

ALNAP

An inter-agency forum working to
improve learning and accountability in
the international humanitarian system

Symposium Version
28/9/00

KOSOVO HUMANITARIAN PROGRAMME EVALUATIONS: TOWARDS SYNTHESIS, META-ANALYSIS AND SIXTEEN PROPOSITIONS FOR DISCUSSION

Background Paper for the ALNAP Symposium
Learning-from-Evaluation:
Humanitarian Assistance and Protection in Kosovo
to be held in Geneva 17-18th October 2000

Raymond Apthorpe
Evaluation Training and Research Adviser
ALNAP Secretariat
ODI, London

Acknowledgements: Every peer-reviewer needs a peer-reviewer. I am most grateful to John Borton, Coordinator of the ALNAP Secretariat, not least for working with me jointly for two days on this assignment, and to Kate Robertson, Administrator, for her valuable comments on, and editing of, the text. I am grateful to John, Felicity Heyworth and Nicola Norman, for identifying and collecting the Kosovo evaluations considered in the odyssey that is this assignment.

CONTENTS

Section A. Purpose and Audience of this Paper

1. Objectives of this Paper
2. Approach and Methods: Meta-Analysis *and* Synthesis
3. The Evaluative Reports Considered
4. Difficulties Encountered

Section B. How the Crisis in Kosovo was Perceived.

1. Introductory Remarks
2. A Crisis of Humanitarian Principles Arising From NATO's Involvement
3. A Multi-Ethnic Society in Crisis: What does this Mean?
4. A Crisis of Protection of Human Rights
5. A Crisis of Dominant Donors and Unrestrained Funding
6. A Crisis of Lack of Rapid Response Capacity
7. Unbalanced Participation of the Big UN Players: a Crisis of Coordination
8. "The Small NGOs", a Crisis of Proliferation
9. A Crisis of Scale, Speed, Unpredictability and Unpreparedness
10. Concluding Remarks: Calling a Conflict a Crisis

Section C. A Synthesis of the Evaluations' Main Messages on Humanitarian Principles, Human Resources and Operations

1. Introductory remarks.
2. Principles, Policy and Strategic Planning for Humanitarian Practice
 - 2.1 To prepare or not to prepare?
 - 2.2 Humanitarian aid and conditionality in relation to civil and military activities
 - 2.3 Humanitarian evacuation and humanitarian transfer
 - 2.4 Armed protection
3. Organisation and Management Structures and Human Resources for Emergency Response
 - 3.1 Lack of standing rapid response capacity in the system as a whole
 - 3.2 Training and briefing in humanitarian arts and skills.
 - 3.3 Weaknesses in internal vertical management
 - 3.4 Organisational Learning
4. Operations, Technical and Standards Issues
 - 4.1 The non-camp populations
 - 4.2 Needs assessment
 - 4.3 The Sphere Standards
 - 4.4 Social learning about the affected populations
5. Concluding Remarks and Propositions for the Symposium

Section D. Meta-Analysis: Evaluating the Evaluations

1. Introductory remarks.
2. Lesson-Learning Evaluation and Lessons (As Ever) To Be Learned.
3. Opaqueness of Evaluation Methods.
4. Independent Evaluation: Achieved, Necessary, Relevant?
5. Where Does the Comparative Advantage of Evaluation Practice Lie?
6. Attribution of Bias
7. Concluding Remarks and Some Propositions for the Symposium on Evaluation Practice as Critical Inquiry

Annex 1. Master List of Evaluative Studies Obtained and Considered

Annex 2. Consolidated List of the Main Findings and Propositions for Consideration at the Symposium

Annex 3. ALNAP Questionnaire to the Evaluators

Annex 4. ALNAP Questionnaire to the Commissioning Agencies

Section A. Purpose and Audience of this Paper

1. Objectives of this paper

The Active Learning Network for Accountability and Performance in Humanitarian Assistance, ALNAP is committed to maximising the learning and accountability benefits that may be derived from any particular humanitarian programme. Two of the three themes adopted by ALNAP Full Member agencies for its 2000/01/02 workplan are 'Making the Evaluation Process More Effective' and 'Strengthening Accountability Frameworks in the Humanitarian System'.

Aware that an unusually large number of evaluations of the humanitarian response in Kosovo were being undertaken, ALNAP Full Members endorsed a three-stage process to allow available evaluations to be subjected to a meta-analysis and synthesis. The findings of this first draft would form the basis of discussion at a Symposium attended by some of the principal actors in the evaluation process and the operations evaluated. This version of the paper represents the first stage of that process. The second stage will be the Symposium, and the third an expanded and finalised version of this paper, to reflect the principal outcomes of the Symposium.

The objectives of this paper are to:

- (i) overview the nature, scope and quality of the humanitarian programme evaluations of humanitarian responses in Kosovo; and
- (ii) identify recurring themes in relation to the broad categories of humanitarian principles, operations and human resources.

The limited time in which to prepare this paper and the broad nature of the materials it covers have necessitated an inevitable degree of selectivity, abstraction and generalisation.

2. Approach and Methods: Meta-Analysis and Synthesis.

Twenty person/days were eventually allocated to reading and analysing the evaluations and preparing this paper. A further eight days have been allowed to expand and finalise the paper in the light of discussions at the Symposium. This will form the core of the first of the new Annual ALNAP review of humanitarian programme evaluations covering 1999-2000.

This paper straddles its two tasks of 'meta-analysis' and 'synthesis' somewhat precariously. The former providing quality control through an evaluation of the evaluations (also known as meta-evaluation), and, the latter bringing together the evaluations' main contents. While there is inevitable overlap in these two modes of inquiry, this paper seeks to present the synthesis first. More so than with individual evaluations, synthesis "can be proactive ... not afraid to aim to improve, not prove."

I proceeded in this task by 'moral marinating' in the sea of studies to be absorbed, the 'emancipatory reading' to avoid drowning. A sort of silent peer-review, but by one peer only (a team would undoubtedly have come up with something different).

The approach throughout is as strongly inductive as possible, working as it were up from the materials considered (rather than down from a deductive schema). However, induction is never induction alone. I draw also on five weeks spent in Kosovo and Macedonia in August-September 1999 undertaking research to develop a potential model for the Ombudsman Project¹ and three days in August 2000 spent peer-reviewing one of the Kosovo evaluations, another invaluable opportunity to learn. I am most grateful to all concerned.

Two questionnaires (see Annexes 3 & 4) were devised to elicit complementary material from the evaluators and the commissioning agencies. Responses are still being received and results will be reported to the Symposium and incorporated into the final version of this paper.

Finally, with the principal objective of the present exercise being to overview a set of studies, this paper avoids "ranking" or "grading" and, at this stage at least, direct quotation. All the documents reviewed are publicly available enabling readers to make up their own minds about individual cases.

¹ Recently renamed the 'Humanitarian Accountability Project'

Where the ALNAP Secretariat holds electronic copies, executive summaries (and other excerpts) of the studies in our sample are available on request. **Documents covered by this paper are listed in Annex 1.**

3. The Evaluative Reports Considered

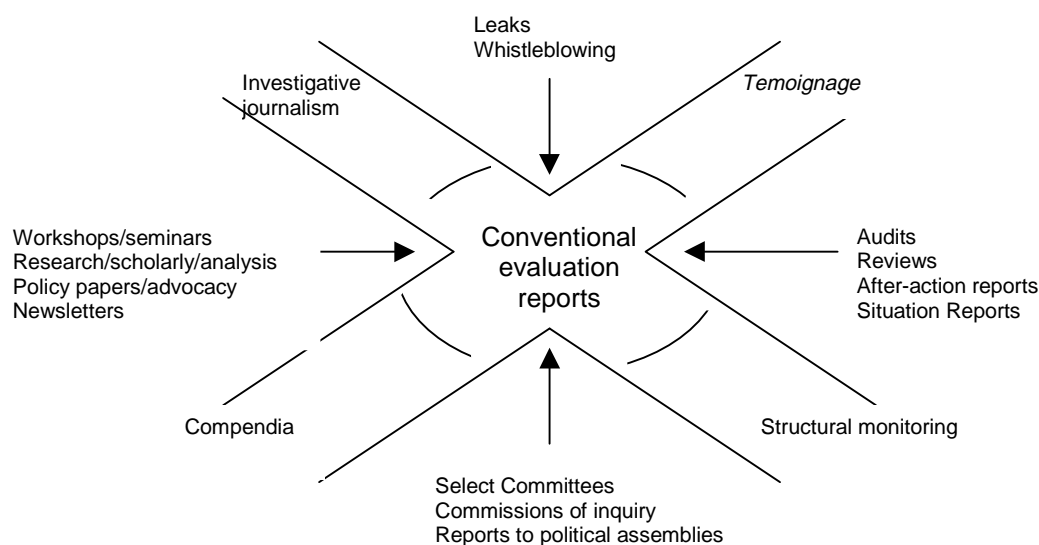
At the time this work was commissioned, the perception in the humanitarian world was that there were many closely overlapping evaluations to overview. However, as documents were gathered it became apparent that they were very mixed in nature. In the final analysis only 8 or 9 (for the sake of differentiation) fit the current criteria for 'conventional'² commissioned evaluations. The remainder comprise: internal learning workshops/seminars reports, reports to political assemblies, mid-term and after-action reviews, policy papers, public relations information prospectuses, research studies and missions, value-for-money and gender audits, newsletters, investigative journalism and advocacy and personal testaments.

Perhaps the term 'evaluation' is misleading and still open to interpretation.

All the conventional evaluation reports meet or exceed their Terms of Reference (TOR) and include good executive summaries. One is donor commissioned and the remainder split evenly across single agency or umbrella organisation commissions, whether in-house, by external consultants, or a mix.

The following diagram illustrates the range of genres of different evaluation-type documents in our sample (for expository purposes, it includes the rogue genre of leaks and whistle-blowing - which this paper does not draw on):

Figure 1



Inevitably the figure is not complete nor each document type completely discrete.

² Definition of a 'conventional evaluation' as used in this synthesis and meta-analysis.

A formal process designed to draw lessons to improve subsequent performance and/or improve accountability that has the following characteristics:

- It is commissioned by an organisation and evaluates the humanitarian response performance of either that organisation or of a partner organisation
- It is undertaken by a team of 'outsiders' – usually consultants/researchers or by a mixed team of consultants and 'insiders' i.e. personnel from either the commissioning organisation and/or the organisation being evaluated
- It assesses performance against evaluation criteria including several (but not necessarily all) of the following: impact, effectiveness, efficiency, appropriateness, coherence, connectedness
- It is increasingly most likely to assess humanitarian principle and policy too.

It articulates findings and draws conclusions and recommendations for the agencies to use in learning lessons

Overlap in the material is considerable with issues sometimes pursued with equal or greater clarity or wider reference in the non-conventional evaluation sources.

4. Difficulties Encountered

Evaluation reports that are not included are those undertaken by Tearfund, Catholic Relief Service, Spanish Red Cross Society, German Red Cross Society, CARE-International (Main Report), DFID/UNICEF, and ECHO. Some of these are still in progress whilst others may have been kept back for 'internal use only'.

Given the bulk of the documents obtained, the scale and detail of the programmes they cover (including *their own* not inconsiderable problems of coverage), and the diversity of *genre* (including *the familiar problem of apples and pears*), any serious attempt to fulfil the task set this paper will inevitably engender a high degree of generalisation.

Despite all of these studies being placed in the context of Kosovo and its surrounds, no two go into exactly the same spatial, temporal and organisational aspects of it, or have equal depth, breadth, or credibility. They were undertaken at different stages of an ongoing dynamic situation, the changing nature of which is seldom sufficiently - if at all - recognised in the ways in which reports state their conclusions and recommendations. Instead of a systematic phase-by-phase approach to the presentation of these stages, the usual practice is an annexe is attached of "main events", with "history" and its analysis being left largely as that. This does not amuse historians.

Finally, none has a system-wide focus encompassing the humanitarian organisations involved, or humanitarian and non-humanitarian facets. Neither synthesis nor meta-evaluation can provide a good substitute.

Section B How the Crisis in Kosovo was Perceived.

1. Introductory remarks.

Section B distills and reviews, with some elements of meta-analysis, how these studies report back on the perceptions of those responsible for the humanitarian programmes. The evaluations' principal conclusions and recommendations, including the conduct and achievement of these programmes and the main messages on humanitarian principles, human resources and operations, are synthesised in Section C.

The reality is that the Kosovo conflict was seen as comprising many crises, differently responded to by different agencies and their different programmes.

2. A Crisis of Humanitarian Principles Arising from Nato's Involvement

All the studies recognise NATO's involvement as huge and decisive. Most see this as a contradiction of humanitarian principle. Whilst many recognise the practical advantages of compromise of principle, others speak of this involvement as being "beyond the humanitarian pale", threatening the end of humanitarianism. Relations with militias other than NATO are not often addressed.

3. A Multi-Ethnic Society in Crisis: What Does this Mean?

While all the documents in our sample perceive - like the humanitarian programmes they evaluate - the crisis as one of a multi-ethnic society in crisis, none ponder or probe the modes and meanings of "ethnicity" as the affected populations themselves understand and act on these. The presumption is that these meanings are all of the "identity" type of "ethnic hatred" and "ethnic purity" discourse. Is this why they do not explore ethnicity as just one cultural and social marker among others, an indicator of economic standing and condition, a symbol of geopolitics, history and place? Or are they ignorant of sociological and social anthropological trends in ethnicity studies? This void makes it difficult for an induction-led synthesis and meta-evaluation to critique the studies effectively on this point.

4. A Crisis of Protection of Human Rights

Some of the studies perceive the main crisis in Kosovo as one of human rights. One or two, therefore, directly query the validity of "no famine", "no morbidity", as a key success indicator, saying that, as it was not mainly an assistance crisis, this is beside the point.

However, it could be said that in the case of the majority of Kosovo evaluative studies the focus on humanitarian assistance deflects attention from what ought to be at least an equal focus on human rights advocacy and protection.

A puzzle in itself is that none of the studies, other than the NATO ones, seriously address armed protection, although they make reference to "physical security". While of course not all protection is armed protection, even the evaluations which have a protection focus and give space to "end of humanitarianism" issues, fail to cover it.

5. A Crisis of Dominant Donors and Unrestrained Funding

Many of the documents refer to the overall response as being variously unstable, uncoordinated and inefficient. Some refer to a crisis driven disproportionately by the geo-politics of pre-eminently the US and the UK.

Many regret the heavy bilaterality of the response, reflected not least in the extent to which donors contracted humanitarian assistance resources and responsibilities to (their national) INGOs rather than to the multilaterals. Where the former accepted the lack of impartiality and neutrality with a degree of embarrassment, it did little to influence the bilateral approach.

Largely relating to this bilateralism is the flood of funds and available aid. As is usual elsewhere the relief response was donor driven, but, unusually, not by a paucity of aid but by its abundance.

6. A Crisis of Lack of Rapid Response Capacity

While a lay-person might assume the initial lead action and assessment following the crisis was undertaken by first-alert emergency units, structures, organisations and personnel, various documents point out this was not the case. Standing rapid emergency response capacity was simply unavailable.

I was myself surprised that I had to be reminded by one of the documents that even UNHCR is not by mandate or historically a relief organisation. If so, what is its comparative advantage as lead-agency in an emergency?

No emergency readiness on the part of the humanitarian relief system to react rapidly and specifically for humanitarian rescue? No organising concept of "rescue" as in "prevention" rather than "cure" within the humanitarian relief system? Both odd and serious. Are "durable solutions", "relief-development continua", "sustainability" and the like to blame for crowding it out?

7. Unbalanced Participation of the big UN players: A Crisis of Coordination

Where the documents perceive crises of leadership and coordination, UNHCR serves as the chief whipping-boy getting the lowest marks for achievement (e.g. for its missing "middle management") although higher marks for effort. Where human resources are central, WFP gets the highest marks (e.g. for "fielding its best people"). However, OCHA's absence, - therefore no marks - which is surely a very pertinent matter for evaluation, is not addressed in any of the documents (unless I have overlooked it). Perhaps this is because only a system-wide evaluation could plausibly have been expected to do this: or similarly, assess what UNDP's role might have been.

Of course, "unbalanced participation" is not the same as "lack of effective coordination" or "poor leadership". The prevalent view in the documents is that, as a whole, it is not perceived as "a crisis in coordination", despite critique of leadership failings and the critical impact that the lack of coordination had on the response.

8. "The Small NGOs": A Crisis of Proliferation

An unusually large number of NGOs appear to have been involved in the response. One document speaks of Kosovo having attracted "one NGO for every day of the year". However, apart from referring to their "great number", precious little else is said about the 'small' and 'smallest' actors other than they "make coordination difficult". Significantly, the number of INGOs is not mentioned.

9. A Crisis of Scale, Speed, Unpredictability and Unpreparedness

All of the documents speak of, and treat the crisis as one of huge and rapid displacements of hundreds of thousands people, first out of Kosovo, then, equally rapidly and, from the agencies' perception, unexpectedly back again. This was of course the aspect of the conflict most visible in the media.

The location of refugees to a great extent outside camps, with what were initially called "host families" and later, because of the payments involved, "private accommodation", creates both problems of social learning and access.

10. Concluding remarks

That most of the document, and in particular the conventional evaluations, speak of "the Kosovo crisis" - or crises - but not of "the Kosovo conflict", deserves comment. Whether their calling a conflict a crisis was by habit or choice is not clear, however, such parlance is not without effect of its own, whether intended or not. It lifts discourse on to an abstract plane (where the atmosphere is thin and heady and great for "crisis management" dreams). It allows - or actually encourages - big ideas to reign over organised interests. For example, if the KLA or even the Mother Teresa Society appear in these documents, it is only to be "written out" from the humanitarian scene on the grounds that they are "political" not "civil society" organs.

"Crisis" discourse is prone to just asking "what?" questions (e.g. "what is the crisis?"). In contrast, "conflict" discourse triggers more "who?" questions (e.g. "who is fighting who about what?"). Where our documents do bring "the people" in, it is for the greater part with reference to their economic vulnerability, poverty and needs "to be targeted". Their interests and organisational, political and religious ideas are omitted, and where they appear, reduced to the sorts of local aberrations that can be dismissed as, just local colour. In short, the documents, and the conventional evaluations in particular, are low on political economy sensibility.

Section C. A Synthesis of the Evaluations' Main Messages on Humanitarian Principles, Human Resources and Operations

1. Introductory Remarks

This Section provides, in thumbnail snatches, a synthesis of the documents' principal conclusions and recommendations of the documents under consideration which can be broadly categorised under the following three headings:

- Principles, Policy and Strategic Planning for Humanitarian Practice;
- Organisation and Management Structures, and Human Resources for Emergency Response;
- Operations, Technical and Standards Issues.

The nature of this exercise and its limitations, however, obliges a high degree of selection and generalisation. A more detailed and variegated treatment of each would be ideal.

2. Principles, Policy and Strategic Planning for Humanitarian Practice

2.1 To prepare or not to prepare?

All the documents consider the preparedness issue. All agree that the scale and speed of events were not anticipated. Most say preparedness needs to improve but few make proposals as to how this can be achieved, with little attention paid to organisational learning issues or to the substantive specifics found to have been lacking.

One document, after remarking that the Kosovo case confirms the historic tendency of early warning systems to be unreliable or inadequate, recommends that investment in the development of rapid reaction rescue and associated mechanisms should be prioritised over and above investment in early warning capacity.

2.2 Humanitarian aid and conditionality in relation to civil and military activities.

Many of the documents discuss at length whether the military can pursue military and humanitarian objectives simultaneously. Most find this a contradiction of humanitarian principles while agreeing with the INGOs concerned that practicalities in the Kosovo context limited the scope for objection on principle.

None of the documents, however, address how organisational mandates, implicitly or explicitly, act as forms of conditionality. The civil society focus insisted on by many INGOs for their "rebuilding society" paradigm cannot be apolitical, **especially** in the wake of a heavily political conflict. How aid discourse can construct "civil society" as non-political is difficult to understand.

This paper suggests that for issues of humanitarian principle to be analysed effectively, to the point of reaching comprehensive conclusions and recommendations, a broader canvas is needed than this set of documents. Issues of mandates, organisations' identities, concepts of independence and expertise in both programme design and management, as well as tensions in the managing and doing of evaluations, need to be part of the picture to avoid addressing neutrality and objectivity in the abstract.

2.3 Humanitarian evacuation and humanitarian transfer.

The practices of humanitarian evacuation and humanitarian transfer are examined in some documents and noted as innovative, with varying implications for humanitarian principle. Although administrative problems are explored, particularly in the case of evacuation (e.g. eligibility, fairness in selection, etc) the value of both is recognised, particularly so in the case of humanitarian transfer in respect of future emergencies, especially in this region. These innovations are seen as instances where humanitarian practice is in advance of principle.

2.4 Armed protection.

Like "humanitarian assistance", "humanitarian protection" comes in many forms, often portrayed as imperative to the protection of human rights. The armed protection of minorities is very prominent in Kosovo (one personal testament claims that "50%" of NATO's efforts or resources - it is not clear which - were expended on this activity).

Although not all protection in a conflict is armed, armed protection is nevertheless much in evidence and needed in Kosovo. However, even those evaluations which give space to consideration, conclusion and recommendation on whether or not military involvement compromises humanitarian principle or service, for the most part do not address armed protection either in this particular context or when pondering protection at large.

While it is a welcome development in humanitarian programme evaluation that protection is considered along with assistance, there remains much to be achieved in terms of reaching evaluation excellence in this area.

3. Organisation and Management Structures, and Human Resources for Emergency Response

3.1 Lack of standing rapid response capacity throughout the system

Many of the documents refer to inadequate specialist rapid response capacity within the various organisations that make up international humanitarian relief system. This is seen as the result of organisational mandates not focussing primarily on the provision of relief and where emphasis is strongest on excellence in development. In either case, a result is inadequate provision of specialist rescue-oriented capacity.

In the absence of dedicated rescue capacity, the responsibility for initiating and managing the humanitarian response fell on operational and regional desks, reported as inappropriately staffed for this purpose. 'Rescue' here is as in health indicators studies, where 'preventative' contrasts with 'rescue' with 'curative'.

3.2 Training and briefing in the humanitarian arts and skills.

The evaluations portray Kosovo as yet another case where a large-scale humanitarian response pressured agencies into the recruitment and deployment of personnel in breach of agreed good practice as regards training and pre- and post-briefing. Various, rather conventional, recommendations are made that this should be remedied immediately. The implication in practical terms (not spelt out by the reports) is that only specialised provision at field level could compensate for continued 'bad' practice higher up the organisational ladders.

3.3 Weaknesses in internal vertical management

Several reports indicate weaknesses in the internal vertical management of those organisations within the humanitarian response system. Weaknesses and difficulties were identified at regional coordination level; between head and country offices; between the country office and the points of delivery/camps; in the mis-match of specialists and generalists; between routine and volunteer personnel deployed; as well as in conditions of service and duration of deployment.

These weaknesses are found to be responsible for much of the variation in programme performance, however, in most cases, little provision is made in the evaluation TOR and team composition to allow organisational and managerial dimensions to be considered and addressed competently. Perhaps this is why so little progress continues to be made in these areas.

3.4 Organisational learning

Like others before them, the reports in the main say or imply that past lessons are still *not* being implemented by organisations. However, a welcome development is the extent to which these reports make explicit comparison with other earlier complex emergencies and humanitarian responses. Attempts are made to derive models from precedents, for example, on inter-agency relations (including civil-military), registration and coordination. It is still uncommon for evaluations to explicitly compare assessments of one situation and case with another.

The question this background paper must ask is how well such comparisons are made, and which judgements could and should have been informed by them. For example, from Rwanda genocide and post-genocide documents something could well have been learned about why a nuanced approach to issues of ethnicity is required, and how this might be done. (This said, even these often fail to probe deeply enough into "the two tribes" interacting as occupational and cultural classes or castes instead of, or as well as, essentialistic identities).

More, and better, recourse to such explicitly comparative analysis may contribute greatly to organisational learning. However, such knowledge has to be historically situated and itself evaluated to ensure appropriate delivery into the cocoon of a-temporal and a-paradoxical "good practice".

4. Operations, Technical and Standards Issues

4.1 The non-camp populations

A substantial proportion of the total caseload did not go to camps. The reports reveal the system experiencing intensified problems of registration, needs assessment, assistance delivery standards, and protection standards with regard to the non-camp population. Relationships between host families and refugees were more assumed by agencies than known. The implications for the response were inequities with over-focus on the camps.

4.2 Needs assessment

The reports show the significant extent to which needs assessment was strongly influenced by supply considerations and blueprint approaches as exemplified by responses in kit or package form. In addition, cross-cultural factors appear to have been under-estimated. In respect of food, needs were over-emphasised compared to shelter requirements (and winter needs identified too strongly with shelter needs).

Some - but not all - of the reports found psycho-social needs assessment and provision to be dogged by definitional and conceptual problems.

A UNHCR unit dedicated to compiling area needs assessment, arguably receives less attention in these evaluations than is due.

Evaluations of complex emergencies everywhere acknowledge structural, seasonal and situational aspects of need. The Kosovo reports do likewise with seasonal dimensions related particularly to "winterisation".

4.3 The Sphere technical standards

Several documents consider whether the Sphere standards were used and if so whether they were appropriate or not. A general finding was that where inadequate and/or inappropriate supply-driven pressures combine with inadequate needs assessment, Sphere minimum standards may not be attainable, even where they are considered appropriate.

Most of these reports tend to accept the value of the standards and propose their greater dissemination and expansion beyond the five technical sectors, e.g. to include community services. At the same time, one report strongly disputes this, holding that such standards are irrelevant distractions in situations where operational flexibility and strategic priorities must carry the day.

4.4 Social learning about the affected populations

With a couple of interesting exceptions, the set of reports gives inadequate attention to establishing the affected population's perceptions of the nature of the crises, their needs in respect of these, and their responses to the assistance or protection provided.

In the view of this synthesis and meta-analysis, this continuing neglect reflects both the agencies' **and** the evaluators' failure to understand that, even in circumstances which to outsiders look particularly chaotic, the interests and social organisations of affected populations can make positive contributions to humanitarian programmes.

5. Concluding Remarks and Propositions for the Symposium

Below, provocative propositions for Symposium consideration are linked to the reports' main findings.
For ease of reference, the main findings and propositions have been collated in Annex 2.

1. Principles and strategic planning for humanitarian practice.

C. 1.1 Concern: Where margins of unpredictability are great, investment in preparedness might be better spent on developing complementary (and better) alternatives.

Proposition: Emergency reactive response capacity should have priority over improved early warning

C.1.2 Concern: The evaluation of aid and conditionality in relation to civil and military activities, requires a broad canvas which includes mandates, agendas and organisational identities of all actors, not just neutrality and other abstract ideas.

Proposition: There is no aid that is unconditional

C. 1.3 Concern: The innovative practices of humanitarian transfer and evacuation achieved much on the ground, despite difficulties of principle (and international law).

Proposition: Both humanitarian evacuation and transfer should be added to the package of immediately deployable humanitarian sector responses, with emphasis on humanitarian transfer.

C. 1.4 Concern: Although armed protection can seldom, if ever, be provided by NGOs and similar channels, there is no reason for principled practice analysis of humanitarianism to ignore it.

Proposition: If armed protection is not included within the framework and analysis of the humanitarian system, then it will inevitably become the responsibility of external actors with other objectives

2. Organisation and management structures, and human resources for emergency response

C. 2.1 Concern: The lack of specialist standby rapid-rescue response capacity across the relief system is an absurdity.

Proposition: The organisation and management of rapid response is subverted by the lack of expression of the concept of 'rescue' on both strategic and operational planes.

C. 2.2 Concern: Training and briefing are still not central to the humanitarian sectors activities, with the resultant lack of professionalism becoming entrenched.

Proposition: A specialist organisation is required to provide mobile field-based training and briefing on-site facilities for agency personnel whose organisations continue to fail to do this.

C. 2.3 Concern: Weaknesses in internal vertical management, open to fairly easy redress, continue to undermine aspects of humanitarian implementation.

Proposition: The 'no-man's land' of middle management represents a neglected portion of the management chain which requires explicit attention by agencies and evaluations.

C. 2.4 Concern: Structured programme monitoring, a crucial means of ensuring that assistance and protection are actually delivered to those in need, is still not standard practice.

Proposition: That an unmonitored humanitarian programme is unworthy as well as unevaluable.

3. Operations, technical and standards issues

C. 3.1. Concern: Besides those internally displaced by conflicts, many of the externally displaced are refugees not in camps. Can existing deployments in the humanitarian system reach and cope with these?

Proposition: The humanitarian system cannot reach a large proportion of the affected populations.

C. 3.2 Concern: Inappropriate situational and cultural needs assessments continue to be driven by supply pressures and preconceptions to the detriment of the affected populations.

Proposition: This trend will not be reversed until humanitarian actors undertake serious social learning about the affected populations.

C. 3.3 Concern: The Sphere technical standards remain controversial both in respect of their value and applicability. They are, however, becoming more widely known and perhaps accepted, with some seeking to expand their coverage.

Proposition: The coverage of Sphere standards - as benchmarks - is inadequate.

C. 3.4 Concern: Social learning about affected populations continues to fall short of the codes of conduct to which agencies are signatories.

Proposition: The current inadequacies in the practice of social learning exist because there is neither the desire nor the techniques, among organisations or evaluators, to correct this shortcoming.

Section D. Towards Meta-Analysis: Evaluating the Evaluations

1. Introductory Remarks

The remarks by way of meta-evaluation in this section continue the practice in this paper of covering all the documents seen, but with the principal attention paid to the conventional evaluation reports.

These Kosovo reports may be the best of the 1999-2000 crop of humanitarian programme evaluations worldwide - indeed there is no reason to think they are not. In two particular respects they may be in the forefront of trends: the integrated way in which they tackle both humanitarian principles and emergency response operations, and their trend towards making explicit contrasts and comparisons of Kosovo with other complex emergencies. Shortcomings, where they exist, could be those of the current state-of-the-art generally.

One of the broader issues raised in Section B includes the extent to which it is wise for evaluators to share too many basic characteristics with those they evaluate. This is a perennial question for meta-evaluations to ask - and perennially difficult to answer.

2. Lesson-Learning Evaluation and Lessons (As Ever) To Be Learned

Lesson-learning is stated as a principal objective of most reports, documents and workshop outcomes, whether for single agencies or umbrella organisations, with the learning tending to focus on those lessons yet to be learned rather than those already learned.

Given that the Kosovo evaluations, like so many preceding them, report on the low uptake of past lessons recommended (research is needed to test this), this meta-analysis must ask where the problem(s) lie(s). Is it with the current approach to doing "lesson-learning evaluation", the humanitarian system as a whole or in part, or does the problem lie with both?

If the lesson identified by an evaluation is not matched by a competent recommendation, then a credibility gap emerges that inevitably undermines the process through which the lesson might be effectively learned by the organisation(s) concerned. All the standard appropriateness, timeliness, effectiveness, etc. criteria now used by humanitarian programme evaluation should also apply in an evaluation of evaluations.

3. Opaqueness of Evaluation Methods

Unexceptionally, the Kosovo evaluations provide background information on "team's method of work", "people interviewed", etc. This kind of brute methodological information (say M1) is usually of a simplistic nature and of limited use in meta-analysis, pointing to the need for a second more detailed level of methodological information.

This second level (M2) might include analysis of: the division of labour and time in relation to what was planned and then done; issues of evaluability encountered and how these were resolved (or not); an outline of difficult relations with parties involved and how these were handled etc. The Kosovo sources provide virtually nothing at this level.

A third level of information (M3) might include: informing readers and would-be users of how "findings" were established to the evaluators' satisfaction; how the cognitive process developed through the interpretation of findings into conclusions; and, how the transition was made from conclusions to recommendations.

As for "hard findings" in the sense of "hard data", probably no evaluator (or meta-evaluator) would disagree that this is particularly difficult to establish with any degree of confidence in an evaluation of emergency response. Research pieces, with less time-constraints and assuming access, are usually more successful.

As for drawing conclusions from such data, there is probably also wide agreement that in a consultancy this is more like drawing a sketch, than drawing water from a well. Style and composition are as integral to the exercise as technical methods.

Reaching and judging recommendations is yet another process. The question of whether recommendations are true or valid, science or art, tends to be given much less weight than whether recommendations are feasible, affordable, familiar, etc.

This is no straight-line process, and the initial problem for a heavily induction-led meta-analysis is that, as a set of documents, the Kosovo evaluations are more opaque than transparent, providing no M2 or M3 level information. Is there a similarly opaque reason for this?

4. Independent Evaluation: Achieved, Necessary, or Relevant?

The conventional evaluation teams would all claim to be 'independent', even though paid and facilitated by the agencies they evaluate. Most of the non-conventional evaluative studies would also probably claim independence, although agency lesson-learning groups and the like are difficult to classify in this respect. The meaning and understanding of 'independent' may vary, but, 'not having been manipulated by the commissioning agent' and 'delivering a report that holds the views of its authors only' (and perhaps also the beneficiaries') is probably the most common interpretation. In general, however, an "independent evaluation" is held to be one where the authors are external consultants (or researchers).

Probing further, an evaluation team made up principally, but not entirely, of outsiders, may be better described as 'mixed' rather than 'independent'. Mixed teams could include a commissioning agency staffer from its evaluation, policy, planning or even a sectoral, department. More 'mixed' teams might include members drawn from two agency's programme staffers joining outsiders for an evaluation of a programme in which both agencies are intimately involved.

To deconstruct "independent" still further, other senses include emancipation from the strictures of a particular discipline or school or religion or region or culture, and whether evaluators report only to commissioning agencies, excluding affected populations' representatives. This ambiguity stems from a multi-stakeholder context and *which* stakeholder's perception of independence is being applied. It is also necessary to ponder the implications for our conundrum of whether a gender balance has been achieved in an evaluation team and its thinking.

Is independence in all or only certain of its guises more critical than say insider-understanding and knowledge of the commissioning organisation; or, substantive sectoral competence in say water or sanitation; or, excellence in a particular type or method of evaluation? These are all issues in respect of the credibility of an evaluation, not all of which can be attained at once. Other attributes may for certain purposes (for example lesson learning) be more crucial than maintaining independence in its strictest sense. In only some of these senses of "independent" are the Kosovo evaluations what they claim to be, and those senses in which they are independent may prove irrelevant or even a handicap.

5. Where Does the Comparative Advantage of Evaluation Practice Lie?

There has been insufficient diversity of approach in the Kosovo evaluations to meet the expectations set by their TOR. However, the TOR themselves do not explicitly call for diversity of approach in meeting diverse objectives (e.g. accountability, audit, lesson-learning); in addressing the different issues of principle versus practice, of organisation and management, or of human resource; or in dealing with different time frames. Some TOR call for impact analysis without making suitable provision for the task to be done.

In only one or two instances in our Kosovo sample is a scenario-like menu of options provided in the recommendation section, where the reasoned consequences of following one as opposed to another option is outlined but leaves the choice as to which "is best" for the client, to the client.

The above suggests that the most characteristic form of evaluation currently practiced is contracted-out management consultancy, where evaluators are paid primarily to approach the task from a management perspective. Other perspectives such as that of expert witness, or advocate of an ideology, or voice of the voiceless, are not characteristic of current practice, while not being entirely absent. These would involve acting and thinking as advisors whose comparative advantage lay in analytical guidance rather than in recommendation of management lessons.

Compared to the tasks of social research and formal inquiry which take longer and require region and situation specific intelligence and skills, conventional consultancy (rather than research) evaluations complain constantly about the lack of available empirical data to describe patterns of operational interaction within the humanitarian system. Either such intelligence is simply not available or, in the case of region-illiterate appointed consultants, not known, and its capture is thwarted by time and/or resources and/or skills constraints.

6. Attribution of Bias

If “independent” is difficult to pin down, “bias” may be even more so, especially if considered to be ever present in social and other inquiry. The key question, particularly for a meta-evaluation, becomes one about how best to detect and demonstrate it. The author of an earlier synthesis of “independent evaluations” commented on the extent to which he was able to trace “evident bias” given that each evaluation shared the same TOR, albeit evaluating different aid programmes. The evaluators’ own influences in terms of their institutional affiliations, previous work etc. strongly shaped their frames and discourses.

Since the Kosovo evaluations do not have shared TOR, it is not immediately apparent how this heavily induction-led meta-evaluation might go about attributing bias. The task is therefore postponed until questionnaire results are in. (see Annexes 3 & 4 for questionnaires)

As for this meta-evaluator’s bias, there is no reason to think him any less influenced by his own beliefs and experiences - e.g. there is insufficient distinction between the products of consultation and research; that apolitical analysis is taken too far; that sociological and historical analysis deserves a greater role in evaluation; that the reality of the foreign-aid land, as part of the broader picture in which humanitarian work constitutes a principal part, must be addressed, in addition to the specifics of individual policies and programmes; and, that outcomes-oriented evaluations may best meet humanitarian concerns.

7. Concluding Remarks and Propositions for the Symposium on Evaluation Practice as Critical Inquiry

The following four provocative positions are posed for symposium discussion on the basis of the foregoing meta-evaluation of our sample of evaluative reports. Each is introduced by a few words indicating the area of concern which gives rise to it.

D.1 Concern: Evaluation through consultancy cannot compare with research, committees of inquiry, select committees and the like, when it comes to establishing “hard findings” with great confidence.

Proposition: Evaluation must find ways, with lowered expectations of achievements as regards “findings”, to focus on “conclusions” and “recommendations” by developing frames and discourses which permit this.

D.2 Concern: There is continuing ambiguity in expectations of evaluation as seen in their TOR. Expectations are not matched by variety of approach in the undertaking of evaluations.

Proposition: “Independent evaluation” is often no more than a credibility-placebo, and no guarantor of quality of output. When does independence actually serve evaluation objectives?

D.3 Concern: Lesson learning evaluations have not engendered the level of organisational learning, and system take-up sought or recommended. Conventional evaluation’s neglect of organisational learning may be as much to blame for this as the organisations concerned.

Proposition: Evaluation TOR and team composition need to be more comprehensive if lessons learning critique is to be lifted above the level of anecdote, excuse and praise. To address the missing element of organisational learning, evaluators need to develop new skills, aptitudes and attitudes.

D.4 Concern: Evaluation for critical practice may not be well served by evaluators always being asked to act and think from a management perspective, or by the conventional evaluation genre of contracted-out management consultation.

Proposition: Evaluators should think and act more as witnesses, advocates, spokespeople, that is as advisors.

ANNEX 1. Master List of Evaluative Documents Obtained and Considered (*=those classified by this Author as 'conventional' evaluations)

Commissioning Org.	Title	Authors	Type of Study
*ACT	Report of an Evaluation of EUBK91	Silkin, Kieffer and de Klerk	Evaluation
ACT	One, Two, Many...but not yet whole. Mid-term Evaluation Emergency Appeal	Silkin and Bourman	Mid-term Review
British Red Cross	Kosovo Crisis – IWD response, Achievements, Challenges: What Next?	Hanley	Internal Learning Review
*British Red Cross	Kosovo Schools Rehabilitation Programme Review	Tiller	Evaluation
Brookings Institution	Lessons of Kosovo	Daalder	Commissioned Lesson-Learning Study
Brookings Institution	The Lessons of Kosovo's Two Wars	Haas	
Brookings Institution	Unlearning the Lessons of Kosovo	Daalder and O'Hanlon	
CAFOD	Kosovo Emergency Programmes: lessons learning workshop		Workshop
CAFOD	Albania Emergency Preparedness Workshop		Lesson Learning Workshop
CARE International	The Kosovo Crisis – What can we learn?	Harvey et al.	Public summary of internal evaluation
DANIDA	Real-Time Evaluation of the Humanitarian Response to the Crisis in Kosova	ETC UK	Test of real time method
*DRC	Lessons Learnt from the DRC Involvement in the Kosovo Crisis 1998-99	Paludan et al.	Evaluation
DEC	Learning Lesson Study	Hallam, Kleyn and Kelly	Umbrella Learning Study
*DEC	Independent Evaluation of Expenditure of DEC Kosovo Appeal Funds - Phases 1 & 2	Wiles et al.	Evaluation
*DFID	Coordination in the 1999 Kosovo Refugee Emergency: the Emergency Management Group Albania	Telford et al.	Organisational Review/Evaluation
DFID/ESCOR	Solutions and responses to the humanitarian crisis in Kosovo	Barutciski	Desk International Law Study
EGDI	From one crisis to another: organisational learning in UNHCR	Suhrke	Seminar paper
* Groupe URD	La Mouvence Associative Albanaise, la crise Kosovar et les strategies des ONG Francaise	Grunewald et al.	Mission d'Etude
Groupe URD	Civilian humanitarian and military actors in Albania	Grunewald et al.	Mission d'Etude
ICG	'Kosovo Report Card' ICG Balkans Report No100, 28/8/00	International Crisis Group	Informed independent comment
ICG	'Kosovo: let's learn from Bosnia: Models and Methods of International Administration', 17/5/99	International Crisis Group	Informed independent comment
ICG	'War in the Balkans: Consequences of the Kosovo Conflict and Future Options for Kosovo and the Region', 19/4/99	International Crisis Group	Informed independent comment
ICRC	'Intl. humanitarian law and the Kosovo crisis: lessons learned or to be learned' Int. Rev. of the Red Cross No 837	Burger	Legal advisor's testament
ICRC	Conflict in the Balkans: human tragedies and the challenge to independent humanitarian action	Krabenbuhl	Delegate's testament
ICRC/IFRC/Nat.Soc	Technical Workshop on the Health and Relief Lessons Learned in the 1999 Balkans Crisis	ICRC/IFRC	Lesson Learning Workshop

*ICRC/IFRC/Nat.Soc	Balkans Evaluation: an examination of the role of the Intl. Red Cross and Red Crescent Movement's Response	Stone, Anema and Wissink	Evaluation
ICVA	The independent evaluation of UNHCR's emergency preparedness and response to the Kosovo crisis	Thomas, MacMillan and Schenkenberg van Mierop	Newsletter
MSF Holland	MSF Response in Macedonia to the Kosovo Refugee Crisis: A New Humanitarian Order?	Schulte-Hillen	Policy Review
NATO	Kosovo One Year On - Achievements and Challenges		Prospectus
NATO/EAPC	Compendium of views and experiences on the humanitarian aspects of peacekeeping		Compendium
NATO/EAPC	Seminar on the Kosovo experience with regard to the compendium on humanitarian aspects of peacekeeping		Seminar
Netherlands Government	Evaluation of the Kosovo Crisis		Evaluation
Netherlands Govt/Brown University/HLC	NATO and Humanitarian Action in the Kosovo Crisis	Miner, Van Baarda and Sommers	Research (interview/workshop)
OXFAM	Humanitarian response to Kosovo refugee crisis		News Briefing
OXFAM International	Winning the Peace? Some lessons from the EU's Aid to South Eastern Europe		Policy Paper
Refugee Studies Prog.	Learning from Kosovo		Journal (FMR) issue
SIRIUS	Preliminary Report: Kosovo and Operation Allied Force - Lessons to be learned		Lesson Learning Study
UK House of Commons	Kosovo	IDC	Parliamentary Inquiry
UK Ministry of Defence	Kosovo: Lessons from the Crisis		Internal Lesson Learning Study
UK National Audit Office	DFID Emergency Aid: the Kosovo Crisis		Value-for-money Audit
UN	Consolidated Inter-Agency Appeal for Southeastern Europe Jan-June 2000		Mid-term Review
*UNHCR	Kosovo Refugee Crisis: an independent evaluation of UNHCR's emergency preparedness and response	Barutskisi, Sandison and Garlock	Evaluation
Urgent Action Fund/Women's Committee	Gender Audit of Reconstruction Programmes in South Eastern Europe	Corrin	Gender Audit
US Dept. of State	Inter-agency review of US Government's civilian humanitarian and transition programmes		Lesson Learning Review
US Dept. of Defence	Report to Congress - Kosovo/Operation Allied Force: After Action Report		After Action Report
*WFP	Full Report of the Evaluation of the Kosovo EMOP 6040		Evaluation
WHO	Inter-agency medical/health taskforce: Humanitarian crisis in the South Balkans		Guidance Document

ANNEX 2

A Consolidated List of Propositions for the Symposium Break-Out Groups to Ponder and Probe, and perhaps Resolve.

The sixteen propositions that make up this consolidated list are collected here for convenience from the concluding parts of Sections C and D.

1. PRINCIPLES, POLICY AND STRATEGIC PLANNING FOR HUMANITARIAN PRACTICE. *(Section C)*

C. 1.1 Concern: Where margins of unpredictability are great, investment in preparedness might be better spent on developing complementary (and better) alternatives.

Proposition: Emergency reactive response capacity should have priority over improved early warning

C.1.2 Concern: The evaluation of aid and conditionality in relation to civil and military activities, requires a broad canvas which includes mandates, agendas and organisational identities of all actors, not just neutrality and other abstract ideas.

Proposition: There is no aid that is unconditional

C. 1.3 Concern: The innovative practices of humanitarian transfer and evacuation achieved much on the ground, despite difficulties of principle (and international law).

Proposition: Both humanitarian evacuation and transfer should be added to the package of immediately deployable humanitarian sector responses, with emphasis on humanitarian transfer.

C. 1.4 Concern: Although armed protection can seldom, if ever, be provided by NGOs and similar channels, there is no reason for principled practice analysis of humanitarianism to ignore it.

Proposition: If armed protection is not included within the framework and analysis of the humanitarian system, then it will inevitably become the responsibility of external actors with other objectives

2. ORGANISATION AND MANAGEMENT STRUCTURES, AND HUMAN RESOURCES FOR EMERGENCY RESPONSE. *(Section C.)*

C. 2.1 Concern: The lack of specialist standby rapid-rescue response capacity across the relief system is an absurdity.

Proposition: The organisation and management of rapid response is subverted by the lack of expression of the concept of 'rescue' on both strategic and operational planes.

C. 2.2 Concern: Training and briefing are still not central to the humanitarian sectors activities, with the resultant lack of professionalism becoming entrenched.

Proposition: A specialist organisation is required to provide mobile field-based training and briefing on-site facilities for agency personnel whose organisations continue to fail to do this.

C. 2.3 Concern: Weaknesses in internal vertical management, open to fairly easy redress, continue to undermine aspects of humanitarian implementation.

Proposition: The 'no-man's land' of middle management represents a neglected portion of the management chain which requires explicit attention by agencies and evaluations.

C. 2.4 Concern: Structured programme monitoring, a crucial means of ensuring that assistance and protection are actually delivered to those in need, is still not standard practice.

Proposition: That an unmonitored humanitarian programme is unworthy as well as unevaluable.

3. OPERATIONS, TECHNICAL AND STANDARDS ISSUES. *(Section C.)*

C. 3.1. Concern: Besides those internally displaced by conflicts, many of the externally displaced are refugees not in camps. Can existing deployments in the humanitarian system reach and cope with these?

Proposition: The humanitarian system cannot reach a large proportion of the affected populations.

C. 3.2 Concern: Inappropriate situational and cultural needs assessments continue to be driven by supply pressures and preconceptions to the detriment of the affected populations.

Proposition: This trend will not be reversed until humanitarian actors undertake serious social learning about the affected populations.

C. 3.3 Concern: The Sphere technical standards remain controversial both in respect of their value and applicability. They are, however, becoming more widely known and perhaps accepted, with some seeking to expand their coverage.

Proposition: The coverage of Sphere standards - as benchmarks - is inadequate.

C. 3.4 Concern: Social learning about affected populations continues to fall short of the codes of conduct to which agencies are signatories.

Proposition: The current inadequacies in the practice of social learning exist because there is neither the desire nor the techniques, among organisations or evaluators, to correct this shortcoming.

4. EVALUATION CONCEPTS, TECHNIQUES AND PROCESSES (*Section D.*)

D.1 Concern: Evaluation through consultancy cannot compare with research, committees of inquiry, select committees and the like, when it comes to establishing "hard findings" with great confidence.

Proposition: Evaluation must find ways, with lowered expectations of achievements as regards "findings", to focus on "conclusions" and "recommendations" by developing frames and discourses which permit this.

D.2 Concern: There is continuing ambiguity in expectations of evaluation as seen in their TOR. Expectations are not matched by variety of approach in the undertaking of evaluations.

Proposition: "Independent evaluation" is often no more than a credibility-placebo, and no guarantor of quality of output. When does independence actually serve evaluation objectives?

D.3 Concern: Lesson learning evaluations have not engendered the level of organisational learning, and system take-up sought or recommended. Conventional evaluation's neglect of organisational learning may be as much to blame for this as the organisations concerned.

Proposition: Evaluation TOR and team composition need to be more comprehensive if lessons learning critique is to be lifted above the level of anecdote, excuse and praise. To address the missing element of organisational learning, evaluators need to develop new skills, aptitudes and attitudes.

D.4 Concern: Evaluation for critical practice may not be well served by evaluators always being asked to act and think from a management perspective, or by the conventional evaluation genre of contracted-out management consultation.

Proposition: Evaluators should think and act more as witnesses, advocates, spokespeople, that is as advisors.

ANNEX 3. ALNAP Questionnaire to Evaluators.

1. Your previous experience of doing evaluations of humanitarian programmes.

- a. Before this Kosovo evaluation, had you carried out any other evaluations of humanitarian programmes? If so where, when and in what capacity?
- b. What specific lessons in relation to the evaluation of humanitarian programmes did you learn from this earlier experience?
- c. How did you apply these lessons in undertaking the Kosovo evaluation?
- d. Did you undertake it as an external consultant or an in-house staff member?

2. Your Balkans background.

- a. What was your Balkans background (including language) before the Kosovo evaluation?
- b. How did your Balkans experience, or lack of it, effect your capacity to undertake the evaluation?

3. Your personal background as an evaluator.

- a. Have you at any time attended a course, workshop, or conference in evaluation (particularly humanitarian evaluation), or given a course or similar training on it?
- b. Did you start your evaluation career in economic development evaluation, and if so how, when and where?
- c. Do you have any opinion about the value or otherwise of membership in evaluation societies or the like, and if so, what is this opinion based on?
- d. What do you feel might enhance your skills as an evaluator of humanitarian programmes?

4. Your experience in doing the Kosovo evaluation

a. As regards TOR,

- (i) If you discussed/negotiated these at the outset, what were the key issues you raised, and were they resolved?
- (ii) As the evaluation proceeded, did you consider that further modifications of the TOR were needed, and if so, did you obtain them?
- (iii) In your opinion, what should be the key features of TOR for evaluations of humanitarian programmes? Should there be different TOR for the team as a whole, or for each of the members including the team leader?

b. Relations with client(s) throughout the evaluation

- (i) What relations did you have with the client(s) in relation to their management of the evaluation? Please illustrate how this effected your work.
- (ii) How much time did you spend with the client(s) both at head office and in the field?
- (iii) Ideally, what form would you recommend that relations with the client(s) should take?

c. As regards your method of work/methodology.

- (i) Besides what you have recorded about this in your evaluation report, is there anything you would wish to add here?

d. When finalising and presenting the final report to your client(s).

- (i) How did the finalisation and presentation process go? Did you present your evaluation on time or was an extension sought and granted?
- (ii) What were your impressions of the client(s)' reaction to the evaluation? For instance, did you feel that it was taken seriously and that it would make a difference?
- (iii) How and when do you feel the evaluator's involvement in an evaluation should end? Does it end with the submission of the report, or are further input and/or services necessary or desirable?
- (iv) What, if anything, do you do with an evaluation once you have submitted it? (e.g. publish in some form, use it as training course material, etc.)

5. How do you think the evaluation report should be judged?

- a. Given how difficult it can be to arrive at firm factual findings, evidence, conclusions, and recommendations, do you think each of these parts of a final report should be judged separately? If so, against which criteria?
- b. What aspects of "evaluation quality control" do you consider to be important, and how would you propose they should be tackled, by yourself and by others?

6. Recruitment.

- a. If you were an external consultant, were you recruited by the client(s) directly, or through advertisement and/or tender?
- b. What in your experience has been the most effective approach to recruitment?

ANNEX 4. ALNAP Questionnaire to Commissioning Agencies

1. TOR

- a. What were the objectives set by the TOR, how were these decided, and how long did the process take?
- b. Have lessons emerged this time that will influence how you would draw up TOR next time, for a similar humanitarian evaluation? Please illustrate.

2. Team composition.

- a. What type of team did you originally aim to get?
- b. If impossible to get, what kind of team did you have to settle for?
- c. Next time for a similar type of evaluation, would you go about this task differently, and if so how?
- d. Was the Kosovo evaluation by external consultants only, a mixed team of internal and external members, or an internal team?

3. Team recruiting.

- a. Was recruitment of external consultant(s) by search or tender, and how were in-house staff selected to work on or with the team? Why was this particular approach taken?
- b. In hindsight, would you adopt a different approach?

4. Relations with the evaluators throughout the assignment, including finalising their report.

- a. What was the nature of your relations with the team once it had been recruited? Please illustrate, outlining how particular aspects affected the work undertaken.
- b. What part did you play in the evaluation process, in particular in the final stages?

5. Relationship with the evaluation team after they had submitted their final report.

- a. Do you consider your relations with the team to have ended (in the case of external consultants), as regards this particular assignment, once it had submitted the final report?
- b. If not, what further input and/or services did you require from the team?

6. Assessment of the product once a draft had been delivered.

- a. What was the commenting and vetting process used by the commissioning agency in the finalising of the report?
- b. Specifically, what "quality control" criteria were uppermost?
- c. Ideally, what would you propose as good practice in this process?

7. Action taken since the submission of the final report.

- a. What processes were used in considering whether and how to take action in response to the evaluation? Do you consider these were adequate? If not, please illustrate.
- b. Did these result in action being taken, and if so what action?
- c. If no action was taken, why not?

8. Good practice for managing evaluations.

- a. Has this evaluation led to your learning new lessons (or confirming old) as to how to manage humanitarian evaluations?
- b. What changes would you like to see, internally or externally, to improve the management of humanitarian evaluation?

9. Cost of the evaluation

- a. What was the financial cost of the evaluation?
- b. How does this amount compare to the cost of the programme being evaluated? (Under 1% of its budget, 1-2%, 2-3%, 3-4%, 4-5%, 5-6%, 6-10%?)
- c. In your view, should the proportion of funding spent on evaluation be changed?

10. Functional responsibilities of the commissioning section within the organisation

- a. Were you responsible for the commissioning as well as the management of the evaluation?
- b. If not, did the person(s) responsible for commissioning and managing the evaluation work in/or represent an evaluation, a policy, a strategic planning, an operations, or another department in your organisation?
- c. Is that department single-purpose (as a evaluation, a policy, a strategic planning, and operations, department, and if so which) or multi-functional?