CARE'S HUMANITARIAN OPERATIONS:

REVIEW OF CARE'S USE OF EVALUATIONS AND AFTER ACTION REVIEWS IN DECISION-MAKING

Monica Oliver
Georgia State University

CONTENTS

APPENDICES

      Appendix 1A - List of Evaluations Reviewed
      Appendix 1B Meta-evaluation checklist
      Appendix 1C Meta-evaluation matrix
      Appendix 2A Interview protocol
      Appendix 2B Map of CARE emergency response decision-making structure
      Appendix 3 sample report cover sheets

EXECUTIVE SUMMARY

Purpose of the Study

Over the past few years, high-profile disasters and conflicts have been subject to increasing media attention and public scrutiny. The tsunami in Asia captured the global public's attention and heightened its awareness of humanitarian aid organizations and their role in responding to victims' needs in times of crisis. With this increased exposure and higher volume of donations earmarked for crisis response, NGOs are expected more than ever to hold themselves accountable for their own actions.[1] Moreover, impact measurement, which has long been a priority for development programs, has become an increasing area of focus in disaster relief and response. However, as is evident in the current literature, measuring impact is easier said than done in a field where the urgent nature of the situation often precludes collecting baseline data or devising an evaluation strategy prior to responding to the crisis.

Evaluation activity and research can both assess impact from a certain perspective and offer a road map for honing one's practice. This supposes, of course, that the evaluation results are digestible, accessible, and received into a learning-friendly context. CARE International's policies aimed at high-quality programming and effective evaluation indicate CARE's commitment to consistent good quality and continuous improvement of policies and programs. This study reviews CARE International's evaluations of emergency response over the past five years and investigates how well CARE internalizes recommendations and lessons-learned from the evaluations. The purpose of the study is to assess CARE's learning environment and use of evaluations and to reflect on how CARE might more effectively use its evaluation findings to improve its operational performance, inform its policies and better understand the impacts (both intended and unintended) of interventions, taking tips from its own experience and that of its peers as available.

Methodology

This study engaged a three-pronged methodology. The first step involved reviewing and synthesizing evaluation and After Action Review (AAR) reports on CARE's response to humanitarian crises over the past five years. A checklist was used to analyze each of 23 evaluation reports so as to identify the common themes and trends emerging from five years of lessons-learned and recommendations. Secondly, the researcher interviewed 36 individuals involved in various aspects of emergency response for CARE, from Country Directors to Evaluators to Procurement Officers. The interviews attempted to capture the actual and perceived instances of evaluation use. Thirdly, through the interviews and other inquiries, the study identified examples of evaluation use by peer agencies so as to provide opportunity for

---

[1] On the heels of calls for greater accountability among NGOs stemming from the much-maligned humanitarian aid activity in Somalia and Rwanda in the 90s (see Reimann, Kim. 2005. "Up to No Good? Recent Critics and Critiques of NGOs." In Henry F. Carey and Oliver P. Richmond, eds., *Subcontracting Peace: The Challenges of NGO Peacebuilding.* Aldershot, UK: Ashgate) comes the ALNAP-housed Tsunami Evaluation Coalition report calling for transparent, publicly available evaluations, and the Clinton Global Initiative calling for increased NGO transparency and accountability in relief efforts.

reflection on how CARE might innovate and integrate other components of evaluation into its learning environment.

Main Findings

While the checklist highlighted several trends among the lessons-learned and recommendations, three themes, in particular, emerged from the meta-analysis of evaluation reports from 2000 – 2005:

Lessons- Learned: Key Trends

- Decision-Making: The evaluation reports repeatedly expressed the need for an established clear decision-making process during emergencies, from the field to the regional office to the CARE secretariat. Though this might seem self-evident on the surface, it is particularly critical for accountability and is not always clear in an emergency situation where temporary deployed staff team with local permanent staff. Confirming lines of authority, including reporting responsibilities, in all ToRs and at the beginning of each emergency response would mitigate this.

- Training: The need for appropriately trained staff in an emergency emerges in a number of the evaluation reports. This includes orientation to CARE and to context-specific operations for a given emergency. Corollary to this is the importance of maintaining an up-to-date roster of persons available to respond to an emergency.   The perception that this does not exist or is not up-to-date could be shifted through regular distributions of ToRs for all personnel at the outset of a response.

- Evaluation and learning: lengthy evaluation reports have proven difficult to wade through, making internalizing of lessons "learned" a challenge. In addition to scaling down the overall size of evaluation reports, prioritizing and categorizing the lessons-learned would go a long way toward their being embraced by those who can effect change at various levels in the organization.

How CARE uses Lessons-Learned

The interviews pointed to several instances of formal[2] use of evaluation data. Significantly, these instances of use stemmed from individual efforts rather than from a structural learning environment; that is, if someone followed up on a recommendation from an evaluation, it was often due to his own initiative rather than due to a mechanism within CARE for follow-up. There are a number of examples of informal use of evaluations; for example, being asked to participate in an evaluation as an interviewee or in an After Action Review heightens the individual's sense of ownership in the recommendations that follow.   The overwhelming sentiment regarding evaluation reports was that they are too long and too tedious to sift through

---

[2] "Formal" here is understood to mean intended, instrumental use; that is, a group or individual deliberately took an evaluation finding and acted upon it (for example, the leadership in India made improvements to its procurement strategy as a direct response to an evaluation finding it inadequate).

given that everyone is working to and beyond capacity already. The genuine desire to do high-quality work and to do better work was strongly evident in the interviews, but just as strong was the perception of not having the luxury of time to go through evaluation reports and utilize their findings effectively.

A scan of other organizations' experiences of evaluation use suggests that much of CARE's experience is common to the sector; the nature of response to complex emergencies is such that impact measurement, accountability, and evaluation utilization are daunting goals. There are, however, existing models, perhaps even outside of the NGO cadre, that might serve as models on which NGOs can draw.

Recommendations

1. Use of Standardized Templates for:

   o Evaluations: There is little consistency among the evaluation reports reviewed in terms of content and methodology. Standardizing evaluations so that there is a minimum baseline set of data and so that lessons-learned and recommendations are easy to identify by area of responsibility, would greatly facilitate the reports' later use.

   o Terms of Reference (TOR). Use of a format would help promote consistency in evaluation reports as far as delineating methods used, including their strengths and limitations.

   o Template or guideline for AARs: The After Action Review is perceived as a very positive form of learning lessons through evaluative reflection. A thorough how-to for conducting one, or at least reporting on one, would facilitate the use of AAR findings.

2. Annual synthesis of priority themes to coincide with CARE's planning cycle: It is very evident from the interviews conducted for this study that CARE employees are time-starved from the operational level all the way up to senior management. The current typical lengthy report format discourages reading evaluation reports and identifying recommendations relevant to the individual's job. A yearly synthesis and prioritizing of important recommendations culled from evaluation reports and After-Action Reviews would assist in shaping CARE's policy and planning agenda. Several of the individuals interviewed envisioned this yearly synthesis as coinciding with the end of the calendar year in December, in anticipation of January planning sessions for the following fiscal year.

3. Cover sheets to summarize reports linked to a searchable database: As mentioned, individuals perceive evaluation reports as too cumbersome to be practical for incorporating specific lessons-learned. A "cover sheet" for evaluation reports, to be completed by the evaluator, would categorize lessons-learned into areas of specialty, such as human resources, external relations, procurement, etc. This should facilitate the use of the report findings not only by senior management, but also by technical specialists who be interested in more detailed reports. "On-demand" customized reports could be accessed either directly by decision-makers who have some IT skills, or by staff who are responsible for compiling

reports for senior management to access and collate syntheses to help with evidence-based decision-making.[3]

4. <u>Cross-learning opportunities and language:</u> many interviewees felt that other countries and regions could learn from their emergency response experiences, and vice versa. Inviting staff from other countries and/or regions to After Action Reviews and similar events either as a participant or co-facilitator would enable valuable sharing and reflection.   Moreover, systematically translating evaluation reports into other languages such as French and Spanish would enhance wider learning.

Conclusion

While use of evaluations to improve the quality of emergency response within CARE does not appear to be formal or part of an entrenched culture of learning, the informal examples of use are intriguing and point the way toward more effective use through innovative learning mechanisms. As studies by CARE-USA's Learning and Organizational Development Unit have demonstrated, an individual's position and setting within CARE greatly affect that person's use and communication of information,. Consequently, a employee in a CARE Country Office might benefit from a learning exchange visit elsewhere, whereas a senior manager would find a succinct annual synthesis of key lessons-learned trends most useful. The time is ripe for facilitating more effective use in a flexible and inexpensive way. CARE can learn from itself and from its peers to promote better evaluation utilization, thereby improving its emergency response and aspiring to its mission of ending poverty and poverty-related suffering.

---

[3] The Gates-funded ECB project may develop a database shared by IWG agencies.

INTRODUCTION

International attention on emergencies has heightened considerably in the past decade and agencies are increasingly in the media spotlight. CARE International is no exception to this, and with this focus on humanitarian aid has come a dual concern for an agency's capacity to respond appropriately and for an agency's ability to be accountable to its beneficiaries, itself, its peers and its donors. These priorities of capacity and accountability reflect a desire both on an agency level and on a broader level to assure that emergency response programs are of sound quality and that they continuously improve.

As described in their Humanitarian Benchmarks, CARE International strives to uphold its humanitarian assistance programming to a minimum standard of quality. This is evident through CARE's involvement in numerous initiatives: as a major agency among relief organizations, CARE subscribes to the SPHERE minimum standards and to the Red Cross Code of Conduct. CARE is an active member of the Active Learning Network for Accountability and performance in Humanitarian Action (ALNAP) and of the Humanitarian Accountability Partnership International (HAP). Moreover, CARE's commitment to high-quality programs and continuous improvement is evident through its internal policies and practice: the CARE International Project Standards and Program Principles provide such a guideline, as does CARE's Evaluation Policy. Moreover, CARE has commissioned three MEGA meta-evaluations (2000, 2001, 2004, with a fourth pending); the aim of these meta-analyses of CARE's program evaluations is to assure program goal attainment.

This inter- and intra-institutional commitment to accountability in emergency response reflects CARE's ultimate mission of reducing poverty through sustainable programs that respect the rights and dignity of the world's poorest. A major factor in accountability is the ability to look critically at policies and programs in an effort to discern the impact of CARE's response and to pinpoint capacity gaps and areas for improvement. Thinking "evaluatively" about policies and programs requires measuring our relief efforts, disseminating what we learn from such assessments to those who can make the necessary improvements, and putting into action those improvements that are within our means.

Successful learning from the findings and recommendations put forth in evaluations requires an organizational commitment to regular, high-quality program evaluation. Findings and recommendations are of little use unless there is a culture of learning within the organization that promotes dissemination and utilization of such findings from the policy level to the operational level. Such a culture seeks not only to reflect on what has happened, but to influence what will happen so as to carry out the organization's mission ever more effectively.

CARE has increasingly made an effort over the past five years to evaluate its emergency response efforts through a variety of different styles of evaluation, including:

- Real Time Evaluations
- After Action Reviews
- Final Evaluations
- Joint evaluations with other agencies

These evaluations have resulted in a considerable body of information concerning the critical facets of CARE's emergency response activities. The question remains as to how that information has been absorbed into CARE's practice and policies. Current initiatives such as the Emergency Capacity Building Project (ECB) and the Humanitarian Accountability Project (HAP) highlight the desire of the foremost agencies involved in emergency response to enhance their capacity and hold themselves accountable for their actions, and a critical aspect of such accountability is to take reasonable measures to ensure that we try and repeat good practice and don't repeat the same mistakes. For this reason, it is timely for CARE to examine its own evaluation utilization and to ferret out how its evaluation process works and how it might work better.

Evaluation of disaster relief efforts for CARE offers a means for assessing the effect of the organization's response to major emergencies. However, CARE may not have a systematic structure for using its evaluation findings. This does not necessarily mean that the evaluations are ineffective or fall on deaf ears. It is possible that the organizational learning culture circumvents whatever evaluation utilization would come out of a formal structure for use. Also, there is an important distinction to be made between using evaluations to improve future programming in the organization and using evaluations to improve policy, or—even more lofty -- to ameliorate the human condition as emergencies devastate it.

Major Research Questions

1. What are the major characteristics of CARE emergency assistance evaluations?
2. What are CARE's decision-making mechanisms?
3. How does CARE currently use its evaluation research in decision making?
4. Does CARE use findings from humanitarian assistance evaluations to improve policies and programming for future disasters? If so, how?
5. Do CARE's emergency response evaluations influence the organization in ways different from what would constitute direct, instrumental use?
6. How might CARE improve its current ways of evaluating emergency response efforts so that those evaluations are better decision-making tools for the organization?

The answer to the final question should be of interest to any aid organization working in the humanitarian sector, for the opportunity to expand one's influence on changing the deplorable conditions emergency victims face. If use can go beyond program improvement to influence the aid world, the organization will benefit from understanding how it learns from evaluations and how it might facilitate learning with greater reach.

There are three components, or "maps," to this study:

Map I: A **synthesis or meta-analysis** of CARE's emergency response evaluations over the last five years seeks to understand how CARE evaluates, how the organization disseminates findings, and where the patterns are in the recommendations and lessons-learned presented in the evaluations.

Map II: A **stakeholder analysis** of CARE's organizational structure and personnel to attempt to identify who in the organization is likely to use what information for particular purposes and what the channels of communication and influence are.

Map III: An examination of **other examples** in the humanitarian aid world will highlight both how CARE might improve its utilization of evaluations and what CARE is doing well that others might emulate. The original research design (see appendices for ToR) intended a broader "mapping" of examples from elsewhere, even government and private-sector cases. This proved to be daunting in scope. Furthermore, ALNAP (Active Learning Network for Accountability and Performance in Humanitarian Action), a consortium of agencies concerned with inter-agency learning and accountability, chose 2006 to focus its annual study and publication on evaluation utilization among humanitarian aid agencies. Rather than "reinvent the wheel," it was seen as timely for this study and an opportunity to draw from the insights and conclusions gleaned from the ALNAP research.

EVALUATION USE

Current thinking on evaluation use stems from a body of work on knowledge and research utilization that sprang from a public-sector focus on social betterment in the 1960s and 1970s. With the onslaught of social science research came a heightened concern for accountability, and it is out of this period that program evaluation became an established area of practice. That research theoretically results in new knowledge was not a novel concept during this period; of greater concern was how practitioners utilized that knowledge. Social science research provided a means of identifying cause-and-effect relationships between programming and results, which was of keen interest to decision-makers. Nevertheless, the factors contributing to effective utilization of research remained to be identified. Much of the thinking during this period looked at the use of research on a national or policy level; the organizational level was still to come.

The need to distinguish between different kinds of use became apparent. In 1979, Harvard researcher Carol Weiss introduced instrumental use as a term for the classic linear form of use, or the use perhaps, that the researcher intended: the researcher or evaluator proffers knowledge directly to the user, who in turn uses it immediately and as the researcher envisioned. Conceptual use, for Weiss, is less direct and occurs when a piece of knowledge influences an individual's thinking about a policy or program. Finally, symbolic use occurs when decision-makers use research knowledge for political gain or to justify already-made decisions.

More recently, attention toward "process use" has introduced the idea that the very process of evaluating is itself a form of use: an interview can be a type of intervention, data gathering can increase inter-organizational communication, and the evaluator's interaction with stakeholders can provide them with an opportunity for reflection (Patton 1997, Shulha and Cousins 1997). The idea of process use expanded beyond the individual's use to include organizational learning (Shulha and Cousins 1997). The study of evaluation use in an organizational setting led to greater questions about organizational learning and knowledge management. Evaluation

utilization relates to the organizational learning environment. This is particularly true if, as Patton (1994) suggests, the very process of evaluating *is* the learning environment.

The body of literature on organizational learning is vast and spans decades. Argyris and Schon's seminal work on organizational learning fleshes out the seeming paradox embedded in the concept: individual people learn, retain information, transfer information, and so forth, so how can an organization be said to "learn"? What and how does it learn? They get around this paradox by examining what it is to be an "organization." An organization has procedures and boundaries, and significantly, it designates individuals to make decisions for the whole. If individuals can act on behalf of an organization, then they can learn on behalf of an organization. They distinguish between different types of learning: single-loop learning changes either individual assumptions behind organizational strategy, or changes the organization's strategy. Double-loop learning, on the other hand, changes values in addition to strategies and assumptions. This distinction suggests that there are different levels of learning and that each level depends greatly on the learning environment.

Peter Senge brought the term "learning organization" into mainstream usage in 1990 and offered the characteristics of such an organization. Senge named five "disciplines" of a learning organization: 1) personal mastery of individual vision and of objective reality; 2) mental models, or assumptions affecting how we see the world; 3) building a shared vision of the future; 4) team learning and dialogue, and 5) systems thinking, or being able to "see the organization's patterns as a whole…from within the organization." This fifth discipline incorporates the other four and is critical to the organization's evolving as a learning organization with a learning culture. For Senge, managers must learn to strategize, not merely within the scope of their own responsibilities, but about the whole system and in the long term.

Humanitarian aid organizations have characteristics that distinguish them from corporations or from other nonprofits in terms of learning environment. Working in a developing country often means confronting an unpredictable, chaotic setting with little infrastructure. Moreover, cultures and levels of development can differ so much from one to another that it appears impossible simply to apply a program to one place just because it worked in another. This study attempts to examine evaluation use within the learning organization that is CARE.

METHODOLOGY

Map I: Meta-analysis of evaluations of humanitarian actions. This first exercise draws from four main sources to create a checklist (see Appendix 1A) by which to organize the lessons-learned from the last five years' worth of evaluations:

- CARE International Project Standards and Program Principles
- CARE International Evaluation Standards
- ALNAP Quality Pro Forma
- Checklist for Developing Checklists, American Evaluation Association.

Completion of the checklist for each of the evaluations allowed for identification of patterns, trends, and timelines. Also, the evaluation methodology, absence or presence of a ToR, and notation of whether internal or external evaluators conducted the study were recorded for each evaluation.

Map II: Stakeholder analysis – How does CARE make decisions? The stakeholders for these evaluations are those individuals for whom the evaluation findings may guide their designing and implementing current and future policies and programs. The main data source for this is semi-structured interviews of some 35 individuals at different positions within CARE over the summer of 2006. Their insights identified systems (or lack thereof) in which the evaluation findings are situated, as well as the variety of perspectives on how CARE might better streamline the learning mechanisms in place currently. (See Appendix 2C for a complete list of interviewees).

Map III: Examples from other organizations. This section is very much a work in progress, and was drawn largely from ALNAP's research project on evaluation use, which was carried out during 2006 to better understand stakeholder dynamics and the effects of those dynamics on evaluation use. In addition to ALNAP's project, there is potential to tap into a couple of corporate models that CARE employees and others have identified as innovative systems for learning from evaluation findings.

The methodology used for these evaluations consists almost entirely of interviews, document review, and field visits (After Action Reviews used primarily group reflection activities). From a research standpoint, methodological rigor does not come across as a priority in the evaluation reports or in the ToRs. This is not to say that the methods employed are unsound; rather, it is an observation that the design of the evaluations appears on the surface to be less important to the organization than other elements of the evaluation.

Limitations to the Study

Most, though not all, of the interviews conducted for this study were with employees of CARE USA. Though this is in part a function of CARE USA's comparatively significant size and role in emergency response, it is important to acknowledge here that the research findings represent CARE USA more accurately or more thoroughly than they do some of the smaller CARE members. As with any qualitative research, there is a level of subjectivity to this study, particularly in its heavy reliance on interviews that capture individuals' perceptions. A single person conducted the interviews, leaving room for the possibility of a one-sided interpretation of the interview data. Also, the third mapping exercise is not meant to be an exhaustive exploration of other accountability models, nor even a representative sample; rather, has been reshaped from the original research design to draw primarily from the research of ALNAP vis-à-vis its member agencies. As such, it is not representative of the many models out there, but does reflect the practice of some of CARE's peers.

Background Documents Reviewed

The evaluations examined fell between 2000 and 2005. Though 36 CARE evaluation reports were identified as having been completed in this time period, this study focuses on 23 of them, to avoid confusion with those projects that crossed over from relief to development or those whose evaluation reports were incomplete. After-Action Reviews/Lessons-Learned, Real Time Evaluations, Final Evaluations, and Joint Evaluations comprise these 23 examples. The documents are usually the product of a team of evaluators. This evaluation team may be made up of evaluators internal to CARE, external to CARE, or a combination oergency response is only a part of their jobs, and they are working beyond capacity as it is, so they don't have the luxury of sifting through lengthy reports and learning from them.

Conclusions: a frequent interview comment involved the desire to be more involved during the evaluation process. Three interviewees remarked that if they hf both. The conclusions resulting from this analysis reveal both a pattern of recommendation themes and of repeated recommendations over time.

EVALUATION REPORT CONTENT & FORMAT

There are several types of evaluations included in this study:

Final Evaluations take place at the end of a project or program. They look at the entire length of the program, from the beginning of a response to its completion or transition. Typically, an external evaluator or evaluators comes on as a consultant to lead the evaluation.

Real Time Evaluations occur while a project is still in full motion. Their intent is to assess the effectiveness of the emergency response so that the response can be tweaked as necessary for better effectiveness.

After Action Reviews. Also called Lessons Learned workshops, these events take place immediately after an intervention, and aim at capturing the experiences of those who staffed the emergency. The ultimate goal is to reflect on what went well and what might be improved upon for future interventions. These reports are typically shorter than other evaluation reports, as the workshops themselves, in contrast to a full-scale evaluation, are only a few days long.

Joint (Multi-agency) Evaluations. These represent a collaborative effort between multiple agencies addressing a given emergency to assess their coordination efforts and collective impact on the situation.

Key Findings (see Appendix 1C for complete matrix of all findings/patterns)

Map I: Meta-Analysis of Trends

The evaluation reports vary widely in terms of content and format. The reports reviewed are, for the most part, lengthy; the reports range from 2-page summaries to 78-page reviews, with the majority of the reports containing 25+ pages. The exception to this is the After Action Review

summaries, which attempt to capture information from a few days of reflection rather than from the duration of an entire program. Most of the reports contain an executive summary, but in some cases these summaries consist of a few pages highlighting the report's main lessons-learned, whereas in other cases, the executive summary is a mere paragraph or two generally explaining the study. This suggests that those preparing the reports are doing so without explicit guidelines as to whether or how to prepare the executive summary. The lessons-learned and recommendations themselves vary considerably as far as quantity and depth. Moreover, only a few evaluations make a distinction between "lessons-learned" and "recommendations," and none explain what constitutes a "lesson-learned." A firm understanding of what a "lesson learned" is – and what it isn't – could be a useful construct for CARE.

Beneficiaries

The evaluation reports included the occasional call greater inclusion of beneficiaries in project design: Afghanistan 2002, for example, recommended increasing consultation with beneficiaries for emergency projects. Sri Lanka/India 2005 noted that the emergency response suffered in instances where beneficiary consultation was sacrificed for efficiency's sake. However, beneficiary-focused recommendations, or accountability to beneficiaries, were the exception rather than the rule in these reports. It should be noted that for the evaluation reports falling outside of the scope of this study (for those projects concerned with long-term development and rehabilitation), beneficiaries are a more central feature in the recommendations. Furthermore, though the evaluations seem at first glance not to focus much on emergency assistance beneficiaries; this could be a function of language confusion. One evaluator's "beneficiary" may be another evaluator's "community," a term with very different connotations. Notably, the Hurricane Jeanne evaluation for Haiti mentions the needs assessment conducted with beneficiaries there as a "good practice" that CARE would do well to replicate elsewhere.

Decision-making

Decision-making Chain of command themes to emerge from the evaluation reports. Moreover, recommendations to specify or clarify the chain of command cut across regions, type of emergency, and time. Examples of lessons-learned involving chain of command include:

- *CARE's overall disaster planning has not established clear "emergency" roles, responsibilities and procedures for deploying staff. Deploying emergency staff learned by trial and error.* (India 2001)

- *Clarify lines of authority...national offices should follow up with country offices on reporting deadlines* (Afghanistan 2002)

- *Reporting lines within a country office should be re-articulated as soon as possible after arrival of external emergency response personnel* (West Bank/Gaza 2002)

- *Facilitate clarifications of roles and responsibilities of CO and CERT staff* (Iraq 2003)

- *Divisions of labor need to be clearly defined* (DRC 2004)

- *Lines of responsibility and leadership should be clearly defined* (Haiti 2004)

The majority of the evaluation reports reviewed include some version of a recommendation regarding clarifying lines of authority, visits from senior management, and follow-up on responsibilities. The various recommendations related to lines of authority do not all point to the same suggested structure, but it is clear that established lines of communication and reporting are a priority at all levels of emergency response and directly affect the efficacy of the response.

Human Resources and Personnel

- Training for staff deployed to an emergency. This involves orientation to CARE and how it is organized, as well as familiarizing deployed personnel with local procedures and operations. Training local staff, using local training methods as appropriate, is also a common theme. The Afghanistan 2002 evaluation report is particularly detailed in training observations and recommendations.

- The added workload for disaster response is a recurring concern throughout the reports reviewed:

  "Concern has been expressed that many CARE staff now remain with workloads exponentially increased from pre-disaster days…a closer examination of the current division of labour and staff efficiency/motivation would now be useful."
  Evaluation of CARE Haiti's response to Tropical Storm Jeanne

- Four evaluation reports point to a need for having a terms of reference (ToR) for every person deployed to an emergency, or having generic TORs as part of the CO's preparedness plan.

- The need for acceptable living conditions for deployed staff was a concern in three reports.

- Three reports recommended maintaining an active roster of available and qualified persons for emergencies, suggesting that if such a thing indeed already exists, that is not always the perception in the moment.

- Finally, three reports recommended having a senior staff person or senior management person visit the site of the emergency as early as possible into the response. This sentiment is echoed in the interview portion of this study; buy-in on the part of senior management is regarded as critical to morale.

Preparedness Planning

Several of the reports pushed for having an emergency plan firmly in place at the onset of a crisis. Additionally, the reports called for orientation of deployed staff to CARE as an organization, if necessary, and orientation to that particular country office. One report called for

a revision of the CI emergency manual. CARE India's AAR for the tsunami response emphasizes the need for better preparedness in several areas, including procurement, policies, and long term strategy. Five reports mention risk reduction and contingency planning as necessities for better efficiency in the future. As with "beneficiaries" above, "preparedness planning" is a term that is just as often called something else, such as contingency planning or risk reduction.

Procurement and Logistics

Procurement received little mention in the evaluation reports, and in general, procurement-related recommendations were specific to the context in which they were observed rather than being generalizeable CARE-wide. The evaluation report for Albania (2000) recommended a permanent procurement capacity. The India earthquake report mentioned the need for a procurement database. Procurement was of greater priority in the interviews than in the evaluation reports reviewed. Many of the logistics observations concerned communications and the need for adequate devices. The Iraq RTE specifically recommended a minimum standards for procurement for critical items such as vehicles.

Finance

One evaluation report recommended a finance manager for the start of any emergency operation. The Iraq report recommends bringing in an external finance manager if resources allow. Some reports included situation-specific suggestions for soliciting funding (e.g. the 2001 Kenya report suggested simultaneously seeking funding for environmental rehabilitation). A couple of reports recommended a CI emergency fund for a more immediate capability in disaster response situations.

Psychosocial Iissues

Four evaluation reports mentioned the psychological duress characterizing emergency response staff. The evaluation on Tropical Storm Jeanne, for example, comments that psychological support was an obvious staff need and was late in coming. The Darfur Real Time Evaluation report, similarly, highlighted the importance of counseling both for the displaced Sudanese and for CARE staff for coping with the crisis. The Kosovo After Action Review recommended the continuation of provision of counseling to staff. The Multi-Agency evaluation for Thailand and Indonesia noted the great need for psychological healing on the part of both victims and response staff, noting that current resources are not adequate.

Communications

Communications recommendations range from observations about the need for reliable technology to comments about the criticality of a seamless flow of information. This theme also emerged in the interviews. Though many of the communications recommendations were situation and context-specific, the over-arching theme was that lines of communication need to be established between CARE factions for each and every emergency response. The Multi Agency Evaluations asserted that coordination among relief agencies is essential to an effective

response, remarking in more than one instance that the current level of coordination between agencies is not sufficient.

Security

Not surprisingly, security came up only in the conflict-area emergency evaluation reports, such as Afghanistan and West bank/Gaza. Two reports (Afghanistan, Iraq) recommended that CI develop a security protocol. The Darfur report expressed a desire that security plans be in place prior to a crisis in volatile settings such as Darfur.

Project Planning

Aspects of project planning / preparedness came up frequently in the evaluation reports. Some distinguished between short-term planning and mid-to-long-range planning strategies. The need for a situation analysis to precede any response was a recurring theme. The sentiment emerging from the reports is that a recognized emergency preparedness plan would help CARE in responding more rapidly to crises.

Public Relations / Information Management

A number of reports mentioned a desire to cultivate relations with external partners in order to respond more efficiently. One of the After Action Reviews recommended training more individuals to respond to the media's questions, as the few who were trained to respond were fatigued.

Very little mention was made of information management, except in the context of the desire to improve internal lines of communication and reporting.

EFFECTIVE VALUATION REPORT UTILIZATION:
TROPICAL STORM JEANNE

More than one interviewee pointed to the evaluation and After Action Review for Tropical Storm Jeanne in Haiti as a model of how evaluation can effectively inform planning and preparedness. What accounts for this?

Timing. The original relief effort required 500-600 staff. Though it seemed important to assess the relief effort early in the response so as to include staff and avoid losing information, the reality was that the staff was stretched to their limits with the response effort. An initial review in January 2005 following the September storm allowed for the participation of a good cross-section of staff, despite the fact that some had already departed. Moreover, the distribution of the final evaluation report in March allowed for its use in CARE Haiti's AOP **annual planning event** in April. The report identified resource gaps, such as storage and distribution points for potable water, that the planning session was able to address for the following fiscal year. The report was useful in **scenario-building** and subsequent contingency planning.

Morale. The Haiti After-Action Review and thorough subsequent evaluation both provided a forum for staff to **reflect** and highlighted **what they had done well** in the response, rather than remaining limited to where their response effort had fallen short.

Communicability & Follow-up. In addition to two external consultants, CARE Haiti's M&E focal point also joined the evaluation team. This helped the team to ensure their recommendations were realistic for CARE Haiti, there was good follow-up, and the CARE staff member himself learned a great deal. CARE Haiti translated the report into French, which improved its use amongst local staff and partner agencies.

Policy/Advocacy

Policy and advocacy came up frequently in the evaluation reports; they emerged less in the individual interviews. However, one report (Afghanistan) referred to advocacy as an "appropriate" response for CARE due to an appreciation by peer agencies of CARE leadership role. The India earthquake evaluation report suggested that advocacy for beneficiaries would have helped in clarifying their rights to relief assistance.

Learning

Report after report called for training in emergency response for CARE staff, both temporary and permanent. Three reports called for building adequate evaluation resources into all emergency plans and budgets and another three reports highlighted the need for early and/or consistent data collection as far as emergency response. Finally, having minimum standards or guidelines for monitoring and evaluation was a recurring recommendation.

Map II: Interviews with CARE Staff[4]

Interviewees represented different tiers of the emergency response framework within CARE (see Appendix 2B). Respondents ranged from those on the "front lines" directly involved with the emergency response to those on the executive level. There were 33 formal interviews, supplemented with several informal conversations and interview follow-up discussions. Interviews were conducted both by telephone and in person, and averaged about 45 minutes apiece. The interview protocol (see Appendix 2A) provided a loose format for the semi-structured discussions.

The most common refrain from the interviews was the desire for shorter, more pointed evaluation reports. Those on the front lines remarked that they did not have the luxury of time to read **lengthy reports** and do their jobs in the field; those at the executive level commented that they did not need 40 pages worth of information in order to use the reports to make good policy judgments. One person specifically recommended a maximum length of 20 pages; most wished simply for recommendations to be on top, separate from the main text body. Three interviewees called for the recommendations to be categorized by job responsibilities, such as finance, human resources, security, etc.

In addition to concerns about the length, many interviewees felt that the **distribution** of evaluation reports is inconsistent. Some were not sure whether reports were systematically shared. Very few of the interviewees seemed to know where to look if they wanted to locate a repository of reports; only one interviewee mentioned Livelink specifically[5].

The **focus** of the evaluation reports was also of concern to many of the interviewees. The overarching sentiment was that they did not have time to read through and pick out the lessons-

---

[4] see Annex 2C for more details
[5] Copies of all CARE's humanitarian evaluations are currently archived on Livelink.

learned that applied to their specific tasks, nor the time to go through old evaluation reports when dealing with a new emergency.

Furthermore, there was a question among interviewees of **accountability**, some of whom should have responsibility to ensure that lessons are in fact learned:

The key thing is accountability – lessons learned do not become action if there is no accountability, no incentive to change.

Is there an action plan for following up on lessons learned and recommendations? If not, that would be helpful.

Interviewees working at a more operational level experience found evaluation reports are often too "theoretical" or "academic" and would have preferred much more practical guidance. There were suggestions to simplify language and regularly translation into French and Spanish.

Several interviewees mentioned a lack of a **learning culture** within CARE, along with a lack of structure into which learning could be fed and retained. For many such a culture was about attitudes and behavior rather than about organizational structure. Suggestions included looking to other models perceived as successful, such as that of World Vision (one interviewee specifically mentioned WVI's comparatively well-organized procurement system for emergencies) or even corporate models.

The interviews yielded very few examples of lessons-learned from evaluation reports that had led directly to actions meant to address them. Most interviewees agreed this should be happening, but did not feel it was realistic at present. Several attributed this gap to capacity, while acknowledging that if they had been interviewed for an evaluation, or had been asked to participate in an AAR, they would have been more likely to read the ensuing evaluation report. Another frequent refrain was the issue of accountability; there is no incentive for following up on recommendations, and no penalty for not doing so. Interviewees seemed genuinely to want to do their job well and do it better if possible; evaluations were seen as time-consuming and a hindrance, rather than a means to that end.

RECOMMENDATIONS

The mapping exercises highlight a particular strength within CARE: employees really do care about doing their jobs well and are willing to work hard at that in the most adverse of circumstances. When lessons go unlearned, it is not so much a problem of indifference as it is of attitudes, time availability, and confusion over whose responsibility it is to assure that change occurs. The following are recommendations for building a stronger culture for lesson-learning:

1. <u>Standardized Template for evaluation reports</u>. The evaluation reports reviewed are not uniform in terms of form, content and methodology, although recent evaluations are becoming more consistent. A standardized format would be helpful for those looking to

skim the report rapidly and would help evaluators in ensuring that their outputs were in line with CARE's expectations. Fields on the format could be linked to a searchable database to allow easy access to lessons learned in a concise format, either from individual evaluations or in the form of a synthesis (e.g. a summary of recommendations relating to human resources over the past two years).

How: The recommendation is for a **standardized ToR and evaluation format** that would include:

- qualifications of the evaluator including whether he/she has ever worked for CARE and/or "knows" CARE,
- methods used to conduct the evaluation,
- minimum baseline data collection,
- evaluation findings for a list of subcategories,
- findings and recommendations. This last category would have a clear distinction made between a "lesson learned" and a "recommendation," with a section for "good practices," or positive lessons-learned that ought to be replicated.

2. Template or guideline for AARs: The After Action Reviews were perceived positively by most of the interviewees as a means of identifying lessons-learned through evaluative reflection. A thorough how-to for conducting one, or at least reporting on one, would facilitate the use of AAR findings.

   How: The AAR should take place early enough that those who responded to the emergency are still there, but late enough that the AAR does not interfere with the response effort. The review should consist of reflection both on the process of the response and on the end result. In addition to the individuals directly involved in the response, representatives from human resources, procurement, logistics, security, and external relations should be invited to participate. Each attendee should receive at least a summary of the AAR notes and recommendations. The facilitator should be competent in the relevant language for the AAR.

3. Yearly synthesis of priority themes to coincide with planning cycle of CARE: It is clear from the interviews conducted for this study that CARE employees, like those of most nonprofit organizations, are time-starved from the operational level all the way up to senior management. Moreover, the lengthiness of the evaluation reports dissuades people from reading them and from wading through text to identify recommendations relevant to the individual's job. An annual synthesis of important themes and identification of themes on which to focus for the year would assist in shaping CARE's policy and planning agenda. Several of the individuals interviewed envisioned this yearly synthesis being produced in December or January of each year, so it would be available for the AOP and Strategic planning sessions for the following fiscal year. Such a synthesis, with follow-up from previous syntheses, would be appropriate at bi-annual ERWG meetings as well.

<u>How</u>: the person responsible for quality assurance within CARE's emergency group would lead a synthesis exercise in November and December of each year. (It is important that the person leading the exercise be thoroughly familiar with the responses reviewed, and that the person be high enough within CARE to assure buy-in from all concerned. The synthesis report could be triangulated by having another researcher, new to CARE, categorize the data as well. The exercise would involve reviewing any evaluation reports from the year leading up to that point and prioritizing the recommendations listed in each, identifying who should be responsible for follow-up for each of the recommendations. The synthesis would be prepared for planning meetings in January and should be concise.

4. <u>Cover sheet for evaluation reports that can feed into a searchable database</u>: Individuals perceive evaluation reports as too cumbersome to be practical for incorporating specific lessons-learned. The reports are lengthy, and recommendations targeting a specific area, such as human resources, get lost among all the other recommendations. A "cover sheet" for evaluation reports should be completed by the evaluator preparing the initial report, which would categorize lessons-learned into areas of specialty (such as human resources, external relations, procurement, etc0, so as to facilitate the use of the report findings by individuals who are responsible only for a slice of the findings. Though the evaluation reports are now easily accessed in their Livelink location, it is more of a repository than a database. A **database** would allow searching by region, or disaster type, or by job sector, specifically: human resources/personnel, finance, procurement/logistics, advocacy, security, and monitoring/evaluation.

<u>How</u>: Potentially, such a cover sheet could eventually be incorporated into a searchable database allowing users to search for evaluation reports containing information relevant to their jobs. The cover sheet would be no more than two pages long and the evaluator submitting the report would complete the cover sheet. Topics included on the sheet:

❑ Short abstract of the emergency context
❑ Time frame of the response and evaluation(s)
❑ Country & Region
❑ Sector(s)
❑ Type of evaluation (AAR, RTE, joint evaluation, etc.)
❑ Type of emergency (conflict, natural disaster, slow-onset, etc.)
❑ Lessons-learned categorized by job function (logistics, procurement, human resources, external relations, etc.)
   Lessons-learned categorized by organizational level (field/operations, regional, headquarters)

5. <u>Policy on internal vs external evaluators</u>: The interviewees, when queried about the plusses and minuses of using internal versus external evaluators (internal to CARE vs external to CARE), responded predictably that while external evaluators sometimes have too large a learning curve in terms of understanding how CARE works, they bring a fresh perspective. Internal evaluators, on the other hand, know how CARE is structured but can be in a politically awkward situation within the organization or can lack objectvity. They may not always ask the tough questions. Most interviewees agreed that a team of

evaluators, internal and external, is ideal when possible. Creating a "bank" of external evaluators who are familiar with CARE and who are known to be competent would facilitate this. This already exists, in a way; human resources maintains a roster. The recommendation is that this resource be formalized.

How: Though "prior CARE experience" is a criterion for hiring evaluators, there is a perception among interviewees that there is great disparity between external evaluators in terms of their understanding of CARE and their experience as evaluators. Formalizing, even training, a group of emergency evaluators could assure that they know CARE, know the desired format and content of the evaluation, and know how to complete the "cover sheet" (see above).

6. Learning opportunities: Several of the interviewees were of the opinion that other countries and regions could learn from their emergency response experiences, and vice versa. Inviting staff from other countries and/or regions to After Action Reviews, planning meetings, participate in evaluations as team members and other such events on a rotating basis might enable valuable sharing and reflection.

How: Budget for at least one individual from a neighboring region's CARE office to sit in on each After Action Review. Share that individual's reflections and reactions widely. Also, systematically translate evaluation reports into French and Spanish.

CONCLUSION

Evaluations of humanitarian aid missions have the potential to affirm and bolster staff morale, discover and increase good practices, and highlight areas for improvement. CARE has experience with effective evaluation utilization, and with some changes to its structures and systems, has the capacity to encourage a culture of learning while putting the framework in place for improved practice. Annual prioritization and synthesis of lessons-learned, a searchable database, standardized evaluations and reviews, and increased learning opportunities are simple steps that can lead to a host of practical improvements and attitude shifts, as indeed evaluations already have in isolated instances.