



EVALUATION IN THE HUMANITARIAN ACTION SECTOR

A REVIEW OF GROUPE URD'S ACTIVITIES 1997-2002

November 2002
François Grünewald

TABLE OF CONTENT

1. INTRODUCTION	2
2. A BRIEF HISTORY	2
3. APPRAISAL OF IDEAS AND PRACTICES	3
3.1. EVALUATION AND RESPONSIBILITY	4
3.2. EVALUATION AND TRAINING	4
3.3. EVALUATION, CRITERIA, STANDARDS AND INDICATORS	6
3.3.1. Criteria, standards, indicators	6
3.2.2. Achievement criteria	7
3.2.3. Management criteria	7
3.2.4. Impact criteria	7
3.2.4. Situation criteria	8
3.2.5. Collecting and processing data	8
3.4. INTERSECTORAL AND TRANSVERAL	8
3.5. EVALUATION AND ITS POSITION IN THE PROJECT CYCLE	9
3.6. EVALUATOR INDEPENDANCE AND/OR USEFUL EVALUATION	10
4. MAIN AREAS OF EVALUATION WORK	11
4.1. AT THE HEART OF THE PROJECT FOR QUALITY: REMS AND QUALITY METHODOLOGY	11
4.2. BASIC PRINCIPALS OF THE 'REMS' APPROACH	11
4.3. THE QUALITY METHODOLOGY	12
5. CIRCULATION OF ACQUIRED KNOWLEDGE AND TRAINING	13
6. CONCLUSION	13
GENERAL PRESENTATION	16
CONCEPTS	16
TECHNIQUES	16
EXAMPLES	17

*« Il ne suffit pas de faire le bien,
Il faut le bien faire.
Et le prouver »*

*"It is not enough to do the right thing.
You have to do the right thing well.
And prove it."*

*Diderot
and Groupe URD*

1. INTRODUCTION

Groupe URD's main objective to participate in improving humanitarian practice has been written into its statutes since the group was created. This commitment takes many forms, amongst which:

- regular presence in the sector of evaluation of humanitarian action (over the past four years we have participated in over 20 evaluation programmes in four continents, in a variety of contexts including conflict zones, natural disasters and post-crisis situations),
- development of needs assessment, monitoring and evaluation techniques,
- distribution of these techniques and training on how to use them.

Groupe URD is involved in numerous international and national debates on this subject. Publications in French¹ and English, a regularly updated website www.urd.org where all our research documents can be consulted, four to six training sessions on evaluation per year, and the development and implementation of the 'Quality Project' with its investigative missions in Central America, Afghanistan, Africa, etc. are all proof of this commitment. Similarly, Groupe URD has played a major role in setting up evaluative research for the 'Global Study on the participation of affected populations in humanitarian action', with evaluation work carried out in Sri Lanka, Colombia, Afghanistan, the Democratic Republic of Congo and Angola.

Evaluation practice methods are central to Groupe URD's research on quality and improving humanitarian action.

2. A BRIEF HISTORY

Evaluation practice theory first emerged and was initially implemented more than 50 years ago, and, today, a 'qualitative approach'² has become a primary concern for almost all organisations in the goods and services sectors, and in more recent years, for humanitarian NGOs. The complex topics of evaluation and the quality of humanitarian action have recently become central themes of numerous international debates with the aim of improving the quality of humanitarian action and of illustrating how the responsibility lies with all aid actors. It is of course dependant on many factors, including the politics of individual countries, of donor agencies, of institutions and of teams in the field, the mobilisation of financial, material and human resources, which is itself dependant on politics, the competence of actors at all levels, from top management to logisticians and also complex training processes. As a result of all these factors evaluation practice is still a difficult subject to define.

During the second half of 2001, no less than four conferences on this theme and other related subjects were held in Europe and the USA. The debate about quality is often linked to concerns about who should take responsibility for public funds, for networks that divert private donations from their true destination and for recipients/beneficiaries of humanitarian aid.

The highlights of the last decade in terms of the question of quality of humanitarian aid are presented below.

1992 Dutch government requests an evaluation on Somalia
Creation of ECHO. Procedures include Quality guidelines.

¹ Our most recent publication on this subject is "Evaluating humanitarian aid, practitioners' views", ed. Karthala, as part of the "Appraisal of humanitarian practice" Collection by Groupe URD.

² See texts on the Groupe URD website www.urd.org on the subject of the Quality of humanitarian aid, 1994, the review of the seminar on the initial activities of the 'Quality Project' for 2002, as well as our brochure "Let's talk Quality: a discussion about methodology", Groupe URD, November 2002

- 1994 Multi-donor evaluation in Rwanda. This leads to eleven recommendations of which only two relate to humanitarian action. The remaining nine focus on the work environment and in particular the role of politics, of the United Nations, prevention strategies, etc.
- 1995 Red Cross/NGO Code of Conduct for disaster situations
First international seminar on evaluation organised by MSF Holland
- 1996 "Humanitarian evaluation" sector emerges
- 1997 SPHERE Project. 'The Humanitarian Charter and the Minimum Universal Standards'
- 1998 Ombudsman Project. Application controls for SPHERE standards
Francophone and Latin-American NGOs launch an alternative movement.
ALNAP network (Active Learning Network for Accountability and Performance). International network for evaluation of humanitarian action.
Creation of Evaluation Unit within the ICRC
- 1999 SPHERE expands its activities in lobbying, training and publications
ETIKUMA Seminar in Lyon brings together members of the French government, VOICE and NGOs to discuss codes of conduct and standards
Increasing number of evaluation studies in the field
- 2000 Creation of Quality Platform (32 signatories) and Quality Project
- 2001 Quality Project is launched
Humanitarian Accountability Project (HAP), follows on from Ombudsman Project
"Being good at doing good" conference in the Haye. Four approaches to ensure quality where identified.
ALNAP conference in Washington leads to a change in approach. The focus is no longer on standards but rather on the term "quality"
Meeting held by ECHO partners and creation of Quality Group within ECHO
- 2002 Launch of SPHERE evaluation
Continuing development of Quality Project with the launch of an evaluation process in Afghanistan
Conference held by ECHO and partners on the subject of quality
Groupe URD nominated for the *Global Study on the participation of affected populations in humanitarian action*

3. APPRAISAL OF IDEAS AND PRACTICES

It goes without saying that all actors involved in international solidarity are concerned by the quality of their intervention, yet evaluation practices are relatively new to the humanitarian sector, that is to say barely ten years. A whole series of initiatives, some of which are fairly controversial, has been implemented, including the Code of Conduct of NGOs and the Red Cross for operations in disaster situations, the SPHERE project for Humanitarian Charter and universal minimum standards, the Humanitarian Ombudsman project, since renamed the Humanitarian Accountability Project (HAP), to name but a few. A significant minority of French NGOs rallied round the 'Quality Platform' which boasted members from more than thirty associations in over ten countries, including representatives from NGOs, United Nations agencies and the main donor agencies. This collective mobilisation then entered a second phase, focussing on the development of the 'Quality Project'. This project has a strictly proposal approach, based on theoretical reflection and experimental research, and takes into account the complex nature of quality issues and the stakes involved in the highly specific context of humanitarian action.

The 'Quality Project' currently receives input from roughly fifteen NGOs with Groupe URD playing an organisational role. Its objective is to improve the quality of humanitarian action and the service supplied to its beneficiaries. This is to be achieved via the development of methods and techniques, inspired by the Quality Assurance approach, which can be applied at all phases in the project cycle: needs assessment, intervention design, implementation and monitoring, and evaluation.

As evaluation in the humanitarian action sector has become very active over the past few years, so too has the content of its debates evolved enormously. Groupe URD is currently involved in activities that revolve around five main axes:

- the link between evaluation and responsibility (of which accountability is only one component);
- the link between evaluation and training;
- criteria, standards and indicators;
- the role and the place of evaluation within the project cycle;
- the independence of evaluation teams.

3.1. EVALUATION AND RESPONSIBILITY

The topic of *accountability* is complicated and a relatively new concept. There are inevitable divisions over the issue of vocabulary and translation, as the term 'accountability', or the responsibility to account for one's actions, has recently been translated in French by the word 'redevabilité' (to be indebted) which does not necessarily evoke the same idea. Accountability mainly refers to financial and accounting aspects, as well as administrative and legal procedures; i.e. whether contract conditions have been respected. At this stage, as indicated in Annexe 1 prepared by Groupe URD for the 'Quality' working group of the International Federation of the Red Cross, it appears that the term responsibility is richer than that. In a document prepared for a seminar organised by MSF Holland³ in 1995, it is indicated that humanitarian agencies have a responsibility:

- firstly towards those who benefit from their work and for whom they mobilise resources,
- and secondly, towards those who give them the resources that enable them to work.

Thus, accountability can be broken down into *upwards accountability*, i.e. towards donor agencies, and *downwards accountability*, i.e. towards the beneficiaries. However, many questions remain answered, such as "To what extent are humanitarian actors indebted or accountable to beneficiaries?" More recently, on the back of work such as the "People In Aid" project, the idea that we also have a responsibility towards our own teams has also been added to the list.

The means of expressing this responsibility should not be seen merely in terms of accountability. The issue of responsibility, both for the populations affected by the crisis and for the donors (whether private or public), concerns firstly the quality of the assistance and of its **impact** and secondly, how it is accounted for. Certain management tools concentrate on the question of accountability, such as annual financial reports certified by auditors, audits carried out by the donors, commitments made to the Charter Committee, etc. Traditional methods of evaluating humanitarian assistance regard accountability as a means to evaluate compliance, i.e. whether the terms of the contract have been respected. This is the position taken by the SPHERE project.

But can other tools contribute to improving our management of the ultimate responsibility, that of providing the best possible assistance and protection for beneficiaries? Our work over the past five years has prompted us to place the role of evaluation in general and the qualitative approach that we would like to promote, at the heart of our quality method, which itself results from the "Quality Insurance" approach. Furthermore, within the framework of the Global Study of Consultation and participation, we are also very attached to the idea of providing the beneficiaries with a sense of satisfaction and of giving a voice to the populations in distress.

Responsibility towards team members revolves around issues such as fair treatment, transparency, a collective training strategy, payment policy, promotion opportunities, etc. This is an important subject but one where Groupe URD has decided to remain on the periphery, by keeping a watch over developments in this area and by organising regular training sessions on a number of related subjects, in particular team security and crisis management, the relationship between emergency and development, pre- and post-crisis evaluation, food security, etc.

Groupe URD is presently working on two important projects on evaluation and on how to capitalise on humanitarian practice in terms of the quality of their intervention:

- 'Quality Project', with the aim of developing a Quality Method;
- Global Study of the consultation and participation of affected populations and beneficiaries in humanitarian assistance.

It is hoped that, via this work, we will be able to contribute to the debate about responsibility.

3.2. EVALUATION AND TRAINING

The use of evaluation as a tool for training and in order to promote institutional change is a key issue for those responsible for evaluation services in many institutions. In fact, all too often, evaluation is seen as a means of control, and thus something to be suspicious of, before being considered as an activity enabling

³ "You said evaluation", 1995, François Grünewald. Document prepared for the seminar "Evaluation of international humanitarian assistance", Amsterdam, available on the Groupe URD website www.urd.org.

progress. But for many missions, the report is often drafted too late once the programme has ended, it is relegated to a shelf and does not incite any feedback, leaving one very sceptical about the purpose of the exercise. Whereas we should be discussing the priority of “evaluating in order to evolve”, we are preoccupied with the question of “evaluation for control purposes” or “evaluation to provide a witness”. Again, there seems to be no connection between lessons learnt and the resulting action taken which explains the necessity to focus on bringing evaluation results and the action itself closer together. Aside from this observation, several initiatives have emerged.

- Real time evaluation This exercise closely resembles monitoring, except that it is usually carried out by a team which is not involved in the programme, often a team from a regional office or from head office. The first attempts at real time evaluation were carried out by the HCR. It generally involves evaluative visits to the field when the programme is underway, so that any necessary adjustments in operational strategy can be made as a result.
- Regular evaluation with mini seminars (REMS) (EIMS - Les évaluations itératives avec mini-séminaires) This process was developed and tested by Groupe URD from November 1998 to March 2002 with the aim of drawing some lessons from the humanitarian aid following the Mitch hurricane and the earthquake in Salvador. It is a complex process linking exchanges, co-ordinating interviews, sending evaluation missions in the field and giving feedback to local NGOs and international actors. The final phase included a series of workshops and conferences in France, Switzerland, USA and the publication of several articles in specialised press, and a certain number of ideas were drawn up concerning improving training processes.
- A programme combining evaluation and research Groupe URD used this approach in the Albania-Kosovo crisis⁴. As of spring 1999, real time evaluation was assessed by an evaluation team in Albania and research carried out over a longer time-scale over the summer 2000. At each stage, a feedback seminar was held. Thus, after the first mission in the field, a series of debates was organised with interested NGOs. After the second mission, a mini-seminar was organised in Tirana with Albanian and international NGOs, while a larger event was later organised in Paris. The Global Study on consultation and participation will also establish itself in this category of evaluative research.
- Learning Office The Learning Office innovation was part of the ALNAP network (Active Learning Network for Accountability and Performance). The feasibility of this process had been examined extensively in several concrete examples, India with the Orissa hurricane, Timor, Sierra Leone, Kosovo). This work resulted in a proposal, which, following a long debate and varying hypotheses, led to a pilot project in Malawi in 2002, in the context of the food crisis of southern Africa. One of the main stakes in this exercise is to see how the Learning office can be put in place in parallel with other institutions, such as the OCHA, whilst avoiding duplication.
- Capitalising on the learning process Collective work within the ALNAP framework resulted in strategic reflection on the problems of evaluation for learning. It shows that the problems linked to learning have by in large been put aside, so great is the pressure from donor agencies in favour of accountability.
- Post action review The international network World Vision has recently undertaken a very promising initiative. This learning tool of ‘post action review’ or post operation review’ has its origins in military practice. World Vision organises, in the weeks or month after the start of an operation, an ‘on-the-spot workshop’ in order to assess any errors made and lessons learnt. The results are re-injected into the operation underway but also into the institutional dynamics upstream.
- The Quality Method At the heart of the ‘Quality Project’, implemented by the Groupe URD, can be found an approach in which evaluation plays a central role, the Quality Method. Resulting from work on quality insurance, this method proposes an evaluative approach lasting the duration of the action, with an initial emphasis on detecting and managing ‘critical points’. The second element of this method is taking systematic recourse to evaluation: repeat self-evaluation and evaluation by others at certain times.

Though and exchange of these experiences continue. The aim is that evaluation reports do not gather dust on shelves, their lessons unheeded. Evaluation for advancement.

⁴ This approach is written into a research and evaluation agreement between Groupe URD and Coordination SUD, a network of NGOs involved international solidarity, based in France.

3.3. EVALUATION, CRITERIA, STANDARDS AND INDICATORS

3.3.1. Criteria, standards, indicators

The debate about criteria, standards and indicators is an issue which is central to needs assessment and monitoring practices, and to the evaluation of humanitarian aid. The exact meaning and the use of these terms is often subject to great confusion.

We refer to criteria in order to express an opinion or our approval in terms of an objective. A criterion has two components:

- a norm or a standard can be used to indicate the desired value of the criterion, the performance level to aim for. The definition of a norm is often contextual.
- an indicator is the variable by which the criterion is measured and allows us to indicate the difference between the standard and the observed situation.

Donor agencies demand evidence of criteria and standards, and obviously indicators to show that the standards have been attained. At this point, one has to be careful since certain methodology gimmicks (such as the use one single criterion to qualify all elements of a programme), methods often invented by epidemiologists, and the desire or the need for a scientific appearance to grant legitimacy, can have a dramatic effect. For example, donor agencies, under the aegis of Colombia University, USA, are currently debating the use of the crude mortality rate as a global indicator for assessing the impact of humanitarian aid, illustrating the absurd lengths to which all this can lead. Groupe URD maintains a critical watch over this dynamic.

Researchers working in the sector have identified several main issues to which criteria can be applied. The most well known are: effectiveness, efficiency, pertinence and impact. Others have since been added to the list: connecting or continuity between emergency aid and development, the influence of aid interventions on conflict dynamics, on social dynamics, in particular the gender issue, etc.

These criteria must also have a certain number of characteristics. On the one hand, there are the SMART criteria: sensible, measurable, adapted, reproducible and timely.

However, less relevant criteria have also been put forward. In general, criteria should make sense. It may seem obvious to state this, yet so often, criteria are employed which are based on ratios whose common denominator is tainted by a large margin of error. So, for example, participative dynamics should be selected with the population concerned. Furthermore, criteria should be specific and easy to use. Measuring and processing them should not be a costly procedure. They can be qualitative as well as quantitative. Humanitarian aid has a tendency to drown itself in huge quantities of data that is collected but never processed, and often it is not even possible to process. Where a population is threatened, so too do we find huge quantities of buried data. There is an increasing tendency for volunteers and evaluators to spend day and night in front of their computer screens with very little contact with the population supposedly at the centre of the action.

Below, we highlight the main limitations for the use of criteria. Essentially, they can be used to assess phenomena that are both measurable and expected.

Criteria	Measurable phenomena	Immeasurable phenomena
Effects/Impact Expected	INDICATORS CAN BE USED	
Unexpected		

The four main sets of criteria are outlined below with their use, limits and validity.

- achievement;
- management;
- impact;
- situation.

3.2.2. Achievement criteria

These are most frequently employed by donor agencies for evaluating humanitarian aid in order to check whether “we have achieved what we said we have achieved”. As humanitarian aid is often relatively simple logistically and theoretically implemented over a short time span (compared with the pace of development projects), the use of these criteria depends upon three hypotheses.

- Causal hypothesis This assumes that achievement indicators can be considered as proxi-indicators. One does not measure the impact a programme to build wells has on public health, but a hypothesis is made, based on a fact, that such a programme will have an effect. Hence, it is enough to check that the wells are well built, within the specified time span and allocated budget.
- Non-visibility hypothesis This implies the position taken by NGOs whose business verges on dishonesty, playing with taxpayers’ money. Therefore, controls must be established to ensure that contracts are respected.
- Accountability obligation hypothesis In theory, a public body is obliged to account for taxpayers’ money. In this scenario, it is believed that taxpayers want information on how their taxes are spent, rather than on the impact of these achievements.

Compliance criteria can also be included in this section. They can be seen to measure the difference between the contractual agreement and what is achieved in reality. When this type of criteria is employed, participative evaluation is rarely practised and the evaluation exercise is managed in a controlling fashion, very similar to an typical financial audit.

3.2.3. Management criteria

These are criteria which should above all examine organisational procedures: management of financial department and human resources, communication strategy, decision making procedures, etc. References to this type of methodology can be found in the large body of texts written about the theory of organisations. Groupe URD is not involved in this issue unless its impact on aid and the quality of aid is threatened. However, two Swiss NGOs, MEDAIR and TDH have carried out an interesting piece of work in this field, testing the use of ISO standards in their institutions. We will continue to follow this approach.

3.2.4. Impact criteria

The total impact of an intervention includes those effects that are expected and unexpected, direct and indirect. These can be positive or negative, they can be immediately obvious or alternatively take a while to become apparent. And this is where things become complicated. On the one hand, criteria assessing the quality of humanitarian intervention and indicators used to measure them can be different for actions that are apparently similar, depending on whether the context is of rapid-onset emergency, extreme emergency or post crisis.

However, on evaluating a programme with a varying pace, it appears that an element that could have been considered characteristic of quality at a given time can become synonymous of inappropriate adaptation at another time. Finally, understanding an intervention’s impact requires in theory to be able to compare the situation **with and without the programme**, whereas at best, for ethical and practical reasons, the evaluator can only reasonably compare the situation **before and after the intervention**.

In-depth reflection on causality approach and related proxi-indicators is indispensable at this point, but must be treated with caution. We must bear in mind that an underlying causal approach is effectively based on the mere probability that an action X has an effect on Y.

3.2.4. Situation criteria

In general, humanitarian aid takes place in contexts that are highly volatile, sometimes even dangerous. It is important that any possible contextual changes are routinely examined for the duration of the programme and this is undertaken by an important part of all monitoring procedures. But the evaluation process itself must also be capable of assessing the situation and observing any changes from when the initial needs assessment was drawn up. Indeed it is this comparison alone which enables us to fully understand the effects of humanitarian intervention. Above all, what is involved is trying to avoid attributing positive or negative factors to the humanitarian programme, when they are rather results of a change in the situation, completely removed from the programme itself.

Evaluation work must also attempt to understand the reciprocal and in general complex interactions between the aid carried out and the developing situation. Various projects have attempted to define an appropriate methodology and criteria. We would simply like to refer to: firstly, the project by Mary Anderson "Local Capacity for Peace", from whence the henceforth famous debate on Hippocrates' maxim "First, do no harm" originates; and secondly, the project, which is currently underway, known as "Peace and Conflict Analysis". Nevertheless one must note that for the analysis of situation criteria, the evaluator must take on a position of optimum ignorance (in the Chambers sense of the word) and be modest.

It is in this context that Groupe URD, in liaison with other researchers, is reviewing the Log Frame. Indeed it is due to the fluctuating nature of the humanitarian aid context that the multi-scenario approach is still being applied inappropriately. This approach, which leads to the definition of several plausible courses of events, inevitably implies that the last column of the Log Frame (hypotheses, risks and constraints) should in fact take second place. From this point, we can envisage various scenarios and define specific criteria indicating "when to move from Plan A to Plan B".

3.2.5. Collecting and processing data

Attention must also be paid to the manner in which data is collected and processed. The influence that the epidemiology sector has had on humanitarian aid means that often our social sciences expertise is forgotten. The crisis-affected populations are not "homogenous groups of people in which variables are distributed randomly", but stratified systems that can be differentiated geographically, socially and economically. In the majority of cases, the variables under analysis do not follow a normal even distribution. However, often, humanitarian actors do not think twice about collecting and processing data that is neither following a normal distribution course, are seldom continuous, or discrete, and calculate averages and standard deviations as if they were. Too much time is wasted filling collating data whose pertinence is dubious, the collecting process is tainted with numerous bias, and processing time consuming and costly. Errors are also made when numerators is controlled by denominators very approximate.

3.4. INTERSECTORAL AND TRANSVERAL

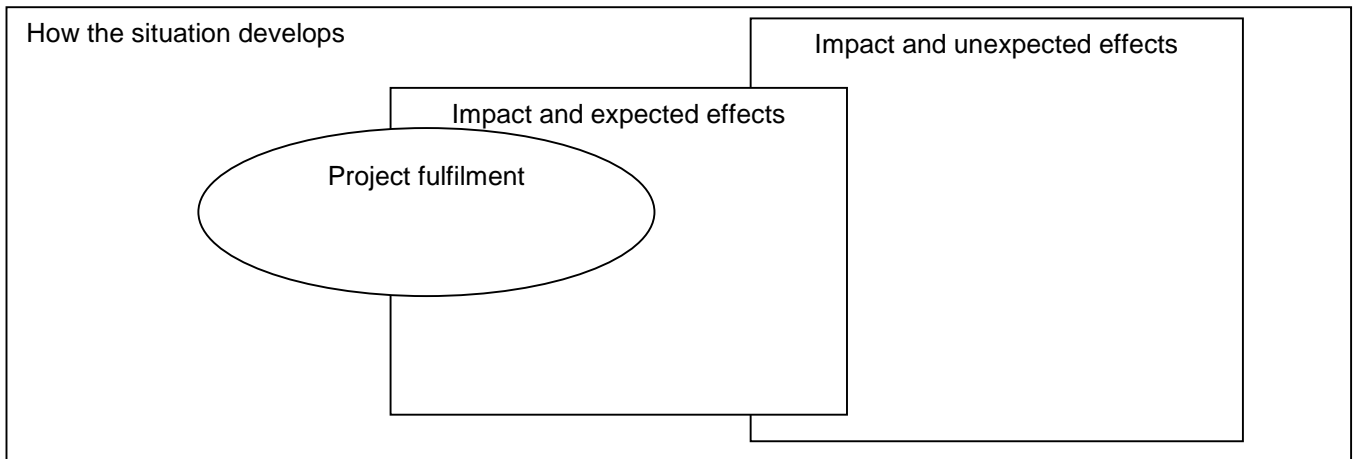
If humanitarian interventions are often sectoral in nature in terms of mandates and the experience of aid agencies, the needs of populations in distress are holistic. But even when a specific programme within a precise technical domain is being evaluated, a full understanding of the impact often requires applying a combination of disciplinary fields simultaneously. This is an approach whose operational implementation requires investing in a truly multidisciplinary team, whereby both pure science (health, agriculture, construction, etc.) and social sciences (communication, management) work in harmony. Multi-sectoral evaluation practices, thanks to interdisciplinary work, have now become a skill in which Groupe URD is highly referenced. The table below defines the current areas of investigation.

Combining sectoral fields with transversal issues

Sectors	Health		Economic security		Human establishment		Protection		Education	
	Health	Water and water treatment	Food aid	Agricultural rehabilitation	Habitat	Urbanism	IHL	Human Rights	Primary	Other
Pertinence										
Impact on populations										
Impact on social cohesion										
Impact on security and how the conflict develops										
Connecting: links between emergency and development										
Viability and durability										
Effectiveness										
Efficiency										
Participation, consultation and partnerships										
Coordination										
Population movement										
Gender										
Environment										

3.5. EVALUATION AND ITS POSITION IN THE PROJECT CYCLE

Evaluation practice in a humanitarian context shows just how dangerous and unproductive it can be to separate evaluation from the rest of project cycle management. It is at the beginning of the cycle, during the needs assessment and conception phase, that a monitoring and evaluation strategy should be drawn up, and a certain amount of consideration given to criteria and indicators. In addition, that humanitarian aid evolves within fluctuating and volatile contexts, it is not rare to see programmes changing dramatically, in so much as for their objectives, as their strategy and operational implementation. Occasionally the whole range of criteria has to be partially amended or even completely reviewed. This issue has notably led to a reappraisal of the relevance of the now mechanical use of the Log Frame.



3.6. EVALUATOR INDEPENDENCE AND/OR USEFUL EVALUATION

This subject is quite controversial as it is a sensitive topic and it touches on the credibility of evaluation as much as on its impact in terms of a practice that can induce change. There are presently four schools of thought based on the following principals.

- the importance of evaluator independence;
- the belief that in the end, what counts is the willingness to participate in evaluation practices and exercising responsibility, rather than formal independence;
- the development of a participative approach to evaluation;
- finding the best combination of the above approaches.

However experience shows that, in practice, it is not all that simple. In general, independent evaluators are tied to a market, that of evaluation. Evaluation work is sometimes extremely laborious, as much for those being evaluated as for those doing the evaluation. This explains the reticence of so-called independent evaluators to carry their investigations to their natural conclusion. In addition, external evaluators are often criticised for not fully understanding the context and for the way in which the institution functions. All the same, once the report has been made, the external evaluator is not obliged to fulfil any "after-sales service". The result of this is that the evaluation exercise does not give rise to enough real benefit and increases the number of reports gathering dust on shelves.

The turn of events can be completely different when the evaluator is truly motivated by the idea of "evaluation for advancement". In this case the evaluation can be extremely thorough, directed by a good knowledge of the institutional culture and see results leading to concrete developments. However, if the internal evaluator feels that his own advancement within the organisation might be linked to any observations he might make, if he abuses his position in order to settle personal disputes, then evaluation work may have a catastrophic impact on the structure.

Participative evaluation work is not common in the field of humanitarian aid. The working culture of humanitarian actors tends to view the populations with whom they work as passive victims, or even worse as dishonest people continually trying to get more aid via illegal means. Participative evaluation work in a crisis environment must always take into account the difficulties and sometimes even dangers that are inherent to participation, and in particular any action that may engender security problems for both affected populations and for evaluators.

Groupe URD has developed a multi-approach strategy based on:

- traditional evaluations commissioned by donor agencies or international institutions. The European Union, the Food and Agricultural Organisation of the United Nations (FAO), the WFO, etc. have regularly called on the expertise of Groupe URD's personnel.
- inter-NGO evaluations as a support for learning processes. It is important to note here that the key is to draw a lesson from experience and to share this information collectively and not to give good or bad ratings to individual actors. Reports and presentations are drawn up with this end in mind. However, at the request of certain organisations, more detailed and precise reconstruction sessions can be organised 'behind closed doors'. This is the precedent of the REMS methodology.

- 'evaluation – research' combinations, in order quickly uncover any important issues that were revealed during the evaluation and to put forward propositions.
- 'capacity to trigger an evaluation on operations or topics that seem important to build upon. This self-assessment then becomes a project requiring financing. In the past, when a subject or situation seemed to be of importance and ad-hoc financing did not exist, Groupe URD has taken the decision to draw on own resources.
- 'Quality methodology', which, as mentioned previously, is coupled to the project cycle. Based on self-evaluation and self-learning processes, this methodology assists in the identification and management of critical moments.

Independent, yes, but capable of empathising! For Groupe URD the key to evaluation is to be able contribute to improving humanitarian practice and not to hand out good or bad marks. Groupe URD represents four schools of thought: independent; support; usefulness, where possible via the reconstruction workshops as part of our working method, and to initiate learning.

4. MAIN AREAS OF EVALUATION WORK

4.1. AT THE HEART OF THE PROJECT FOR QUALITY: REMS AND QUALITY METHODOLOGY

On the basis of research undertaken for the Project for Quality and previous experience, Groupe URD is expanding its work in this field, by concentrating on two domains:

- developing a 'Quality tool', similar to 'Quality Insurance';
- regular evaluation with mini seminars (REMS).

The results of our research are regularly presented during international conferences and have begun to spark a great deal of interest.

	REMS	Method for Quality
Why?	Learning	Project monitoring / Assessment / Learning
What is involved?	Lessons to be learnt	Monitoring a project from A to Z
Who?	External evaluation	Internal evaluation: self-evaluation
Where?	Several NGOs with a common subject	Each individual NGO
When?	Three separate visits	More or less continuous process for the duration of the project cycle
How?	Building on experience and learning inter NGOs, with the support of an external team	Resources within an NGO

4.2. BASIC PRINCIPALS OF THE 'REMS' APPROACH

This approach is based on eight main principals:

- Real time evaluation. This exercise is carried out by evaluators from outside the project and enables those who are caught up in daily running of the project and who are not necessarily able to distance themselves from events, to benefit from a constructive viewpoint, and if necessary to reposition action. In this way, actors may benefit from information gathered in other similar situations and the question of exit strategies can be tackled almost immediately.
- Various monitoring exercises and the Learning Office approach. This approach requires a presence on a daily basis and the evaluation mission should be a high point, a moment to breath new life into the field teams. They must therefore include a key event in the field, in the form of a reconstruction workshop and an off-the-cuff debate. At this point, it is possible to take stock of certain constraints, but also reformulate questions for which head-office must find a solution. These mini seminars are often moments of great creativity, as everyone becomes aware that others are asking the same questions, but it is possible to identify, thanks to collective discussion, some original solutions.

- On return from the field, an on the spot reconstruction workshop must be held for head-office personnel, followed by discussion based on the assignment report. This structured discussion, looking at the main questions raised by the report and any eventual subjects of disagreement, is vital to the institutional learning process.
- The process must be put in place early in course of the crisis response, in order to benefit from a range of observations on the crisis itself, crisis management, the role of the actors involved, the constraints, etc. that is as complete as possible. Initiating the process early and sustaining its activity over a long period, ensures that facts, problems encountered and constraints are not forgotten, and that a true learning process can commence.
- The team should remain the same for the duration of the operation and it should be relatively small, albeit multidisciplinary. A group of three to four people is an ideal number in terms of cost, support for field teams, mobility, etc. All team members should be experienced, both in specific sectors and in the field of evaluation. Finally the team must have good communication skills and confirmed pedagogical experience.
- A standby arrangement should be developed so that a team can be mobilised quickly. The Groupe URD team is a good example of a team that is multi-disciplinary yet easy to mobilise.
- The team should have access to a resource centre and be able to research information from a distance/online on which to base its preparation work prior to departure, and question and answer investigation during the assignment. Again, Groupe URD's resource centre is fully equipped and is already running smoothly for such exercises.
- The methodology must be easy to reproduce and thus relatively inexpensive. It is a matter of principal that evaluation operations should not be unwieldy in terms of resources to be mobilised, and a question of reproduction, that the methodology can be applied to the majority of major crises. Thus the cost of opportunity/appropriateness for this kind of evaluative methodology must be attractive.

An evaluation procedure must itself be able to undergo evaluation. So that the final evaluation of the impact of this exercise might be attained, a log frame has been devised to establish objectives and hypotheses.

Level	Description	Indicators	Means of verification	Hypotheses and constraints
Main objective	To develop a simple methodology for fast learning with the aim of improving humanitarian practices	Is REMS methodology introduced at the outset of the crisis preparation phase? Do donor agencies budget for evaluation procedures?	Reports Interviews Social audit	Interest shown by actors Interest shown by donor agencies
Specific objective	To test the 'regular evaluation with mini seminars' (REMS) methodology during a particular crisis, in order to compare it with other approaches, as part of the ALNAP project	Are programmes capable of adapting rapidly? Do actors request information on the experience of others? Has a cost-effectiveness analysis been carried out?	Reports Interviews	Interest shown by actors Interest shown by donor agencies Available time
Activities	3 assignments 3 mini seminars Publications	Number and timing of assignments Number and timing of mini seminars Number of participants Number of publications	Reports Interviews	Available time Security Available resources
Resources	3-4 specialists Documents and publications Simple logistics Pedagogical material Workplace for seminars		Reports Interviews	Quality of preparation work for mobilising resources Available finances

4.3. THE QUALITY METHODOLOGY

Following a brief description above, this methodology was the reference of an ad-hoc brochure titled « You said Quality – contribution to a debate », which is available from Groupe URD and other partners involved in the Project for Quality.

5. CIRCULATION OF ACQUIRED KNOWLEDGE AND TRAINING

In our opinion, the ultimate step of the post-evaluation service can be broken down into two activities, building on and circulating knowledge and training.

Thus Groupe URD regularly publishes the results of evaluation research and methodology research in specialist reviews and has written a collection of publications entitled « Humanitarian practice », ed. Karthala. Groupe URD also organises and participates in numerous workshops and conferences both in France and internationally, the most recent of which took place in London, Washington, Delhi, The Haye, Geneva and Paris.

In terms of training, Groupe URD is equipped with the following resources:

- A resource and training centre, capacity 20-25 people.
- A website which is regularly updated with the results of the most recent fieldwork.
- A series of training modules, which can be part of a diploma qualification (therefore upstream in the human resources management system), or as part of a scheme requested by an organisation. Training modules include on the one hand evaluation and evaluation techniques, and on the other, modules on a variety of topics which are regularly updated and added to thanks to the results of evaluation research and lessons learnt during assignments.

Annexe 2 presents one of the evaluation modules available at the Groupe URD training centre.

Within this domain, Groupe URD teaches over 120 days of training per year in Paris, Lyon, Geneva, Lausanne, Pisa, Aix en Provence, Grenoble, Nantes, Santander and Brussels, etc.

6. CONCLUSION

For several years now, Groupe URD has carried out in-depth research in the evaluation of humanitarian aid. For any given humanitarian operation, our research aims are threefold: to sift through the evaluation results; to thoroughly investigate all the critical issues of the multiple processes which make up the operation; and in particular to assess the evaluation process itself.

At length, the research undertaken should give rise to methodology proposals with the aim of improving learning processes both during the project cycle and in more global terms, whereby evaluation is a strategic tool to encourage institutional change.

This approach is supported by the debate about humanitarian NGOs quality policy and is characterised by the conviction that:

- The responsibility of humanitarian organisations is twofold: firstly towards the populations whom they wish to help, and secondly towards private donor agencies and institutions that support them.
- In order to do the right thing, generosity and technology are not enough. Despite good intentions, mistakes are still possible.
- The only way of passing judgement on a given action and learning lessons, is to implement a real evaluation programme, one that is both in-depth and systematic at all levels.
- The only way of moving on from the lessons learnt on the back of the evaluation programme is to implement a concrete learning process. Too many evaluation reports do not serve their purpose. This learning process should combine the reconstruction of and building on events and both at field level and head/regional offices.

Groupe URD will continue to organise quality training on the evaluation of humanitarian aid for executives, decision makers and those in charge of evaluation, as well as supplying training modules introducing the evaluation debate and evaluation practices to both entities and institutions that are responsible for training volunteers.

We will continue to build upon experience from the knowledge gained from evaluation assignments and to spread these ideas, to publish the results of our research into evaluation practice methodology, and to participate in exchanging ideas on an international level.

For further information, please consult the following websites: www.urd.org, www.globalstudyparticipation.org, www.qualityproject.urd.org.



THE RESPONSIBILITY⁵ OF HUMANITARIAN ACTORS A COMPLEX CONCEPT IN A NUTSHELL

Véronique de Geoffroy⁶ & François Grünewald⁷
Groupe URD
5th May 2002

The word « responsibility » is one of the strongest words in the French language. It encompasses legal, moral and ethical connotations. There is only one word in French, “responsabilité”, while in English there are three to describe these various facets: liability, accountability and responsibility.

1. VIS-A-VIS INTERNATIONAL HUMANITARIAN LAW

Legal: The term ‘responsibility’ can refer to the obligation to fulfil certain tasks according to established laws and rights. This can also encompass the concept of “duty bearing” or condemnation should the organisation fail to fulfil these obligations.

Breach of International Humanitarian Law (IHL): States, Command chains and individual armed personnel are responsible for ensuring that IHL is respected. If IHL is breached, it is therefore the responsibility of individual States or armed personal to inform and legally condemn.

IHL does not evoke the responsibility of humanitarian organisations and staff in terms of legal implications. It is nevertheless clearly specified that humanitarian actors have responsibilities within the IHL framework:

- as humanitarian actors, they have to respect IHL and its principles (especially the concepts of impartiality and independence);
- as witnesses of war crimes, humanitarian actors have a role in reporting events, at an ad-hoc level;
- as an operator in relief operations, they should ensure that the response best suits the needs.

2. MANAGERIAL RESPONSIBILITY VIS-A-VIS NATIONAL LAWS

As a channel of resources from States and individual donors, humanitarian actors also have various responsibilities, which in most cases are already grounded into their national legal framework. This includes:

- legal responsibility for the management of funds;
- legal responsibility vis-à-vis their staff;
- contractual responsibility in the implementation of projects when contractual agreements have been signed (including respect for stated objectives, budget lines, and timeframe).

⁵ Definitions and concepts are taken from the Grand Robert Dictionary, the Practical Dictionary on IHL, MSF (Françoise Bouchet-Saulnier), the Forum on “War and accountability”; ICRC and from current work on Quality undertaken by the Quality Project/Groupe URD.

⁶ Fellow researcher and IHL adviser, Groupe URD

⁷ Chairman, Groupe URD

Here, liability is the most useful concept. Yet, people can be guilty (“coupable”) of dishonest behaviour or responsible (“responsible”) for failing to implement an action.

3. OBLIGATION TO MOBILISE RESOURCES OR OBLIGATION TO PRODUCE RESULTS

The debate surrounding the “obligation to mobilise resources” versus the “obligation to produce results” is one that is deeply grounded in the medical sector. Indeed, in this field, medical personnel are held responsible for the quality of their diagnosis, for mobilising resources but cannot be held responsible if, in the absence of error and in respect of existing knowledge, the treatment fails.

As our society leans increasingly towards legal action as a solution for all types of disagreements, following the similar trend in North America in previous years, so too do we focus increasingly on our “obligation to produce results”. Humanitarian actors are worried that this trend, through the extension of the concept of “responsibility” towards “liability”, will distort the process. NGO’s are ready to assume full responsibility in terms of “obligation to mobilise resources”. Yet, the complex nature of humanitarian action in turbulent, often dangerous and manipulated situations means that they cannot be held responsible for their results. Often, too many factors are involved which can prevent even the best-resourced and most professionally designed operation from succeeding.

4. QUALITY OF ACTION AS A RESULT OF DUAL RESPONSIBILITY

Responsibility, accountability, liability. These three concepts can be of great use when referring to various aspects of the humanitarian network, as aid is transmitted from the fortunate few from rich countries (donors) to the affected population (beneficiaries). Yet, if the main issue of concern is “What happens at the end of the chain?”, how can these concepts be integrated into the operation?

This is where the importance of “the search and struggle for quality” can be introduced. The search for quality should primarily result from our duty to support the affected population in the most appropriate way. The way in which this idea is expressed is important and it is on purpose that we did not use the phrase “to deliver the best quality assistance”. But donor agencies and taxpayers also have a similar viewpoint. They want the money that is made available to humanitarian actors to have the best possible impact. This includes firstly the appropriateness, effectiveness and efficiency of the programmes is recognised and secondly, that reporting and financial accountability is carried out. However, at present a negative trend is affecting humanitarian aid, which places elements of responsibility related to accountability and liability before the issue of impact. “Being good at doing good⁸” ends up as “being good at respecting contracts”. Beneficiaries do not count anymore.

It is with this idea of “doing the right thing well⁹” that the moral, professional and ethical connotations of responsibility meet the issues related to quality.

⁸ Conference on Quality, The Hague, 2001

⁹ Paper published in ‘After Action Review’ in World Vision

GRUPE URD TRAINING MODULE

FROM NEEDS ASSESSMENT TO IMPACT EVALUATION: CONCEPTS AND METHODS

GENERAL PRESENTATION

While there has been an increasing demand for Programme Quality and donor agencies continue to press for greater transparency, it has taken time for evaluation practices to break through into the field of humanitarian aid. Often confused with auditing, control mechanisms or monitoring individual performance, it is essential that project or programme evaluation become an integral part of the culture of humanitarian actors. This module can be integrated into diploma courses (such as NOAH, various post-graduate diploma, BIOFORCE or IFAID), or as part of a training strategy for NGO, donor agency and international agency staff. There are two available versions of this module, depending on the participants' needs and level.

Evaluation culture A module which aims at making the principals, issues and methods of evaluation more accessible to humanitarian actors so that they feel more comfortable about requesting, accommodating and using evaluations.

Evaluation practice A module for personnel with specialised training who participate in field assignments, although this module is also appropriate for desk managers and operations managers.

CONCEPTS

As evaluation is not a simple subject, this module outlines key issues and concepts before going into practical details of the methodology. The main concepts are:

Evaluation for advancement It is important that evaluation concepts are placed within a process aiming at Quality and institutional change. Although some claim that evaluation already makes up part of the managerial duties and institutional learning processes, there are still far too many reports that remain unread. Thus a great source of learning material is lost.

History of evaluation and Quality concepts At this point in time, when evaluation is central to many initiatives, such as the SPHERE standards, the Quality Platform and the Humanitarian Accountability Project, it is important that participants study the definition of the word 'evaluation' and examine the history of the ideas and concepts, the role of evaluation and the methods involved.

Communication, partnerships, transparency and empathy are key to successful evaluation Carrying out an evaluation involves many steps, in which the personnel and beneficiaries of operations, projects or programmes under evaluation can play a major role. Issues, techniques and social mobilisation.

Implementation, results, effects, amendments, impact What should we be looking at? Why? Which tools should we use? How can the Log Frame be effectively used?

Management of the evaluation process Have evaluations become the new alibis for humanitarian actors or are they a real tool towards a Quality approach? How can the different processes best be managed, from when the evaluation is commissioned, to preparing the Terms of Reference, recruiting staff, implementing the assignment, giving feedback, presenting the results and ensuring that the dual objective of providing a learning tool and providing a comprehensive report is achieved? How should this process be inserted into project cycle?

TECHNIQUES

The modules expand upon various techniques.

The dynamics involved in participation On the interface between communication and social sciences, participative techniques are key to ensuring that affected populations are involved throughout the whole operational process, from needs assessment, to evaluating the impact of the operation and analysing any possible side effects. A well-informed approach to these techniques is imperative and when these tools are managed efficiently, they can enhance operations very effectively.

Statistics Whether in the interest of scientific precision or IT jargon, the use of statistics has become increasingly widespread, notably due to the fact that computers have become more user-friendly, software has become more complex and requests by donor agencies for tables, diagrams and charts have increased. Nevertheless, it is important to be able to compile truly constructive statistics and to avoid falling into traps.

The complexity issue If the idea of the “average” victim is illusory, how should one go about analysing situations, proposing complex responses and then evaluating these complex responses in similar situations? Detailed analysis of typology and zoning, etc. must be included in the scope of reality analysis procedures. Techniques in establishing functional theoretical models and recording complex realities, such as the Differential Vulnerability table, are discussed.

References, indicators, qualitative and quantitative approaches A good evaluation must be based on hard facts. Which information should be included? How much detail? To what degree of precision? To what comparative level?

Specific and thematic techniques Depending on the audience and explicit requirements, it is possible to prepare a programme including more specific techniques. However, a background to social science theory (ethnology, anthropology, economy) and techniques employed in communication studies (role play, how to conduct meetings and interviews, etc.) will be available to all participants.

EXAMPLES

So that the training module is as realistic and as easy to apply as possible, numerous examples taken from evaluation studies carried out by Groupe URD or its members, or available on the Internet, are referred to during the course. Each one of these examples has been chosen to highlight a particular type of evaluation within the overall approach to evaluation.

Evaluation of global crisis-management in South Sudan. Crisis-management in Bar-El-Gazhal, 1998
Inter-agency and multi-disciplinary evaluation. Post-Mitch emergency aid and rehabilitation in Central America

Economic evaluation. Evaluation of an EU budget line (either in Africa in an ACP country, or in Asia/Latin America as part of the ALA agreement)

Evaluation of an operation carried out within a certain sector by an agency or a group of agencies. In the field of food aid, agricultural rehabilitation, healthcare, protection, anti-personnel mines, prevention/preparation.

ORGANISATION

This module already exists as a condensed training over one day or semi-condensed over three days. The shorter version mainly consists of presenting the concepts. The semi-condensed version includes more thorough presentations, case studies and group work on specific cases. Wherever possible, examples taken from participants' own experience will be used. Indeed, prior to the training, a questionnaire is sent to future participants in order to identify their queries, needs and experience within the field of evaluation.

Depending on the requirements, the number of participants and the length of the training one or two members of staff are required for both these modules. Our teaching staff are primarily specialists within the field of humanitarian aid with extensive evaluation experience.

To conclude the training, participants are invited to evaluate the module itself, applying the various techniques (feedback on needs assessment, identifying a range of needs and requirements, creating a Log Frame, list of achievements, a study of the improved learning skills, an analysis of opportunities of put these skills into practice) in order to demonstrate the pertinence of the techniques.

The main research documents and a relevant bibliography are provided at the end of the module enabling participants to expand their knowledge of a certain number of ideas. Groupe URD's database on evaluation, developed within the scope of the Quality Project, can be consulted online on Groupe URD's website www.urd.org.