold that signifies an emergency. ... [F]ar below the [mortality] threshold of greater than 1[death]/10,000[people]/day, the key Sphere indicator for an emergency. ... Data for this period are not available from Kosovo or the rest of the Balkans. Since the return of the refugees to

Kosovo a comprehensive health study undertaken among the Kosovar population indicates similar low levels' (DEC Vol. II, 2000c: p21).

There are also different types, scales, and durations of emergencies, each perhaps with their different requirements. With UNHCR the problem was partly that it 'responded to the Kosovo refugee crisis as if it were a "normal" emergency. Standard routines for a smaller or slower emergency were followed (although not always attained)' (UNHCR, 2000a: p x).

It is to their credit that some of the sources draw attention to these important considerations. Training issues that need to be informed by them are mentioned under the heading 'Human Resource Weaknesses' below.

A further area of concern is the use of indicators. Normally in evaluation a different concern is considered to deserve a different indicator, as is the case in most of the Kosovo reports. However, to gauge the success (or failure) of a programme in terms of, say, a 'no one died of famine or disease' claim, when neither famine nor disease constituted the emergency is, as some evaluations say, downright misleading, if not aberrant.

### 3.3.3 **Organisation and Management of Human Resources**

#### ***Lack of Standing Rapid Response Capacity***

As mentioned earlier, another striking story to emerge is that of an international humanitarian community that lacks a rapid response capacity. Rescue here does not refer to the humanitarian evacuation programme that eventually got underway in Kosovo, but to a strategic concept for initial agency reconnaissance. The US government's DART (Disaster Assistance Response Team) approach, which emphasises the assessment required for such rescue, may come close to what the international humanitarian system lacks.

Most evaluative reports, including those by the military, lament this woefully inadequate situation. For example: 'Recommendation 1: If ACT is to be operational [in emergencies] it must have an emergency management team that it can deploy immediately to make preliminary assessments and to prepare programmes. Where a multi-agency response is involved, this should include preparing common programming' (ACT, 2000: p10).

The lack of a specialist standby rapid response, whatever may account for it, is an absurdity. If one is ever established, particular attention should be paid to the rapid assessment skills required and whose capacities are to be built – ‘foreigners’ or ‘nationals’. What are needed in the first instance are rapid response assessment teams, not rosters of active service personnel. One evaluation notes that emergency response capacities have yet to be mainstreamed into UNICEF

Although the above is illustrative of the wider picture, it is encouraging to note that in response to a recommendation put forward by the evaluation WFP commissioned, the Emergency Preparedness and Response Steering Committee was created at executive staff level to ‘establish a better framework for early warning, preparedness and response to rapidly evolving emergencies’ (WFP Executive Board, 2000b: p1).

UNHCR’s evaluation noted that UNHCR was ‘not established to provide relief in emergencies ... If members of the General Assembly want to do this, they will get a refugee agency that is quite different from that they established 50 years ago.’ It noted, however, that: ‘It is clearly possible to turn UNHCR into a superbly efficient rescue service’ (UNHCR, 2000a: p xiv), if and when this is the policy decided.

### ***Training and Briefing***

Several evaluations comment on the paucity of available training material and consider this to be partly responsible for poor practice. This is another example of a crisis of the agencies’ own making. Happily some evaluation Terms of Reference called for an assessment of personnel orientation, guidance and training before and during deployment, but the evaluations generally paint a very gloomy picture. When does poor practice qualify as malpractice?

In terms of briefing: ‘[S]ome pointed out that if they had been briefed about field work and the importance of cooperation ... by a veteran field worker, the message might have been more convincing than coming from a desk worker’ (DRC, 1999: p25). Debriefing is also crucial, as is the use of the information gathered from the undertaking.

Agencies were noted as having recognised the useful combination of in-house training with external courses by organisations such as RedR, Sphere, and Merlin, but the reports, despite their vocal

criticism. offered little remedy other than one interesting suggestion that a travelling contingent of field-craft trainers should be established.

The symposium background paper, on which this chapter is based, proposed something akin to the 'Learning Support Office'<sup>5</sup> currently being developed by ALNAP.

### ***Human Resource Weaknesses***

Poor programme performance can often be attributed to poor human resource management (HRM). Unprofessional employment and people management appears to have been much as usual in this complex emergency. A familiar pattern emerges in respect of short-term assignments, lack of appropriate training, even lack of briefing on arrival. Gender is reported to have been mostly 'forgotten again', and the high level of staff turnover a destabilising factor, particularly where agencies lack systems and procedures to ensure continuity and enable new staff to become effective more rapidly.

Shortcomings are also identified in relations between foreign and national staff; between country offices and points of delivery; in the mismatch of specialists and generalists (mostly the absence of the former where required sectorally); between routine and volunteer personnel; and in respect of conditions of service and duration of deployment/assignment. One evaluation notes that: '[P]eople who have been deemed unfit for previous operations in the Balkans appeared again in the Kosovo operation' (DRC, 1999: p25).

Weaknesses in the management of organisations, especially middle management, are easy to diagnose. A common finding in many of the evaluations is that: 'In massive emergencies the agency should ensure the rapid release of middle managers by the immediate adoption of directive, rather than voluntary, deployment practice' (UNHCR, 2000a: p xvi).

These weaknesses are responsible for many of the undesirable actions and outcomes in the programmes performed. However, other than in particular cases (primarily the O&M reports), little provision is made either in the Terms of Reference or the evaluation team composition to allow organisational and managerial dimensions to be competently addressed. This is probably why so little progress continues to be made in the evaluation of these areas, even though the evaluations make a few

more proposals on the remedy for this than they do for training and briefing.

All the evaluations portray Kosovo as yet another case where a large-scale humanitarian response pressured agencies into recruiting and deploying personnel without training, although UNHCR personnel did report that they had received a certain amount of appropriate training.

The problem seems to lie particularly within INGOs, where even senior staff members were found to be unaware of their organisation's commitments as signatories to codes that affirm, amongst other things, certain human resource management norms: 'Although 5 of the 12 DEC agencies are signed up to the pilot People in Aid "Code of Best Practice" ... some senior staff in DEC agencies were unsure whether their agency was a signatory or not. It is not possible to say how well agencies conformed to the code' (DEC Vol. I, 2000c: p xxvii).

### ***The Importance of Existing Presence and National Staff***

It appears to be the exception rather than the rule that agencies insist on existing presence and national staff as a condition of response. However, in addition to improved timeliness and contextual understanding, it is suggested that agencies working through local partners are less likely to pursue their own agendas.

Two evaluations clearly note existing presence and national staff as enhancing factors: '[T]he effectiveness of this church work is due to a combination of factors: 1) they already operate on the ground ... ; 2) they had staff in place (whether national or missionary) who spoke the local language .. ' (Tearfund, 2000: p27); and in the case of UNICEF: 'The predominance of national staff in middle management positions had a direct impact, in that it facilitated a deeper understanding of needs and realities in the three countries visited than is the case for international agencies managed almost uniquely by internationals, most of whom do not speak the relevant languages' (UNICEF/DFID, 2000: p5).

### ***Lack of Social Knowledge and Learning***

Another important aspect of the story to emerge from most evaluations is the 'once again' factor. Once again 'foreign' humanitarian workers arrived ill-equipped in terms of their sociopolitical and cultural knowledge of the conflict-affected populations. Despite the Code of Conduct that the big INGOs

have all signed up to (introducing an ethical dimension), social and cultural ignorance of beneficiaries was the norm. The heavily loaded term 'beneficiaries', as discourse analysis would say, accounts for a large part of the problem.

An obvious example is the criticised UNICEF 'Superman Project' developed as part of a mine awareness campaign. As well as conveying the wrong message – that Superman will save you – when the intention was aimed at inculcating safer behaviour: 'Handicap International psychologists recommend the identification of role models within the community whose behaviour children can copy, rather than the use of fictional heroes ... [T]he age range of the target audience (7-14 years) is too wide, there are no cluster bombs featured although cluster bombs form part of the threat, and the material is in any case not immediately available in Serbian' (UNICEF/DFID, 2000: p24). This highly illustrative example is typical of many from other agencies.

Since it is well recognised that people in the face of a crisis deploy a range of coping strategies, including the decision to move, why did the programmes and their evaluations discover so little about them? One factor is that results-based management is neither oriented nor sympathetic to social knowledge needs, another is the genuine difficulty of systematic consultation in the first phase of an emergency response. This is, however, no excuse for a lack of consultation in evaluations that come later with more time.

The majority of evaluations do too little to cross the knowledge divide but the very nature of complex emergencies is that they are just that, complex: 'Consultations and cooperation with local communities are part of DRC [Danish Refugee Council] project practice. But emergencies, it was argued, left no time for consultation other than those with local authorities. The principle, all agreed, was valid. But the question was how [to do so] in the heat of the emergency' (DRC, 1999: p22). Agencies' basic day-to-day work inevitably involves adjustments, and learning, clearly not undertaken in a social vacuum.

Even where evaluation time is dedicated to learning about beneficiaries, through in-depth individual and family interviews, the focus is mainly on discovering whether assistance provided was appropriate rather than learning, for example, why a return 'more spontaneous than organised, and self-resourced than assisted' happened (UNICEF/DFID, 2000: p32).

### ***Lack of Organisational Learning***

The question of organisational learning relates to much of the above. Like most humanitarian evaluations before them, the Kosovo reports state that past supposedly learned lessons still have not been learned (or at least where learned have not been implemented). One evaluation records a member of staff asking: 'Why are questions about contingency planning and related staffing policy only now discussed. Since many years WFP has been handling emergency operations all around the world, and WFP's budget has shifted from the original 80% for development to almost 80% for emergency programmes. Why are constraints like these raised only now?' (WFP, 2000b: p27).

The fact that the evaluations make many explicit comparisons with earlier complex emergencies might help speed along the information base required for such learning. However, where previous evaluations are the source of such wisdom, use of them should include critical – and transparent – examination of their methodology.

While the need to focus seriously on the process of organisational learning is not being addressed, there are signs of more robust attempts at lesson learning from Kosovo. Evaluation Terms of Reference and team composition should be revisited to acknowledge the linkage between lesson learning and accountability, rather than adopting the view that addressing both represents overload.

### ***The UN Lead Agency Role***

UNHCR was selected as UN Lead Agency in Kosovo due to its refugee protection mandate under international law, although it is not a humanitarian relief organisation by mandate or arguably capacity. Why 'refugee crisis' should trump 'food crisis' or 'integrated and system management crisis' (therefore WFP and UNDP) is not explored, but the evaluations do note a lack of conceptual clarity around the 'Lead Agency' status, and its voluntary (consensual) nature. None however discusses it seriously nor makes recommendations.

Accreditation, particularly of the large numbers of NGOs, appears not to be part of the process of participation. Although raised as an issue, no evaluation offers recommendations, and overall the evaluations pay insufficient attention to the Lead Agency as an issue.



### ***Delivery, Distribution and Impact***

All the reports agree that, for whatever reason (they never address why): 'Uniform tracking, monitoring and evaluation systems and approaches were not applied in the emergency' (UNICEF/DFID, 2000: p7). Other similar statements: 'Neither the ICRC nor the Federation has a monitoring system that provides sufficient feedback' (ICRC, 2000: p ii); 'of particular concern is the absence of a system which reconciles logistics statistics (what was sent out) with relief statistics (what was distributed to beneficiaries)' (ibid.: p iii)

According to one of the military reviews: 'The US ambassador in Albania did not have a good understanding of the [Marine Air Ground Task Force's humanitarian] assessment team's role. ... The consequent lack of coordination and cooperation resulted in the departure of the assessment team before it had completed its mission' (US Department of Defence, 2000: p106). Although UNHCR's Village Needs Assessment Unit in Pristina was seemingly much appreciated it is not assessed in any of the evaluations.

One reason why programmes do not monitor their work in this area may be that agencies still live in the charity paradigm and are simply unwilling to dedicate funds to this purpose. This sits particularly badly in the case of well or over-funded programmes.

Most evaluations that mention 'needs assessment' conclude that remarkably little was done. The extent to which, in this emergency as in so many others, relief supply-side considerations and blueprint approaches ruled, was also noted. This was exemplified by responses that at one point threatened a mini-crisis in assistance because they were so inappropriate to needs: 'Unsolicited "supply driven" in-kind donations clogged the system (second-hand clothes). In some cases they contravened standards and guidelines for emergency aid (milk formula)' (UNICEF/DFID, 2000: p12).

The following appears typical of the scene as a whole: 'Uniform tracking, monitoring and evaluation systems and approaches were not applied in the emergency, nor was adequate capacity available for these functions' (UNICEF/DFID, 2000: p7).

Similarly 'There is an absence of meaningful reporting. Copious information – like the volume of inputs distributed – is available on the result of projects, but [there is] a paucity about the impact

on the beneficiaries along with their families and communities, limiting to a great extent possibilities for more pro-active planning in projects' (ECHO, 2000k, 2000l, 2000m: p4).

As per usual, cultural factors as they relate to needs appear to have been ignored, despite cultural sensitivity supposedly having been a factor in some recruitment. For example, most of the detailed accounts of psychosocial assistance observe that this was the area of service most dogged by definitional and conceptual problems, as well as culturally inappropriate 'Western' ways of diagnosis and treatment. However: 'The psychosocial programme was appreciated by a large number of recipients ("*Very useful and we needed it*" was a typical comment), even though psychosocial support was relatively slow in coming as it was treated as part of health and not as emergency relief' (ICRC, 2000: p 1).

It appears that situational food needs in Kosovo were over-emphasised, for example, in comparison to shelter requirements, and winter needs identified too strongly with shelter needs.

This synthesis takes the view that an unmonitored programme, besides being unworthy, is, in certain respects 'unevaluable', particularly now that standard evaluation of humanitarian action criteria – e.g., appropriateness, effectiveness, efficiency, connectedness, coverage, complementarity and coherence – feature widely in evaluation Terms of Reference.

### 3.3.4 **Technical and Standards Aspects of Emergency Operations**

#### ***The Sphere Technical Standards***

Despite the Sphere Technical Standards becoming more widely known and discussed, with supporters seeking to expand coverage beyond the current five technical sectors, the evaluation reports (and ALNAP Kosovo Symposium plenary discussions) indicate that they remain controversial, both in respect of their value in principle and their applicability in practice, where operational flexibilities and strategic priorities must carry the day. The principal findings of the evaluations as a whole are summed up in the following: 'The evaluation team found that awareness and application ... of the Sphere Standards was poor within most DEC agencies, their international networks and local partners. ... Some DEC agency staff questioned the relevance of [the] Sphere Standards in a European setting . . . [H]owever, the evaluation team felt that most of the Sphere Standards were applicable,

particularly those relating to assessment, monitoring and information systems' (DEC Vol I, 2000c: p xxvii).

### ***Codes of Conduct and Operating Principles***

As already noted, evaluations found little evidence that codes of conduct were familiar to the majority of agency or implementing partner personnel: 'The evaluation team found that awareness of the Code of Conduct and its principles ... was poor and little attention given to their application. ... Many local partners had never heard of either the Code or Sphere Standards, and where DEC agencies worked through local or international networks, other agencies in the network did not necessarily share the same commitment to standards. There was very little monitoring of adherence to the Code and standards in agencies' planning and implementation of programmes and no requirement from the DEC to report against them' (DEC Vol. I, 2000c: p75).

### **3.3 5 Conclusion: Saving Graces**

The findings, conclusions and recommendations above amount to what can only be called a severe critique of the aspects of the programmes they address. That the findings and recommendations synthesised here are neither new, nor Kosovo-specific, is a damning indictment of the humanitarian system.

That the less defensible aspects of the international response (including failure to reach the majority of those presumably in need) did not have a greater negative impact owes more to luck than judgement. It deserves to be remembered that, in this instance, needs were reported as apparently lower than typical in complex emergencies.

The expression 'saving grace', used in one evaluation, is taken up in another: 'The independent evaluation of UNHCR performance during the Kosovo refugee crisis acknowledges that the two main "saving graces" were the hosting by families and the refugees' ability to pay for rent and food. .. It also notes that this situation was unsustainable. In other words, the third "saving grace" was the short duration of the crisis' (DEC Vol. II, 2000c: p67).

This synthesis adds three saving graces of its own: the stability of the governments in the region; NATO's humanitarian contribution (though NATO is not normally considered part of the international

humanitarian system), and the strength of public opinion in terms of not wanting 'another Bosnia'

Almost without exception, the evaluations fail to explain what is arguably the most prominent social feature of this conflict and its response. That is, how the affected populations themselves saw and managed both the conflict and the vast and rapid movements that constituted their response to it. As well as elements of spontaneity and reflex, there must have been some institutional, structural and social capacity factors.

This synthesis cannot treat the issue substantively due to the absence of social information and analysis in the reports considered. As a result, it is hard to reconcile, for example: 'The generous mentioned provision of more or less timely and adequate food to the target groups by the WFP and other major food providers' (WFP, 2000b: p20); 'mortality and morbidity rates were generally well within emergency levels' (UNICEF/DFID, 2000: p13); and, 'the finding that the majority of refugees and internally displaced persons were in fact well nourished meant that there was little need to distribute the high protein biscuits' (ibid.: p19)

That a good part of the story is, anyway, one of a human rights crisis turned into a humanitarian crisis is more to the point.

## 3 4 **Meta-evaluation: Kosovo Evaluations as Learning Tools**

### 3.4 1 **Introduction**

This last section is an exercise in meta-evaluation of the broader enterprise of evaluation of humanitarian action itself, as revealed by the reports. The aim is to determine and then deconstruct some aspects of its nature and scope as currently practised, in order to look for the comparative advantage, and disadvantage, of such evaluation as a learning tool. The purpose is to come up with, and consider proposals for reform.

The dual objectives of evaluation, lesson learning and accountability, may make it a less effective learning tool than activities dedicated solely to learning – e.g., lesson learning studies or learning workshops. If so, where ought the emphasis to lie in terms of a reasonable expectation of the commissioned evaluation type? This chapter shares the concern expressed in three or four of the Kosovo studies that evaluation is perhaps currently asked to do too much, and forced to grapple with too many types of expectation, for which insufficient provision is made.

Part of the problem with present practice is that evaluations are more in line with what could be called knowledge enhancing practices than contributions specifically to knowledge management and implementation. As lesson learning practice, it is not critical enough – indeed in some regards distinctly *uncritical*. This can be traced to the fact that evaluation is often restricted in its focus to the stated policies and objectives of the programmes being evaluated. Assuming these to be sound, it proceeds to verify their attainment.

Other limitations stem from the lack of serious social and political analysis apparent throughout. This leads to remarkably simplistic ideas about treating the affected populations in terms of having ‘ethnic problems’ only. As a result, the social, cultural and political local context of programmes almost never receives the quality and quantity of attention it ought, even in those reports that at least attempt to do so

Such limitations are clearly also a manifestation of the constricting timeframes imposed on the consultancy framework. The adoption into humanitarian discourse of the phrase ‘humanitarian crisis’ instead of ‘Kosovo conflict’ has insulated the