JANUARY 2013

Feinstein International Center

Strengthening the humanity and dignity of people in crisis through knowledge and practice

Source: IRIN 2013

The Use of Evidence in Humanitarian Decision Making

ACAPS Operational Learning Paper

James Darcy, Heather Stobaugh, Peter Walker, and Dan Maxwell

acaps

Tufts UNIVERSITY | Gerald J. and Dorothy R. Friedman School of Nutrition Science and Policy

©2013 Feinstein International Center. All Rights Reserved.

Fair use of this copyrighted material includes its use for non-commercial educational purposes, such as teaching, scholarship, research, criticism, commentary, and news reporting. Unless otherwise noted, those who wish to reproduce text and image files from this publication for such uses may do so without the Feinstein International Center's express permission. However, all commercial use of this material and/or reproduction that alters its meaning or intent, without the express permission of the Feinstein International Center, is prohibited.

Feinstein International Center
Tufts University
114 Curtis Street
Somerville, MA 02144
USA
tel: +1 617.627.3423
fax: +1 617.627.3428
fic.tufts.edu

Acknowledgements

This study was supported by a grant from the Assessment Capacities Project (ACAPS), and we gratefully acknowledge their support. We would like to thank all the agency staff, government officials, global experts and other key informants whom we interviewed in the course of this research. Responsibility for any errors in this report lies solely with the authors.

Contents

1. Introduction **5**

2. Decision making in the humanitarian sector **7**

2.1 Overview 7

2.3 Decision making in a conflict context: DRC case study 9

2.4 Decision making in slow-onset emergencies: Ethiopia case study 13

2.5 Decision making in fast-onset and recurrent crises: Philippines case study 15

3. The generation and use of humanitarian evidence **19**

3.1 Introduction 19

3.2 The generation and use of evidence from early warning systems 21

3.3 Evidence from needs assessments 22

3.4 Coordination of needs assessments 24

3.5 Response analysis: The use of evidence in choosing the type of response 25

3.6 Needs assessment and response analysis tools 26

3.7 Changing circumstances and new evidence 28

3.8 The use of evidence in innovative responses 29

4. Conclusions **31**

4.1 Overview 31

4.2 Specific conclusions 31

5. References **34**

6. Annex **37**

6.1 Research methods 37

6.2 Individuals interviewed 38

1. INTRODUCTION

1.1 This paper reports the results of a study undertaken during 2012 by Tufts University for the Assessment Capacities Project ACAPS,[1] as part of the latter's "Operational Learning" strand of work. This study is designed to complement the work of ACAPS on strengthening needs assessment by addressing the question of how assessments and other sources of information and analysis are used by humanitarian decision makers. The study is based on a combination of literature review, case studies, and key informant interviews.[2]

The pressure to demonstrate that responses and claims about impact are grounded in evidence has been growing over recent years. Humanitarian donors are increasingly under similar pressures to those faced by other public sector colleagues—including those in the wider development sphere—to demonstrate effectiveness and account for impact in ways that they have not previously been expected to do. This is partly a matter of showing that the *policies* on which their decisions are based are themselves well-founded and evidence-based.[3] But it relates also to individual programs, the way response decisions are justified, and how impact and outcomes are measured. These pressures are in turn felt by the agencies that donors fund, as part of a wider accountability agenda. Accountability on these grounds demands more than just reference to the internal logic of programs, through logframes or otherwise: it demands that reference is made to external evidence for claims made. The questions raised here are therefore of far more than just academic interest, both because they concern real outcomes for real people, and because current practice is increasingly under challenge.

The study is based around three main questions. First, how do decision makers in the humanitarian sector currently use information and analysis? Second, what factors, other than information and analysis, are influential in making decisions? Third, what would enable better-informed response decisions? The first two questions are empirical ones, though not always easy to answer, since the process of decision making is often opaque. The decisions involved range from funding decisions by donors to program design decisions by implementing agencies. The third question is more speculative, though the study draws on examples of good practice to derive its conclusions. The issues relating to "response decisions" are not limited to the choice of response type (food aid, cash, etc.), but extend more generally to the *appropriateness* of intervention decisions, including questions of timing, relevance, fit with local preferences, etc. In this regard, the paper understands a well-informed decision to be one that takes due account of data and information most relevant to the crisis context, and combines this with experience-based knowledge to determine what intervention is the most appropriate in that context.

In order to address the overarching questions above, the study looks first at some of the main processes of decision-making in the humanitarian sector and the factors that appear to have most influence on decisions of different kinds. It goes on to look at the way information and analysis is currently generated in the humanitarian sector—both through formal and informal means—and related questions of relevance and credibility. These two topics are then brought together in addressing the question of the use of information by decision makers,

[1] ACAPS is an initiative formed by a consortium of three NGOs (HelpAge International, Merlin, and Norwegian Refugee Council) in collaboration with other humanitarian actors, including the IASC Needs Assessment Task Force. The ultimate goal is to facilitate more effective, efficient, and appropriate humanitarian responses. The aim is to promote better informed and more evidence-driven responses, specifically by supporting a process of coordinated needs assessments that are timely, coherent, and appropriate to context, whereby results are accessible and relevant to decision makers.

[2] Details of the methodology and sources consulted are contained in the Annex.

[3] The wider evidence-based policy movement, described by Stern et al. (2012) is perhaps best exemplified by developments in the medical sphere over the past two decades. Evidence-based medicine is concerned primarily with the justification for claims about the efficacy and safety of a proposed course of treatment for a given medical condition.

and what might enable more informed and evidence-based response decisions.

1.2 Humanitarian contexts are almost by definition "non-ideal." Decisions often have to be made quickly, sometimes with relatively little access to current information or accurate data. The question about informed decision making may therefore come down to what constitutes a "well enough" informed decision in the circumstances; or what constitutes "good enough" information and analysis on which to base a response. There is an accountability dimension to this: in reviewing decisions made, can the responding agency claim to have made reasonable use of the information available in formulating their scenarios and response plans? Allowing for the constraints of crisis settings, did the agency take reasonable steps to inform itself, including talking to those affected by the crisis and the intended beneficiaries? There is a natural presumption that it should—whether through a formal needs assessment process or otherwise—and that this should be one of the core criteria of evaluations.

Whatever the quality of information, no assumption can be made that the increased availability of good information and analysis will *in itself* result in better informed decisions. In reviewing the way decisions are made in practice, the study considers the ways in which such information is used (or not) at different points in the process, which varies across different kinds of decisions in different contexts. There is no single template here, but some patterns do emerge.

2.1 Overview

Previous work by ODI (ODI 2009; Darcy, Anderson et al. 2007; Haan, Majid et al. 2005) reviewed the ways in which decisions are made by managers in the humanitarian sector, providing a reference point for the current study. In general, these studies found that external information *per se* was found to have limited relevance for decision makers. Most decisions appear to be made within quite tight parameters: the range of options being limited by previously decided questions about strategic priorities, available resources, and so on. In the jargon, these decisions are highly "path dependent." What mattered as much as external information was the understanding people had of the institutional framework for decisions, the implicit values and assumptions that they applied in making decisions, and the mental models by which they processed available information.

The ODI studies suggest a number of features common to individual decision making. They suggest that decision makers with limited time to make decisions tend to rely heavily on the judgment of people they trust, both in constructing the humanitarian "narrative" for a given crisis and in defining response options. This can result in a relatively unchallenging attitude to proposals and evidence used to support them. To quote the ODI 2009 study: "Suppose for example that a proposal is received by a donor from a trusted agency colleague…. If this matches the donor's need to find a partner to respond in that particular sector, it seems that the donor is much less likely to test the evidence on which the proposal is based or to seek corroborative evidence. Conversely, if the proposal is from a non-trusted partner or for an approach that is not 'mainstream,' it will need to pass a much stricter test" (ODI 2009, p. 11). This phenomenon was also confirmed in this study, as individuals from large donor organizations expressed that a largely influential factor in deciding which organization or proposal to fund consisted of whether or not they have had a strong working relationship. As one respondent from a large donor agency explained, "It really comes down to trust—we trust [implementing partners] based on past performance."[4]

BOX: A typology of humanitarian decision making

The ODI studies referred to above propose a typology of four main decision types relating to crisis response in the humanitarian sector. We present this here in slightly modified form:

- **Strategic** decisions about whether and how to respond, including macro resource allocations (approach, level, and channel of funding, etc.);
- **Program design** decisions (including targeting);
- **Planning** and micro resource allocation decisions: what resources (money, people, etc.) to allocate and how to allocate them (team composition, budgeting, etc.);
- **Operational** decisions concerning program implementation and modification.

The studies also distinguished *levels* at which decisions were made:

(i) Within organizations: HQ, regional, national, local levels;
(ii) System-wide or inter-organizational.

Maxwell et al. (2012) use a slightly different but related typology of decision making in their work on response analysis. They distinguish "first order" decisions to be about strategy and

Continued on next page

[4] Interview.

Continued from previous page

overall goal-setting. "Second order" decisions involve choosing the intervention modality in which to achieve the general objective set by the first-order decision. "Third order" decisions are about specific program planning issues, such issues as targeting and conditionality. These are supplemented by a range of decisions about cross-cutting issues affecting implementation that must be addressed, such as capacity, risk, gender, logistics, etc. Once these decisions are made, then agencies address more detailed program design and operational questions.

Although the two studies differ slightly in the terminology, conclusions regarding the typology of decision making are quite similar, revealing a fairly consistent pattern across various agencies.

The way information is presented to decision makers appears to be very important to uptake; and the interpretation (and interpreter) of information is just as important as the source. As ODI (2009) notes: "The way in which information is presented can be crucial to its uptake and use by decision-makers. 'Killer facts' were sometimes cited [by respondents] as highly influential on decision-making (e.g., very high reported levels of acute malnutrition) even where these were speculative. Less dramatic facts, such as significant changes in underlying indicators, tended to go unremarked unless presented as part of a case for action. Succinct presentation of information was one key factor in its influence" (ODI 2009, p. 11). This study also confirms these previous findings of the ODI (2009) study. In addition to respondents indicating a need for information to be presented in a concise manner, the information must be provided in such a way that is easily understood by non-technical decision makers. Many expressed feeling overwhelmed by being presented with large quantities of raw data. Furthermore, others identified that without a systematic way of collecting data and comparing across different sources, understanding relative needs and prioritizing action is very difficult for decision makers.

This relates to another important point about information. It can be used as evidence to support a case for action, but somebody generally has to *make* the case and structure the information in a way that supports it. The validity of the process depends partly on the reliability of the information, partly on the credibility of the person presenting the evidence and the case that is made using it. This involves a process of *interpretation* of information. Most decision makers appear to use rules of thumb or mental models when processing information presented to them. They will have in their heads, for example, what constitutes an unusual or significant figure in relation to mortality rates in a country, and will use this to gauge the significance of what they are hearing. They will also have a sense of what constitutes the appropriate response given a combination of different factors. Mostly these are implicit rather than explicit analytical models, and they tend to be highly individualized. More experienced practitioners have developed models that are very sophisticated, and they are able to "sift out" the relevant information from a mass of data presented to them.

BOX: The psychology of decision making

When faced with complex problems or incomplete information, rather than undertake taxing calculations, people tend to resort to simple educated guesses, "rule-of-thumb" thinking or personal intuition. Psychologists refer to these as "heuristics" (e.g., Gilovich,

Continued on next page

Continued from previous page

Griffen, and Kahneman 2002) or "biases." These tend to shape individual decision making in significant ways. One of the main challenges to promoting evidence-based decision making is to overcome inherent biases and habits of thought, and to allow evidence to challenge an individual's normal assumptions. This relates to the subject of incentives: an individual who is encouraged and rewarded for grounding decisions in evidence (or indeed penalized for not doing so) is more likely to challenge their own instinctive responses and to seek out relevant information.

Psychologists have identified a number of biases that commonly affect individual decision making. The "sunk cost" fallacy is one of the most troubling, where people fail to cut their losses and continue investing in clearly failing situations. This suggests that people who have invested time and money in something may have a strong tendency to continue to invest despite clear losses. As Teger (1980) suggests, people can find themselves with "too much invested to quit" and are reluctant to waste their effort. More generally, losses "weigh" more heavily with people than gains. "Prospect theory" posits that individuals are much more distressed by prospective losses than they are made happy by equivalent gains (Kahneman and Tversky 1979).

Decision making is highly affected by personality traits. One such trait is the individual's "need for achievement" (McClelland 1967). People who have a strong need for achievement tend to avoid both low-risk and high-risk situations. They avoid low-risk situations because it is easy to be successful in them and so a genuine sense of achievement is lacking. They avoid high-risk situations because they may not be successful and therefore will not gain the positive feedback they desire; or else the outcome could be attributed to chance rather than their own efforts. In the context of humanitarian decision making, this may well have a bearing on the use of innovative approaches (themselves inherently risky, since they are untried).

The analysis above refers mainly to decision making by individuals, and the way in which they take up and process information. But as noted, the scope for individual decision making is often highly constrained by previously established mandates, frameworks, priorities, and practices. In the same fashion, organizational decisions are generally not made in a vacuum: they are made partly with reference to what others are doing and to joint commitments, e.g., as part of Cluster or other coordination processes. Thus, the basis for decision making is multi-layered, complete with biases and assumptions, in addition to external evidence. The following sections illustrate this analysis with reference to case studies conducted in Ethiopia, Democratic Republic of Congo (DRC), and the Philippines.

2.3 Decision making in a conflict context: DRC case study

2.3.1 *The dominant political narrative*

There are many factors that shape the decision making process in regards to the allocation of funds and response strategies. Perhaps the most important factor that specifically shapes aid allocations for a given context is the prevailing political narrative about that context: the "story" that is told in policy terms about a given situation, its evolution, and related strategic priorities. This is hard to articulate precisely, and there will frequently be more than one narrative, but often it is possible to identify a dominant narrative, at least among the major donors. In the DRC case, this would seem to be one of "post-conflict" (or at least post-November 2011 election) transition and the need for stabilization interventions. However, the major threat to

stability remains the insecurity in the East, understood mainly in terms of insurgency against the state, to combat which UN (MONUSCO[5]) troops actively collaborate with government and Forces Armées de la République Démocratique du Congo (FARDC) under a Chapter VII mandate. Despite this, DRC remains relatively neglected by the major P5[6]/NATO actors and a relatively low global strategic priority.[7]

2.3.2 While the political narrative may largely shape the overall aid allocation and other aspects of international foreign policy, there is a humanitarian narrative within this that recognizes the protracted nature of the crisis and the continuing need for international humanitarian assistance and protection efforts. Yet, reflecting the political narrative, humanitarian aid has been declining for the past four years, with donor emphasis on stabilization, transitional programming and "resilience." The DRC Humanitarian Action Plan is now routinely underfunded,[8] yet humanitarian needs remain as high as ever. In fact, since late 2011, need in the East appears to have increased, with numbers of people displaced by conflict rising from 1.7 to 2.3 million. Elsewhere, levels of global acute malnutrition are chronically high, e.g., 19% in parts of Katanga. Areas outside the conflict zones of the East are relatively neglected, with political priorities playing a significant part in the geographic prioritization of aid.

2.3.3 ***Strategic analysis and planning***

Within the international humanitarian apparatus, strategic analysis and planning at various levels helps inform decisions about humanitarian priorities. In DRC, the annual Humanitarian Action Plan (HAP) provides a common strategic reference point for donors and implementing agencies, resulting from the collaborative analytical efforts of Clusters and agencies. The HAP is different from the standard Consolidated Appeal (CAP) in important ways, notably in that it does not consist of submitted projects but rather an analysis of context and needs by sector. Based on planning figures for anticipated "numbers affected," strategic objectives and targets are defined under various heads of activity. From this, Cluster budgets are drawn up, giving funding requirements by sector. Like the CAP, this is a *projection* of needs in the following year, based on the current year's analysis and apparent trends. So the 2012 HAP was written in August–October 2011. As with the CAP, there is a mid-year review process to allow for appropriate revisions.

The HAP is essentially a consolidated Cluster planning process, put together by OCHA based on discussions at provincial and national level across the various Clusters. While this has strengths, it also has some serious limitations. First, the HAP constitutes essentially a sector-based perspective on the context rather than a genuinely cross-cutting analysis. Second, it does not necessarily reflect the plans of the implementing agencies, although those plans are supposed to be drawn up to be consistent with the HAP. Third, it is a projection based on current priorities and anticipated needs but is not informed by thorough re-assessment of the context. In that sense, the HAP appears to provide only a relatively weak evidence base for current activities. Fourth, the link between budget and activities is unclear, and the basis on which different sectors/Clusters re-evaluate their funding requirements appears inconsistent and somewhat arbitrary.[9]

Thus, although the HAP avoids the problem suffered by the CAP of being a "wish-list" of individual agency projects, it risks irrelevance by being on the one hand too sector-specific and on the other, too broad and speculative. As one

[5] United Nations Organization Stabilization Mission in the Democratic Republic of the Congo (MONUSCO).

[6] The permanent members of the United Nations Security Council, also known as the P5, include the following five governments: China, France, Russia, the United Kingdom, and the United States.

[7] Since the case study was conducted, the activities of the M23 movement, the fall of Goma, and diplomatic pressures around the role of Rwanda in the insurgency have led to an increase in international strategic concern with the situation in the East.

[8] At the time this case study was conducted (July 2012), the HAP was only 35% funded.

[9] For example, in the 2012 HAP, discrepancy between the increase in funds requested by the Protection Cluster (100%) and by other clusters (e.g., WASH at 30%).

donor representative put it: "The HAP is still too vast of a document, you could find everything in it; it's not too helpful in identifying needs."[10] Perhaps its main value lies in the process involved in drafting it and the related series of consultations. This in itself helps promote convergence of thinking and joint prioritization. That said, it is notable that DRC government plans (at central, provincial, or district level) seem to feature only marginally in humanitarian decision making—although the government is involved in the HAP consultations.

Donors and agencies vary in the extent to which they use the HAP as a basis for decision making, but at most it appears to be only one factor in their thinking and not generally a decisive one. Nor does it necessarily reflect the strategic priorities of the various international bodies involved. Individual donors and agencies (or agency families) have their own processes of strategic analysis, planning, and prioritization. So for example, ECHO has their Humanitarian Implementation Plan (HIP), a framework for funding decisions; UN agencies and INGOs tend to have their own country and regional strategies, sometimes agreed across agency "families" (e.g., Oxfam International's "Joint Country Analysis and Strategy"). Some agencies limit their activities to specific areas, as with the International Committee of the Red Cross' (ICRC) "zones prioritaires," based on an analysis of humanitarian priorities and the agency's mandate.

Individual Clusters have their own more detailed plans, formulated at provincial level in the year of implementation using the HAP as a broad framework. Pooled funds from the Common Humanitarian Fund are allocated in line with these more "real-time" (and arguably more "real world") plans. Again, however, the question arises of consistency between these plans and those of individual agencies. *Contingency plans*—related to scenario-based analysis—are partly contained in the above plans, particularly at sector/Cluster level, and partly in the HAP process ("most likely scenarios").

Among the donors, ECHO has their particular own needs assessment process, which is as much a proposal vetting process as anything; as well as the annual HIP. ECHO places considerable emphasis on assessing the capacity of partners to deliver; and indeed the "evidence" they require to substantiate funding applications is substantially to do with the track record of the partner concerned. A respondent within ECHO explained, "All of our programs are monitored on the ground: we see if you had problems in the past and then we consider future funding. This gives us the ability to meet with partners to establish strategies for the following year." Partners are asked to submit a letter of intent before they submit full proposals and these must align with the HAP.

2.3.4 *Processes for diagnostics*

Given the protracted nature of the humanitarian situation in DRC, it is surprising that the 'diagnostic' aspects of practice (assessment, monitoring, evaluation, etc.) are not more developed than they are. All of those consulted for this study remarked on the patchiness of information available to humanitarian decision makers in DRC, particularly from areas outside eastern DRC. Some put this down to problems of scale: Equateur Province alone is the size of France. Access problems and rapidity of change in conflict areas were also cited as major constraints, with international agency presence varying greatly between the East (where it is high) and other areas where it is often very low. Existing tools such as WFP's Comprehensive Food Security and Vulnerability Analysis (CFSVA) were said "not to be usable in a context like this where reliable government data is not available country-wide".[11] Overall, assessment and monitoring were described as "weak' by OCHA particularly with regard to coordinated or joint processes; true impact evaluation was said to be "almost non-existent". Efforts are currently underway to strengthen this aspect of practice.

[10] Interview.

[11] Ibid.

Cluster performance on diagnostics and joint assessment is uneven, with a lack of methodological consistency reported within and between clusters, and between different provinces (including different thresholds for response)[12]. Clusters *are* beginning to use a MIRA-type tool for coordinated assessment led by OCHA. But probably the more significant mechanism is the UNICEF-led Rapid Response to Population Movements (RRMP), which acts as a first-phase assessment and response mechanism for new incidents of mass displacement. This is by far the most advanced coordinated assessment mechanism, though its application is limited to situations involving mass displacement.

The Integrated Phase Classification (IPC) tool, developed by FAO in the Somalia context, is used by the Food Security Cluster to create maps and related analysis of current and predicted food insecurity. This is not based on the full related process of monitoring and surveys that goes with the Somalia version, and some of those interviewed questioned the credibility of the process and the data on which it was based. Nevertheless, it does represent an important attempt at collective analysis that is lacking in some other sectors. Repeated roughly every six months, it is described by FAO as "a tool for the cluster" and something that "helps donors to select priority areas and helps to decide multi-donor pooling". Information coming from multiple sources—NGOs, governments, UN and others—starts at the provincial level and is then brought together at a national meeting.

Few examples were found of formal needs assessments by INGOs. Some, however, have gone to considerable lengths on diagnostics and developed monitoring systems of their own. Most notable in this respect is MSF Belgium which has developed an "antenna" system of health indicator monitoring, working closely with local health authorities. This is sensitive enough to pick up trends in morbidity that then warrant further investigation through more detailed needs assessment.

2.3.5 *Funding for assessments, monitoring and evaluation*

There is reported to be growing pressure from donors for better impact assessment. OCHA is trying to respond to this, and the head of OCHA DRC cited monitoring and evaluation as one of her top priorities – along with making the system more responsive to need.[13] Many commented that the humanitarian cluster system operates in development timeframes, not humanitarian/emergency ones. Thus, it was "too slow to do anything, including allocate funds".

Donors appear to be reluctant to fund assessment and other diagnostic processes; this is cited by OCHA and others as a major constraint on progress. Lack of funding for assessments and monitoring was a recurrent theme from agencies—particularly lack of funding for cluster-led processes. As one FAO representative put it "Food security programming in DRC costs $300m, but we can't even get $1m for cluster process and assessments".[14] One particular issue concerns donor funding for the IPC process. Funding for this—notably from ECHO—has dried up, so its future hangs in the balance. When questioned about this lack of funding, an ECHO representative said responded that "by the time we actually get the information from the IPC, it's too late". He felt that IPC was not a good example when discussing assessment and informed decision making.

Overall, the donor reluctance to fund assessments and monitoring appears at odds with their stated desire to improve the quality of evidence available to decision makers, as well as to determine the medium and longer term effect of humanitarian interventions. One DFID representative explains, "There has been a big push lately to get assessments done right and get good information. It's essentially about

[12] Interview.

[13] Ibid.

[14] Ibid.

finding out what is working…how long these effects last, what happens after the program is over, etc. The idea of 'get in and do a three-month intervention and get out' just doesn't work. There have been many broken and incomplete projects as funding dries up."[15] This was also expressed as the need to bridge the humanitarian-development divide.

2.4 Decision making in slow-onset emergencies: Ethiopia case study

Ethiopia was chosen as a case study to represent contexts of slow-onset emergencies, recurrent drought, and chronic food insecurity. Indeed, Ethiopia is such, but it also represents what response systems can look like under a strong and controlling state, whereby humanitarian action is very tightly controlled by the government. There is strong political pressure for whoever is in power to address the vast needs of the county's large population living in extreme poverty, particularly during recurrent times of extreme food insecurity. The goal is to prove to the world that Ethiopia is no longer the same country it saw on the news during the devastating famines in the 1970s and 1980s. The current regime tries to establish long-term credibility both through its response to vulnerability and through the promise that they are building a more economically vibrant state.

Humanitarian response has attempted to evolve from the previous pattern of periodic scrambles and hype for short-term, emergency famine relief to the current much more systematic approach. The aim is simultaneously to address chronic poverty and acute vulnerability. What was once a massive humanitarian response machine is slowly being reinvented as an integrated relief, welfare, and poverty reduction machine. The creation of the Productive Safety Net Program (PSNP), the largest social welfare program in Sub-Saharan Africa (outside of South Africa) is the main feature of this shift. The PSNP was established in 2005 and provides both cash transfers and in-kind food to food-insecure households in an attempt to bridge the predicted food gaps that arise when food production and other income are insufficient. Furthermore, the Risk Financing Mechanism is an instrument that allows for the PSNP to scale up in times of crisis in districts where it is already operational.

Because of the political imperative to present Ethiopia as a stable and economically growing country, any information that might indicate the contrary is sensitive. Information about hunger, malnutrition, disease, and other humanitarian needs is highly sensitive and tightly controlled by the government. Official needs assessments are conducted on a bi-annual basis through a multi-agency needs assessment methodology led by the government and supported by various UN agencies and NGOs. Teams are deployed and gather assessment information across the country, and the raw data is sent to the national government to be aggregated and analyzed by the government's Disaster Risk Management and Food Security Sector (DRMFSS). Discussions and negotiations then take place between national and regional governments, with DRMFSS giving the final approval and releasing the official figures. These figures are released twice a year (in January and July) in the Humanitarian Requirement Document (HRD)—a document very heavily food-focused. Needs are thus translated into beneficiary numbers, food tonnages essential to meet those needs, and financial requirements.

Because of the nature of this system, the framework within which humanitarian decisions are taken is highly pre-determined and tightly controlled. For agencies and donors, this means that their ability to respond on the basis of evidence—or even to generate such evidence—is highly constrained. The assessment results published in the HRD are the sole source of information that aid agencies are encouraged to quote in proposals and requests for funding. Primary data are not published, and the process by which the national government determines the final figures is opaque. Many individuals interviewed thought these unrepresentative of the actual situation on the ground. While there is no real way to verify the figures, questions are often raised; for

[15] Interview.

example, when the government reports large increases in crop production with simultaneously high numbers of people are relying on food aid. Furthermore, the government recently switched the assessment analysis process from being based on a "livelihoods threshold" to a "survival threshold." Lowering the threshold ultimately leads to smaller beneficiary numbers, leaving what some describe it as a "missing million" from what should be the beneficiary list.

The reality of the situation in Ethiopia is that the final figures have as much to do with a political balancing act as they have to do with needs assessment.[16] All agencies must go by these figures when designing programs and advocating for funding. No independent assessments are conducted. The system is thus highly reliant on the government's definition of need and assessed need, and ultimately its view on the appropriate type and scale of responses.

The food security crisis of 2011 demonstrated some of the factors at play. It is a measure of how the scale of the crisis took time to be become recognized that the projections in the government's HRD of people requiring assistance had to be revised twice: from 2.8 million to 3.2 million in April 2011; and then again to 4.5 million in July. As one evaluation of the response pointed out, "since agencies are only allowed to operate within the parameters set by the HRD, this had a constricting effect on early response, particularly given the long lead time for approval of projects" (DEC 2012). Some interviewed individuals felt strong pressure from donors, media, and headquarters to act immediately based on obvious need; however, they feared that if their agency moved before official HRD figures were approved by the government, they ran the risk of being expelled from Ethiopia all together. The raising of the official humanitarian requirement figures followed significant pressure from major donors, pointing to the fact that an independent "read" on needs can sometimes be essential as a counterweight to official thinking.

The case of Niger in 2005 is another example of this phenomenon: in that case, independent nutritional surveys conducted by Médecins Sans Frontières (MSF) alerted the international community to a food security crisis whose existence the government at the time tried to deny. It took considerable diplomatic effort by UNICEF, FAO, and others with government officials before a response could be agreed—by which time it was largely too late to mitigate the crisis (UNICEF Internal Report 2011).

BOX: Nutritional early warning and response in Ethiopia

As noted in the DEC evaluation of the 2011 crisis response, nutritional early warning and response systems are relatively advanced in Ethiopia. The government primary health care network and delivery systems cover most drought-prone areas, including Somali, Oromia, Amhara, and SNNP Regions. These have the capacity to tackle severe acute malnutrition through community-based management of acute malnutrition (CMAM). From the 2008 drought experience, there are over 8,000 health posts implementing CMAM in most of the priority hot spots in these four regions. Coordinated by the government's Emergency Nutrition Coordination Unit (ENCU) as cluster lead with support from UNICEF and other implementing partners, monthly information on admissions, recovery, referrals, deaths, and defaulters is collected with 80% reporting coverage, with about a month's delay.

Continued on next page

[16] A related point is made in the 2006 ODI report for WFP by Haan et al., A Review of Emergency Food Security Practice in Ethiopia. The authors note that "the current annual needs assessment [in 2006] is a mixture of non-systematic methods, and is ultimately based on negotiations between assessment teams, government agencies and international agencies." While the methods may now be more systematic, the point about negotiated results remains valid.

Continued from previous page

The evaluation records that "in March-April 2011, reports showed a 90% increase in admissions and led to UNICEF and implementing partners … to start the response using contingency/reserve funding and discussions with donors for support" (p. 16). It also notes that "most of those consulted felt that, thanks to early warning, existing safety nets and systematic response to problems like acute malnutrition, a potential catastrophe had been averted" (p. 4), though it records widespread human suffering, patchy response, and great damage to livelihoods, particularly those of pastoralists.

The report notes the significant time lag in the overall response to the crisis, attributed largely to the delay in getting government approval for new responses. It suggests that there is perceived need by many "for a shift of perspective from *crisis response* to *risk management*," something that has only partly been achieved to date. "A more general shift was felt by many to be required away from the use of *outcome* indicators such as nutritional data towards the use of *predictive* (risk) indicators as a basis for early intervention. The necessary complement to this was an agreed policy framework for early (preventive) intervention and specific funds to enable such interventions" (p. 32).

In terms of deciding which type of intervention to implement when responding to a crisis in Ethiopia, respondents suggested that the decision making process for most agencies consisted mainly of reviving the previous response plan with slight updates.[17] The greatest attention is spent on getting local and federal approval for the numbers of people to be assisted and actions being proposed. Because of the nature of crises in Ethiopia—chronic food insecurity, cyclical droughts—patterns of need are reasonably consistent and predictable, though the impact on individual provinces varies year to year, and needs in Somali Region in particular depend in part on patterns of insecurity and displacement, including inward migration of Somalis. Working in Somali Region is particularly sensitive and reliable information tends to be scarce, as it is for SNNPR[18]. One of the results of these factors is a notable gap in the willingness or ability of agencies to consider innovative approaches rather than repeating previous responses.

So what does the Ethiopian case study tell us about the use of evidence in decision making? In such an environment, the issues are less about the accuracy of the assessment data and the credibility of the related processes, more about the political factors that shape the resulting analysis. Effective international humanitarian action in such cases is about negotiation and working with the system as much as it is about obtaining technically correct data. Simply working to improve the monitoring and needs assessment processes—already quite strong in Ethiopia—will not necessarily translate into better decision making and improved assistance. What seems essential to success is the quality of the relationship with local and national officials, building trust, good negotiation and people skills, the use of informal networks (including with donors), and a strong understanding of the system. Those who can achieve this are in a much stronger position to ensure that, at least at more local levels, responses are in fact driven by evidence of need.

2.5 Decision making in fast-onset and recurrent crises: Philippines case study

The Philippines contrasts with Ethiopia and DRC in many ways. Crises in the Philippines consist of repeat natural disasters, including typhoons and monsoon rains, with their associated flooding and landslides. The country encounters approximately twenty typhoons

[17] Interview.

[18] Southern Nations, Nationalities, and People's Region (SNNPR) is a region in Ethiopia. It is one of the largest and most ethnically diverse regions in the country.

every year; in other words, crises are a regular aspect of life in the Philippines.

The state system is responsible for the vast majority of humanitarian needs assessment data gathering, as well as response efforts. Because of the repetitive nature of crises, state legislation plays a significant role in mandating action when crises occur throughout the country. For example, national policies exist whereby a minimum amount of funding (5% of the total funds) for local government must be used for disasters response. Most recently, a national policy requires that a substantial percentage of these funds allocated for emergency response must be allocated toward risk reduction and resiliency building. The goal of this new legislation is that there should be zero casualties as a result of lack of aid. The focus on resilience seems to have made a positive difference in beginning to establish good preparedness for disasters.

The role of the international agencies is described as "gap filling," supporting weaker aspects of the government's response through technical support and development, particularly at the high end of IT information systems, storm tracking and flood monitoring. The government has adopted a Cluster system that semi-mirrors the UN Cluster system: the two do coordinate, but the government mainly takes the lead. The national government must officially ask for assistance from the international community before significant amounts of outside aid can be provided. Any reluctance of the government to ask for help was described by an individual in the National Disaster and Risk Reduction Management Council to be a matter of national pride, stating, "We're not going to beg."[19]

The process for collecting data at the onset of a crisis is led by the government in a systematic process of routine procedures and established templates. All crisis response is monitored and coordinated by the National Disaster and Risk Reduction Management Council (NDRRMC), in coordination with various representatives from line ministries present in their operations room. Data collection takes places through the DRRM offices within the local governments in the affected areas using assessment teams of both officials and volunteers completing needs assessment forms. Each sectoral agency within the local government has unique templates to collect their relevant information. The information is then collated up to the regional and national levels.

The media were also frequently mentioned as having a strong role in providing timely information to the government and verifying the data collected by the assessment teams. Most local DRRM offices have representatives on the ground to verify things with the local authority, but in areas where this capacity is weak, the media are said to be very helpful. The media also have seats in the NDRRMC operations room, reflecting the fact that this triangulation of information is seen as very positive.

The tone of the interviews with both government and international actors regarding the collection of information and responding accordingly was generally positive, although several things were noted as "needing improvement." These included the process of collecting all data through paper and pencil, leaving room for error and slowing the process of aggregating information down. Compounding this problem, much of the information is collected by local volunteers with no particular training, leaving room for questioning the validity of information. But in general, interviewees from both government and the international community concluded that the quality of the information was fairly reliable and useful for planning.

Overall, the government's own data collection and response system is generally agreed to be reasonably solid (despite some weaknesses such as a lack of capacity to do quantitative statistical analysis at central level or to disaggregate data by sex and other key demographic factors). Because of the high frequency and repetitive nature of the crises, the information systems are

[19] Interview.

focused more on repeat reliability than on high resolution data. Experience responding to year after year of multiple crises as well as new preparedness measures have led to reliance on "tried and true" packages of response, rather than obtaining raw data and attempting to identify which response is appropriate among a variety of potential interventions. This approach seems fitting in this context because of the strong contextual knowledge, access and presence throughout the country, communication, and ability to triangulate information beyond the traditional needs assessment methods. All of this contextual knowledge, which precedes the emergency itself, allow agencies to act quickly and confidently based on the formal and informal information they have collected. International agencies, with their relative flexibility of approach, play an important supplementary role where local variations are not always well-catered to respond.

Most agencies reported that they were heavily reliant on government information, even relatively large, independent NGOs. One respondent from such an NGO said, "We start with the government figures, get information from other NGOs, and in some areas we will do a full assessment often with partners." Organizations often validate government information by sending its own assessment team. For example, often if a Barangay (lowest administrative level) is affected, the local government will declare that "all people" in that Barangay are affected. In situations such as these, agencies especially want to clarify the needs situation. The same respondent also suggested there was a "tricky balance between needs and politics when it comes to local assessment figures," particularly when an election was forthcoming. However, generally others felt that there was relatively little manipulation of data for political reasons.

While the government leads the analysis and response to crises, others play a part in coordinating assessments. The Red Cross, assisted by the International Federation of Red Cross and Red Crescent Societies (IFRC), has its own specialized assessment teams—though some reported that the Red Cross process was only weakly connected to the main assessment and information-sharing processes. The UN, coordinated by Office for the Coordination of Humanitarian Affairs (OCHA), leads rapid assessments within 72 hours. For this they use an adapted form of the MIRA[20] and secondary data analysis tools (still under development). A respondent from OCHA explained, "It's a joint assessment between government and international agencies. We use predominantly very simple data entry tools and a variety of personnel (it's mainly who we can find). We support the NDRRMC in doing that." Depending on the scale of the crisis and the request from government, a Flash Appeal may follow within a week.

Although government needs assessments were not highly criticized, some respondents expressed that data could be improved to reflect the humanitarian needs more accurately. As described by OCHA, government data "has an economic slant. It's typically a damage and loss assessment as opposed to damage and needs assessment." The government is working to supplement it with a humanitarian assessment, as encouraged by the international community. The role of the Clusters was felt to be important in building confidence in government data and its accuracy. Others were more forthright, particularly about information from local government. A respondent from a large NGO felt that government information coming up the system "is more of a wish list than data, so the validation process is really important […] local NGOs that are parts of national networks are used to validate the data." The Department of Health and church organizations were singled out as good performers in data collection to help validate information.

How does the system in the Philippines compare with the other two contexts analyzed above? It has some features in common with Ethiopia, both in the recurrent nature of crises and the lead role played by government.

[20] Multi-Cluster/Sector Initial Rapid Assessment.

However, in the Philippines the international agencies play a more obviously "auxiliary" role in relief assistance and one that is secondary to the government, while in Ethiopia the international actors are central to the implementation of government-approved programs and the entire "productive safety net" system. What characterizes the Philippines is the integrated system linking local to national government and the multiple avenues to validate information. For all its flaws, it has proved an effective system for the provision of relief and recompense for damage to cyclone victims. It faces significant challenges in meeting its target of "zero casualties due to lack of aid," but perhaps faces greater challenges in reducing vulnerabilities and preventing the need for relief in the first place. ■

3.1 Introduction

A variety of types and sources of information and evidence are used by humanitarian decision makers. These can be divided into three main categories:

(i) Pre-crisis contextual information (about capacities and vulnerabilities, livelihood patterns, etc.);
(ii) Information concerning the nature of an evolving crisis (e.g., from early warning and assessment data) and the impact of the response to it (monitoring, evaluation); and
(iii) Evidence about "what works" in response to particular kinds of crisis, including best practice, standards, protocols, etc. (ODI 2009).

In practice, category (ii) tends to dominate, with category (iii) being the domain largely of "experts" and specialist advisers. What appears to be critical is the link between the situational analysis, largely informed by (i) and (ii), and the response analysis, largely informed by (iii), but needing to be adapted to the context. In practice that link is not always clearly articulated.

Information is not in itself evidence: it is only "evidential" to the extent that it supports a given hypothesis or proposition; in other words, where it is evidence *for* something. Related to the categories of information above, it is possible to distinguish three main proposition types in the humanitarian sphere:

- That an actual or imminent crisis exists;
- That a given form of response will be (or has been) effective in preventing or mitigating the worst aspects of this crisis;
- That a given form of response is the most appropriate in context, in view of effectiveness plus other factors: alternative response options, local preferences and responses, feasibility, standards, cost, etc.

It is largely in relation to these three types of propositions that humanitarians are challenged to demonstrate an evidential basis for their claims. Our focus here is on the kinds of evidence that relate to response decisions—and related evaluation questions—rather than to policy formulation based on more general propositions about "what works." In the following sections we review some of the main ways in which evidence is currently generated in the humanitarian sector, and the ways in which that evidence is used (or not) by decision makers.

Table 1 reproduced below lists the kinds of process by which information of different kinds is produced, some of which we review in more detail below. In practice, these processes rarely form part of one information "system:" they tend to be fragmented and disconnected, with different actors conducting their own processes (assessments, evaluations, etc.), the results of which are not always shared. With the evolution of the Clusters and with OCHA playing a more pivotal role as an information hub, this situation is changing. But it remains true that even in such established humanitarian contexts as DRC, the various processes involved remain under-developed and poorly integrated.

Table 1. Humanitarian information-generating processes

Component	Information	Links to Response
1) Baseline vulnerability and poverty assessment (BVPA)	• What are the basic livelihoods of groups? • What are known or likely hazards: natural and environmental, social, economic, and political? • Who are the most vulnerable groups? • What capacities, services, and resources (physical, human, social) exist to mitigate vulnerability? • What are coping and risk minimization strategies?	• Long-term development/ vulnerability reduction planning • Emergency preparedness planning • Mitigation planning • Community-based preparedness activities

Continued on next page

Continued from previous page

Component	Information	Links to Response
2. Early warning (EW)	• Indicator trend analysis: Is there a problem shaping up? • Where and how quickly is it developing? • What are the geographic dimensions of the problem? • In what areas should an in-depth assessment be concentrated?	• Activate and focus needs assessment • Contingency and scenario planning • Activate mitigation plans • Geographic targeting
3) Emergency needs assessment (ENA)	• What are the nature, dimensions, and longevity of the problem? • Who are the most vulnerable groups? • What and how much is needed; what is the best response? • How is local coping capacity and provision of services overwhelmed? • What are major logistical and resource considerations?	• Detailed emergency response plans and programs • Detailed targeting • Mobilize resources • Mobilize public awareness
4) Program monitoring	• Are program input and output goals being met? • Are targeting goals addressing need and being met? • Have needs changed since original assessments? • How might program changes need to take place?	• Adjust inputs or logistics • Adjust targeting • Adjust pipeline
5) Impact evaluations	• Is the intervention achieving the intended result? • What adjustments are necessary (response, quantity, targeting)?	• Increase or decrease levels of delivery • Change targeting criteria • Change activities
6) Context monitoring (CM)	• What are the possibilities for exit, recovery/transition? • What are institutional capacities and vulnerabilities? • Does the situation require re-assessment?	• Transition to rehabilitation/ development programming. Re-assess situation • Institutional capacity building
7) Program evaluation and lessons learned	• How can the overall program be improved? • Are humanitarian principles being upheld by program? • What lessons can be learned from experience and mistakes?	• Improvements to overall system: preparedness, response capacity, program, protection, basic needs and rights

Source: Maxwell and Watkins. (2003). Humanitarian Information Systems and Emergencies in the Greater Horn of Africa.

3.2 The generation and use of evidence from early warning systems

In the mid-1980s, a lack of information to predict slow-onset emergencies was widely seen as the biggest constraint to preventing famine in Africa. To solve this problem, large investments have gone into developing and scaling up early warning (EW) systems, whereby data are collected to provide timely notice of imminent crises with the goal of prompting responses *prior* to crises reaching severe levels (Walker 1989). This has seen the emergence of "global" food security systems like FEWS NET and regional systems in southern Africa and the Sahel, together with predictive food insecurity mapping processes like the Integrated Phase Classification, which originated in Somalia. There is little doubt that these systems have helped inform responses that have prevented famine and some of the worst aspects of food insecurity. But despite these efforts, the pattern remains of responses seen as "too little, too late" (Buchanan-Smith and Davis 1995; Buchanan-Smith 2000; Levine, Crosskey et al. 2011; Oxfam 2012). The food security crisis in the Horn of Africa in 2011 (and to some extent the Sahelian crisis that followed) marked the latest chapter in this history of warnings that were not acted upon. While the famine that occurred in South Central Somalia was largely attributable to political factors and access constraints, the lack of concerted action in response to clear warnings from the early warning systems (FEWS NET and Integrated Phase Classification (IPC)) contributed to the failure to prevent famine and provide timely relief (IASC 2012).

In the case of the African food security early warning systems, then, the main issue is not identified as being a *lack* of evidence. Rather, the inability or unwillingness of decision makers to translate information into timely and appropriate responses appears to be the main problem (Buchanan-Smith, Davies et al. 1994; Buchanan-Smith 2000; Levine, Crosskey et al. 2011). Other EW systems, such as those involving cyclone tracking in countries like Bangladesh, Philippines, and the Caribbean have been more successful in that sense. This is in part because they are government-run and closely linked to preparedness and response mechanisms at the local level, with high degrees of community involvement. Getting timely information to people and helping them take avoiding action is key in relation to these "rapid-onset" disasters. Flood early warning systems, as shown in the case of Pakistan, are generally rather less well developed.

Factors that influence the EW decision making process involve a wide range of financial and bureaucratic issues, political factors, and institutional relationships within and between donors, recipient governments, and NGOs. Overlapping EW systems by different stakeholders (national government, NGOs and UN agencies, and donors) can also lead to contradictory information, causing confusion among decision makers and delaying response. The absence of a single and consistent message is a clear hindrance to timely decision making. Donors still rely more heavily on EW data from UN agencies and their own EW systems than information from national governments (as pointed out by Buchanan-Smith, Davies et al. in 1994). The reliability of the data, or "who owns it," is undoubtedly influential for how the data is viewed and used (Buchanan-Smith 2000).

The way information is presented is important. Providing EW indicators in a simple, straightforward manner, such as a summary analysis directly linked to response interventions, is proven to be more effective than "bald" statistics (Thomson, Jenden et al. 1998). Donors and other decision makers tend to focus on outcome indicators, using "crisis" thresholds (primarily relating to mortality and acute malnutrition) and related indicators as triggers for response—thus defeating the purpose of an EW system, which is to inform prompt *early* action. Hesitancy among donors is attributed to the pressure to ensure limited resources are used efficiently, hence a reluctance to act pre-emptively to address a crisis that may not transpire, at least in the ways predicted (Levine, Crosskey et al. 2011). There is also a reluctance to trigger relief mechanisms that may interfere with more developmental efforts to promote resilience and vulnerability reduction. One approach to addressing this problem (pioneered by USAID/OFDA) is to include "crisis modifiers" in existing development programs,

allowing such programs to adapt according to circumstances and meet critical needs as they arise. This is gaining acceptance with other major donors (including ECHO) and the agencies they support (Hillier and Dempsey 2012).

Clearly, evidence from EW systems does not automatically lead to better decision making. How then do we overcome these large systemic barriers in order to improve decision making and ultimately improve response? Recent research conducted by Levine, Crosskey et al. (2011) offers solutions. First, the use of EW information to produce credible predictions must be coupled with improved agency and government preparedness, including established triggers and specific plans for action rooted in livelihood analysis. Secondly, funding mechanisms must be established that are responsive and provide fast, flexible support to livelihood protection. Third, long-term development programs should incorporate "crisis modifiers" and response mechanisms. But the most challenging problem is surely to overcome the political and institutional barriers that appear to make it so difficult to shift from an approach based on crisis response to one based on risk management. Recent donor and agency rhetoric on the need to resolve these issues is encouraging.[21]

3.3 Evidence from needs assessments

The purpose of needs assessments is to identify critical threats (short and medium term) to the well-being of crisis-affected populations in order to inform appropriate external responses and quantify the related resource requirements. This process, in theory, provides the necessary platform for decision makers to allocate resources and design appropriate responses in a timeframe appropriate to the urgency of the situation. However, a plethora of criticisms has been clearly and repeatedly stated by practitioners and researchers regarding needs assessment tools, processes, and systems (UNHCR, WFP et al. 2000; Darcy, Griekspoor et al. 2003; Darcy and Hofmann 2003; Mock and Garfield 2007; Bradt 2009). These include:

- Lack of uniform definitions around common terms;
- Objectives being limited and unclear;
- A blurring between identifying needs and filling perceived "service gaps;"
- Tendency to be conducted by operational agencies in order to substantiate a request for funds—introducing inevitable bias into the process;
- Missing vital information regarding wider contexts (including political, social, and greater economic factors);
- Seen as supply-driven and a "front-loaded" process;
- Disjointed across sectors and poorly coordinated, leading to contradictory information;
- Overlapping in geographic coverage while missing other areas;
- Poor data reliability due to inadequate institutional capacity and poor methodology;
- Too much data collected that go unused;
- Too little or incorrect data collected to demonstrate true needs;
- Data being specifically catered to justify a project design and donor proposal;
- Presenting outputs inadequately to decision makers;
- Results being too slow to drive humanitarian response.

In September 2006, Dartmouth Medical School and Harvard University co-sponsored a multi-agency conference to examine humanitarian health issues among 51 organizations. A Working Group established to discuss evidenced-based decision making concluded that available information was frequently characterized as unreliable, not credible, and "not in a format that easily allowed decision-makers to make informed judgment about appropriate responses" (Mock and Garfield 2007, p. 380). Among the data available, frustration was expressed around the absence of primary data and insufficient attempts to strengthen capacity for analyzing data. Respondents stressed that information for decision making must be available and catered to program managers in the field.

[21] See for example DFID's response to the Humanitarian Emergency Response Review commissioned by the UK government, with its emphasis on anticipation: http://www.dfid.gov.uk/Documents/publications1/hum-emer-resp-rev-uk-gvmt-resp.pdf.

BOX: Who defines "need"?

Some have questioned the whole basis on which "needs" are assessed. As Darcy and Hofmann point out (2003), "assessing needs" is an ambiguous concept. Who or what defines need? If taken to mean a deficit or gap of some kind, particularly a gap in goods or services, then responding to "need" invites supply-driven responses aimed at filling that gap. This interpretation also depends on a logic that suggests that need does not arise until a catastrophic deficit occurs, which fails to explain the humanitarian case for preventive action. Darcy and Hofmann propose an alternative view of needs assessment based on risk and outcome analysis: "A good assessment process combines both [outcomes indicators and risk indicators], correlating evidence about outcomes with evidence about risk." (p. 45).

Whether or not one accepts this view, the concept of "demonstrating need" is not a straightforward one, and what constitutes "evidence of need" may depend on the observer. In particular, the external view may be at odds with that of those actually experiencing the crisis. The extent to which the views of crisis-affected people are used as evidence to inform response is variable. New technologies—in particular mobile phones—are making feasible new methods of consultation with affected populations on an ongoing rather than just a one-off basis.

There have been a number of initiatives in the past decade that aimed at improving the way needs are assessed in order to facilitate better-informed decision making. These include:

- HNTS: The Health and Nutrition Tracking Service (HNTS) at WHO was established in 2007 with the aim to provide impartial, credible, and timely information on mortality and nutrition rates through standardized data collection and analysis methods. The function of HNTS is to provide technical support for existing efforts in the areas of collecting, analyzing, and disseminating data (WHO 2009).
- SMART Project: The Standardized Monitoring and Assessment of Relief and Transitions (SMART) Project sought to improve data collection in three critical data areas of mortality, nutrition, and food security. The goal was to create more reliable and consistent data in a rapid and accessible manner for policy and resource decision making. In 2006, a manual on SMART methodology was published to standardize the measurement of indicators. Since then, the SMART methodology has been adopted for several country-wide programs by UNICEF and other NGOs (Golden, Brennan et al. 2006; UNICEF 2011).
- SENAC Project: WFP's Strengthening Emergency Needs Assessment Capacity Project (SENAC) objective was to reinforce WFP's capacity to assess humanitarian needs in the food sector. This has led to the development of the Comprehensive Food Security and Vulnerability Assessment (CFSVA) manual (WFP 2009a) and the Emergency Food Security Assessment (EFSA) manual (WFP 2009b), which have been used widely in both baseline and emergency assessments across many countries. The SENAC project has reportedly led to significantly improved assessment as well as more consistent and transparent reporting (IFRC 2011).
- IPC Tool: The Integrated Phase Classification (IPC) tool was developed originally by FAO in Somalia but expanded to being applied in over 12 countries. IPC is an analytical tool that establishes common language in order to compare different crises and helps guide decision makers regarding the severity of each crisis (FAO 2008).
- IRA Tool: IASC Nutrition Cluster developed the tri-Cluster Initial Rapid Assessment (IRA) tool in coordination with the WASH and Health Clusters. The IRA tool was designed to guide appropriate

analysis of data collected in order to make essential decisions on immediate response, funding, and/or follow-on assessments. This tool consists of a 37-page checklist that addresses the basic first questions to be asked in the aftermath of a sudden-onset emergency (IASC 2009).

- National Level Initiatives: There have also been efforts on national levels that made advances in improving assessments, such as the Ethiopian Nutrition Cluster Unit in Ethiopia, which helped to standardize approaches in national-level training and in widely available tools and guidelines (IFRC 2011).

The extent to which these initiatives have led to better informed and more evidence-based decisions remains unclear. Many of the initiatives have been criticized as "supply-side" in orientation, and it has been said that it is difficult to interpret and translate data into decision making (Mock and Garfield 2007). A lack of standard indicators and well-defined key terms (such as "humanitarian need," "evidence," "evidenced-based," "rapid assessment") has been noted as a clear barrier to coordination, comparability, and overall improved response (Mock and Garfield 2007; OCHA 2009). IASC's ACE project review of various assessment initiatives and analysis frameworks in 2009 determined these systems contained an insufficient amount of essential information at the early onset of a crisis, significant overlaps in data, and no core set of indicators to improve comparability in measuring needs (OCHA 2009).

3.4 Coordination of needs assessments

In the past (and still today), most agencies had their own nonstandardized survey forms and assessments that often produced conflicting or repetitive results, with very little discussion or coordination between agencies. Agencies also tended to assess situations directly to support individual programs, rather than identifying an aggregate need. An evaluation of the role of assessments in the 2004 Tsunami response found that no cross-sectoral humanitarian needs assessment covered all affected areas in any single country and wider geographical coverage was confined to siloed sectoral surveys (de Ville de Goyet and Morinière 2006). This gap in coordination and neglect to identify a comprehensive need led to a clear call for more joint agency—and multi-sector—approaches to assessments. Thus, the development of common needs assessments (CNA) resulted.

A CNA is defined as "a time-bound, multi-sectoral, multi-stakeholder process of collecting, analyzing, and interpreting data to assess needs and inform decisions on humanitarian and early recovery responses" (Garfield, Blake et al. 2011, p. 3). The purpose is to involve more than one agency in a joint strategy for primary data collection and analysis in a coordinated fashion. By combining financial, human, and physical resources, CNAs aim to generate more comprehensive and timely information for agencies to use in decision making.

Research conducted by Garfield et al. (2011), in a review of recently implemented CNAs, concluded that while there are several potential advantages for CNAs, many challenges and weaknesses still remain (Table 2).

Table 2. Advantages and disadvantages to CNAs

Advantages

- Efficiency (finances, use of resources, less duplication of efforts)
- Timeliness (info. on multiple issues can be collected simultaneously)
- Shared learning (by designing and analyzing together, consensus on humanitarian needs)
- Coherence (balanced inter-sectoral picture for better targeting)
- Coordination across agencies (more likely to coordinate during implementation)
- Effectiveness (define needs early in recovery and better target resources)
- Shared quickly and early program planning (common idea of severity of situation, areas and populations of greatest need, trends for future, gaps, contextual factors, coping strategies, etc.)

Continued on next page

Continued from previous page

Disadvantages

- Results are too late to influence funding decisions
- Extremely expensive
- Results are unclear and too complicated to be used
- Too many goals of different stakeholders
- Excessive focus on quantitative, survey-based data
- Lack of timely analysis and lack of clarity on how to analyze
- Too many tools to choose from and staff are unfamiliar with them
- Excessive focus on quantitative, survey-based data
- Lack of agreement on a common set of indicators
- Delays in implementation due to heavy guidance from headquarters, inadequate skills in field, and rivalry amongst agencies for appropriate processes

Source: Adapted from Garfield et al. (2011). Common Needs Assessments and Humanitarian Action.

3.5 Response analysis: The use of evidence in choosing the type of response

In the past, responding to crises consisted of a small number of prepackaged interventions, thought to be the viable options for responding, yet not based on much evidence (Levine and Chastre 2004). Major changes have occurred in the humanitarian sector over the last decade or so, particularly in the food security sector, whereby a wider range of response options has developed beyond the traditional—or so-called "tried and true"—interventions. The use of new modalities of delivering food aid has grown dramatically with the untying of food aid funds towards purchasing food locally (within the recipient country) and regionally (within neighboring countries). Also, greater emphasis has been placed on cash and voucher programs following the response to the 2004 Indian Ocean Tsunami. Newly created nutrition products have multiplied the possibilities for preventing and treating malnutrition. Furthermore, livelihood support has greatly expanded from the traditional seed and tool distribution to also include a wide array of asset protection, livestock management, and other actions aimed at resilience building.

With such wider variety of response choices relating to food insecurity, the decision making process has also grown. Maxwell et al. (2012) recently conducted research on the process of response analysis, defined as "the analytical process by which the objectives and modality of programme response options in an emergency are determined, and potentially harmful impacts are minimised." Their research shows that indeed very few agencies conduct a formal or structured analysis on the various response options and few base their response choice on evidence that points to the most appropriate response. Ultimately, the decision does not always involve an evidence-based, analytical process (Maxwell et al. 2012).

In fact, assessment evidence has been found to play a marginal, or even negligible, role within agencies (Darcy and Hofmann 2003; de Ville de Goyet and Morinière 2006; Darcy, Anderson et al. 2007; ODI 2009). Maxwell et al. (2012) also identify a number of other factors that are significantly influential in the decision making process, specific to program choice. These include: the capacity and the organizational ethos of the implementing agency; the personal experience of program staff, and a range of external factors, including: donor resources and policy; government policy in the recipient country; media and political influences; the costs of reporting and compliance associated with different resources; the capacity of partner organizations; assumptions about the risks associated with different responses; and a variety of other factors. Unfortunately, an analysis of empirical data is by no means the prominent determining factor in comparison with other factors such as biases, assumptions, political pressures, etc.

Experienced humanitarian staff tend to base decisions mainly on past experiences, instinct, and assumptions. Even when assessment is viewed as a priority for program planning, agencies often violate their own calls for field-validated assessments as a precursor to intervention—often justified on the grounds that it is impractical to wait for a formal analysis to be conducted before choosing an intervention. Expert knowledge from past experiences is certainly crucial for decision making. However, a process where decisions are made without the analysis of evidence may also lead to a pattern of simply justifying whatever is most convenient for the agency or donor. Thus, agencies may fall into a "programmatic inertia" whereby certain types of programs will inevitably be chosen due to individual biases, assumptions, and preferences. This in turn leads to building agency capacity around these interventions, which continue to be the "preferred response" with each new crisis, irrespective of available evidence. As one key informant stated, "people become specialized in something and it becomes more and more difficult to have an open mind and look at other responses options."[22]

3.6 Needs assessment and response analysis tools

In addition to large initiatives aimed at improving humanitarian response through better collection of evidence and achieving better decision making, many efforts have gone into creating improved tools. A wide variety of tools were created—some focusing on specific sectoral assessments in order to attain detailed needs, while others have been designed to be cross-sectoral in an attempt to capture a more holistic picture of the situation.

Table 3. Assessment tools

Joint/common assessments	
PONJA: Post-Nargis Joint Assessment (Myanmar)	
RINAH: Rapid Initial Needs Assessment in Haiti	
McRAM: Multi-Cluster Rapid Assessment Mechanism (Pakistan)	
MIRA: Multi-Cluster/sector Initial Rapid Assessment	
Multi-Cluster/multi-sectoral	
Health, Nutrition, and WASH	IRA: Initial Rapid Assessment
IFRC-Fact Team	Rapid Assessment
UNHCR-WFP	JAM: Joint Assessment Mission
Cluster-/sector-specific	
CCCM	CCCM Assessment Framework
Education	Integrated Rapid Assessment Field Data Checklist RALS: Rapid Education Assessment of Learning Spaces

Continued on next page

[22] Interview.

Continued from previous page

Emergency Shelter	LENSS: Local Estimate of Needs for Shelter and Settlement Risk Mapping and Shelter Response Planning
Health	HeRAMS
Nutrition and Food Security	EFSA: Emergency Food Security Assessment CFSVA: Comprehensive Food Security and Vulnerability Analysis HEA: Rapid Household Economy Analysis MICS: Multiple Indicator Cluster Survey
Protection	UNCHR's Global Assessment of Needs Rapid Protection Assessment Rapid Child Protection Assessment IDP profiling
WASH	RAT: Rapid Assessment Tool CAT: Comprehensive Assessment Tool

Source: OCHA. (2009). Assessment and Classification of Emergencies (ACE) Project; Garfield et al. (2011). Common Needs Assessments and Humanitarian Action.

The MIRA tool[23] developed by OCHA and the IASC Needs Assessment Task Force[24] is now gaining recognition as the preferred tool (often with local variations) in conducting initial, cross-sectoral assessments in rapid-onset disasters. While the results await proper evaluation, this appears to mark a significant improvement on the earlier Needs Analysis Framework developed by OCHA. Crucially, it begins to address the issue of integrated analysis across different sectors, although the analytical model for synthesizing results across sectors remains unclear.

Other tools and guidelines have been created to help facilitate appropriate decision making once the needs data is collected (Table 4).

Table 4: Response analysis tools

Market analysis tools
EMMA (Emergency Market Mapping and Assessment) WFP MAF (Market Analysis Framework) FEWSNET Structure-Conduct-Performance Tool
Livelihoods sector-specific tools
PRIM (Participatory Response Identification Matrix) within LEGS SSSA (Seed Security System Assessment)

Continued on next page

[23] Multi-Cluster/Sector Initial Rapid Assessment.

[24] See http://www.humanitarianinfo.org/iasc/pageloader.aspx?page=content-subsidi-common-default&sb=75.

Continued from previous page

Nutrition sector-specific tools
WHO Decision Chart for Implementing Selective Feeding Programs
WFP Decision Tree for Response Options–Nutrition Intervention Food Products
FAQR Decision Trees (In *Improving the Nutritional Quality of US Food Aid*)
Global Nutrition Cluster's MAM A Decision Tool for Emergencies
Modality-specific tools
MIFIRA (Market Information for Food Insecurity Response Analysis)
Good Practice Review (GPR) Cash Transfer Programming in Emergencies
ECHO Decision Tree For Response Options
Save the Children Risk Assessment Tool
ACF Food Security and Livelihoods Assessment Guidelines
ICRC Global FSA Guidelines, ICRC Guidelines for Cash Transfer Programs
ACF Implementing Cash-based Interventions
Harm-mitigation tools
CARE Benefits/Harms Analysis Tool
Do No Harm
Preventing Corruption in Humanitarian Operations
Process/consensus-oriented tools
RAF (Response Analysis Framework–FAO)
RAP (Response Analysis Project–WFP)
Oxfam Response Analysis Guide

Source: Adapted from Maxwell, Parker, and Stobaugh. What Drives Program Choice in Food Security Crises? (forthcoming)

The acceptance and use of such tools varied greatly among respondents. Although many were aware of some of them, these tools are not as frequently used. Maxwell et al. 2012 identified reasons for such lack of use. First, many practitioners agreed that there are too many tools from which to choose, and it is unclear which are used for what decision. Second, many of the tools are complex and require technical expertise and too much time and too many resources to use. Some respondents in this study felt that tools need to become standardized before they are of any use to decision makers. Yet others thought very much the opposite, expressing disdain for the creation of more decision making tools due to the fact that such tools "remove the ability to think."

3.7 Changing circumstances and new evidence

Almost all donors require some sort of assessment to be conducted in order to support the proposed intervention. But how well does this data correspond with what is actually occurring on the ground at the time of intervention? Although there have been many attempts to reduce the delay in responses through pre-positioned commodities, more flexible funding mechanisms, and early warning, there is still significant time lag between assessments, decision making, and actual implementation. In some cases, there can be a six-month time lapse from when an assessment is conducted to when implementation begins.

Because of such gaps between assessments and implementing, and because situations continue to evolve while programs are being implemented, ongoing assessments and monitoring of frequently changing environments is vital to ensure that interventions are in fact addressing needs as they evolve. However, monitoring information typically focuses on the inputs and outputs of project management, such as supply-chain and commodity tracking, rather than focusing on assessment of the external environment and changing nature of risks (Darcy and Hofmann 2003; Maxwell and Watkins 2003). "Snapshot" surveys at the front end of programs are much more common than any continued surveillance of needs (Darcy, Griekspoor et al. 2003). Little to no ongoing monitoring of needs takes place (de Ville de Goyet and Morinière 2006), and the link between initial assessments and decision making grows weaker throughout the life of a response (Darcy, Anderson et al. 2007).

Using evidence from ongoing assessment is sorely lacking—an obvious gap that is crucial for evidence-based interventions, particularly in protracted crises. However, the answers given by respondents in this study suggest that even if agencies are conducting monitoring, their ability to change the original type of intervention after it has begun is strictly limited. Several individuals interviewed discussed multiple occasions where they had concluded that the originally proposed intervention was no longer appropriate when it came time to implement, as so much time had passed since the original assessment. Yet despite knowing that circumstances had changed, they were unable to adapt the program. The reasons for this, including institutional inertia and the sheer demands of managing and implementing programs, are beyond the scope of this study. But it suggests that there may be little incentive to monitor programs other than as strictly necessary for accountability purposes. This goes to the heart of the problem about the generation and use of evidence: unless there is some clear reason and incentive to do so, responses are likely to remain only weakly grounded in evidence.

3.8 The use of evidence in innovative responses

In most industries, innovative thinking is viewed as an asset, encouraged and rewarded in order to achieve greater success. However, in the humanitarian realm, innovation is quite often a difficult sell, mainly due to numerous actors with different objectives, limited funds, and the belief that a "trial and error" mentality is ethically unacceptable. Rather, the tendency is to stick to the traditional interventions, whether or not the evidence supports them. As one key informant noted, "there is a belief that because something makes sense intuitively, it has to be right … there is a lack of willingness and incentive to change."[25]

In fact, with most innovative ideas, the humanitarian community requires evidence to show that a new intervention is more appropriate and effective than traditional approaches (even though the same evidence isn't necessarily required to prove that the traditional approaches are appropriate). Maxwell et al. (2012) conducted analysis on five of the largest donor agencies' funding requirements for proposed food security interventions. They found that organizations are required to demonstrate a larger body of evidence to support "innovative" food assistance approaches than "traditional" programs. For example, in order to justify the use of cash-based interventions or local/regional procurement of food assistance (as opposed to the traditional trans-oceanic delivery of in-kind food), most donors require agencies to provide a market analysis that ensures such interventions are appropriate and do not cause harmful side effects. Some donors require that agencies also provide a more comprehensive risk assessment or management plan to prove this. With cash interventions, most require agencies to provide analyses of safety, corruption, and protection implications. Such analyses are not necessarily required for the more traditional interventions.

Requiring evidence to support proposed interventions is good in theory; however, when certain interventions require considerably more evidence than others, it influences decision

[25] Interview.

makers' choice of intervention. Several of those interviewed agreed that they often automatically choose the intervention that has fewer requirements in terms of necessary analysis. This cost of complying with the many requirements for innovative approaches persuades decision makers to choose responses based on the amount of work required to propose the intervention, rather than the one that is most appropriate for the crisis. Even when agencies desire to gather the necessary data and evidence in support of innovative approaches, restrictions on time and access prevent them from doing so. One respondent states, "We know we can't get in quickly enough to get information on the markets. So we'll end up doing in-kind food assistance because we simply can't prove the functionality of markets soon enough to support any other interventions."[26] This example is one of many. It is quite ironic that cash is considered "more risky," since an injection of in-kind food aid into a community can have many of the same harmful effects (disruption of markets, security issues, theft, etc.). Yet in-kind food aid requires significantly less evidence to justify its use.

Despite an overall reluctance to adopt new approaches in the humanitarian sector, the presentation of supportive evidence can significantly shift the acceptance of such innovative responses. The most prominent example of this is the body of evidence gathered by Valid International and other agencies that revolutionized the nutrition sector. Dr. Steve Collins presented evidence to prove the effectiveness and efficiency of community-based management of malnutrition. He published his research in August 2001 in the prestigious scientific journal, *The Lancet*, (Collins 2001). He later provided further evidence to major humanitarian actors at a symposium in 2003, which helped to spread the knowledge of the new evidence-backed approach. This evidence sparked further research that demonstrated the large-scale effectiveness of the innovative nutrition programming, eventually winning the support of key players involved in tackling malnutrition (WHO et al. 2007).

Beyond this example, the evidence base proving humanitarian responses to be effective is extremely lacking. Data regarding the comparative value of different types of responses and which are the most effective are missing within the humanitarian community. A strong body of evidence, such as that behind the CMAM nutrition programming, is not well established for other interventions such as food security programs, livelihood approaches, etc. regarding the impact these and other approaches have on saving lives and building resilience.

[26] Interview.

4.1 Overview

Overall, the study revealed high levels of diversity in the contexts for decision making as well as in the use of information and analysis. Some patterns emerge, however. In those contexts where strong governmental systems exist, the generation and use of information is either highly controlled by government (Ethiopia) or else is dominated by government-led systems (Philippines), with international actors playing only an auxiliary role. Most of the key decisions regarding resource allocation are in effect made by local and national government officials in these cases, on the basis of national or regional plans. Domestic political factors represent a significant potential bias which risks distorting the data available. That said, there are checks in most systems. In Ethiopia for example, international actors partner with central and local government in both the assessment of need and the provision of relief. While political bias may affect which areas are prioritized for relief, major discrepancies between assessed and "stated" need are hard to disguise, and the larger international donors have a substantial influence over the recipient government in this regard. Thus, although the validity of the published data may be questionable (Ethiopia), the process of micro-resource allocation and program design is able to a substantial degree to iron out some of the more obvious anomalies at the local level.

In contexts where government is relatively absent from humanitarian decision making, a different set of factors are at play. In the most extreme cases (Somalia, Eastern DRC, parts of Afghanistan), government systems are almost completely absent or bypassed by the international system. Here the dominant political narrative is an international one, and it provides the backdrop for macro-resource allocation decisions. The biases in these cases come as much from pre-determined international strategic priorities as from domestic factors: aid is provided as much in proportion to an area's strategic significance as it is based on assessed need. This is evident in the ebb and flow of funding in response to annual appeals (CAP, etc.), which fluctuates more according to foreign policy agendas like counterterrorism and stabilization than it does according to apparent need (Development Initiatives 2012). This factor also affects those countries like Central Africa Republic whose international profile and related strategic priority is low. The threshold for response is correspondingly higher in such contexts.

In this second category of contexts, the data available come mainly from international agencies. In most crisis contexts, however, there is a mix of government-generated (e.g., National Disaster Management Authorities (NDMA)) and external agency-generated data and analysis. Increasingly the mainly UN-led Clusters or else government-led coordination bodies are attempting to bridge the gap between the two. Joint assessment processes are one feature of this, an attempt to forge consensus and buy-in as well as to streamline and harmonize data collection. This has potential strengths and weaknesses from the point of view of evidence-based responses. The main strengths come from comparability of data and "buy- in" for the process and its results. The main weaknesses relate partly to the often cumbersome and slow nature of these joint processes, and partly to the potential for "group think" to dominate the related analysis. In this regard, independent assessment and monitoring processes (e.g., by individual agencies) continue to be an essential part of the evidential picture, often acting as early warning or corrective to the wider system, whose processes may not be responsive to significant changes at either the micro or macro level.

4.2 Specific conclusions

Despite the diversity of contexts for decision making, it is possible to draw a number of conclusions from the study.

Decision making

1. There appears to be a high level of "path dependence" in most decision making processes in the sector. In other words, the range of options is limited by previously

decided strategic priorities, resource allocation, and other biases. In some cases, these are parameters set by host government authorities; in other cases, they are set more by donors and by implementing agencies. This significantly limits the extent to which decisions are open to influence by evidence, particularly where organizational incentives to generate and respond to new evidence are limited.

2. The extent to which decisions are "predetermined" varies according to the type of decision. In some of the cases reviewed, the dominant political narrative and relative strategic priority given to the country/crisis in question was the factor that had by far the most significant bearing on strategic decisions about crisis response (approach, level of funding, etc.). In some cases of protracted crisis like DRC and Ethiopia, the path dependence is more about programmatic inertia: programs "roll" from year to year without fundamental re-assessment of approach. Where programs are more responsive to context, this tends to be at the lower levels of decision making and at the more local level of programming.

3. Decision makers may be highly selective in their uptake and interpretation of evidence. Personal biases, rules of thumb, and mental models—as well as a variety of (dis) incentives—may prevent individuals and organizations from responding to a situation in the way that evidence appears to demand. It is common for experienced staff to base decisions mainly on past experiences, instinct, and assumptions—even in the face of contradicting evidence. In institutional terms, this in turn leads to building agency capacity around established interventions types, which continue to be the "preferred response" with each new crisis, irrespective of available evidence. As one key informant stated, "people become specialized in something and it becomes more and more difficult to have an open mind and look at other response options."[27]

4. The use of standard predefined response packages is now being challenged, particularly in the area of food security and livelihoods. In evidential terms, this should involve combining evidence about context with historic knowledge about "what works." Yet it remains the case that very few agencies conduct a formal or structured analysis of the various options and base decisions on the evidence that points to the most appropriate response. Even when assessment is viewed as a priority for program planning, agencies often disregard field-validated assessments as a precursor to intervention. Ultimately, the choice of response does not always involve an evidence-based, analytical process (Maxwell et al. 2012).

5. Current processes of decision making tend to be undocumented and untransparent. It is therefore hard to judge whether or how or information and evidence have been used to inform them. In particular, key assumptions are often unstated and therefore hard to test.

Generation of evidence

6. Relatively few documented needs assessments are available beyond the confines of the agencies that conduct them. There has been a rise in the number of joint (multi-agency, multi-sector) needs assessments, and increased focus on the use of the MIRA tool in rapid-onset crises. This is a significant advance, although there remains a lack of genuine multi-sectoral analysis with the result that responses remain largely "siloed." To date, there has been less progress on joint assessment in protracted crises.

7. Even where documented assessments exist, the link between assessment and decision making appears weak. Assessments are still largely front-loaded and used to justify proposals or appeals (Flash Appeal, CAP, etc.). It remains the case that most assessments are conducted in order to substantiate a case made for funding by a particular agency to do a particular thing. Inevitable biases result in a lack of credibility—both of the analysis and of proposed interventions based on that analysis. This appears to be a major distorting factor in

[27] Interview.

the system. It creates a potential incentive to exaggerate the trigger event and its impact in order to secure as much funding as possible for the whole duration of the emergency during the critical "window" at the outset of the crisis.

8. The evidential threshold for new, untried interventions is often higher than for established approaches. For example, in order to justify the use of cash-based interventions or local/regional procurement of food assistance, most donors require agencies to provide a market analysis that ensures such interventions are appropriate and do not cause harmful side effects. Some donors require that agencies also provide a more comprehensive risk assessment or management plan to prove this. Requiring evidence to support proposed interventions is good in theory; however, when certain interventions require considerably more evidence than others, it influences decision makers' choice of intervention. Several of those interviewed agreed that they often automatically choose the intervention that has fewer requirements in terms of necessary analysis. This cost of complying with the many requirements for innovative approaches persuades decision makers to choose responses based on the amount of work required to propose the intervention, rather than the one that is most appropriate for the crisis. Even when agencies desire to gather the necessary data and evidence in support of innovative approaches, restrictions on time and access prevent them from doing so.

9. Situational monitoring and surveillance remains a major weakness in most contexts. However, this study suggests that even if agencies are conducting monitoring, their ability to change the original type of intervention after it has begun is strictly limited. Often, originally proposed interventions are no longer appropriate when it comes time to implement, as so much time passes between implementation and the original assessment. Yet despite knowing that circumstances had changed, they were unable to adapt the program, often due to contractual commitments with donors and hosting governments. This study suggests that there may be little incentive to monitor programs other than as strictly necessary for accountability purposes. This goes to the heart of the problem about the generation and use of evidence: unless there is some clear reason and incentive to do so, responses are likely to remain only weakly grounded in evidence.

What would enable better informed decisions?

10. This study suggests that there are many factors involved in the process of generating more informed decisions. Clearly, there is still room for improvement regarding the processes of data collection and needs assessment, but this is not a silver bullet for achieving improved decision making. First, larger, systemic changes must occur whereby there are better incentives for generating and using evidence in decision making. Second, ongoing assessment and situational monitoring must be more widely adopted. However, in order for this to be effective in improving humanitarian response, the wider humanitarian system must allow flexibility for agencies to adapt programs to meet the changing needs throughout the duration of a crisis. Third, quality analysis and the use of evidence must be highly valued through increased investments in diagnostics. Fourth, the evidence base proving which humanitarian responses are most effective is extremely lacking. Investments must be made in the consolidation of evidence about what works in response to different kinds of needs in different contexts. Fifth, the way evidence is presented is often crucial to its uptake. Knowing how to present it, to whom, and in what form may be essential to informed decision making. ■

REFERENCES

Bradt, D. (2009). Evidence-based Decision-making in Humanitarian Assistance. Humanitarian Practice Network Paper 67. London, Overseas Development Institute.

Buchanan-Smith, M. (2000). Role of Early Warning Systems in Decision-Making Processes. London, Overseas Development Institute, Humanitarian Policy Group.

Buchanan-Smith, M. and S. Davies (1995). Famine Early Warning Systems and Response: The Missing Link. London, IT Publications.

Buchanan-Smith, M., S. Davies, et al. (1994). Food Security: Let Them Eat Information. Bulletin 25.2. Institute of Development Studies.

Collins, S. (2001). Changing the Way We Address Severe Malnutrition During Famine. Lancet. 358: 498–501.

Darcy, J. and C. A. Hofmann. (2003). According to Need? Needs Assessment and Decision-making in the Humanitarian Sector. Humanitarian Policy Group Report 15. London, Overseas Development Institute.

Darcy, J., A. Griekspoor et al. (2003). The Southern Africa Crisis: A Critical Review of Needs Assessment Practice and Its Influence on Resource Allocation. Background Paper for HPG Report 15. London, Overseas Development Institute, Humanitarian Policy Group.

Darcy, J., S. Anderson et al. (2007). A Review of the Links between Needs Assessment and Decision-Making in Response to Food Crises: Study Undertaken for the World Food Programme under the SENAC Project. London, Overseas Development Institute.

de Ville de Goyet, C., and L. Morinière (2006). The Role of Needs Assessment in the Tsunami Response. London, Tsunami Evaluation Coalition.

Development Initiatives. (2012). Global Humanitarian Assistance 2012. Bristol, Development Initiatives.

Disasters Emergency Committee (DEC). (2012). Response to Crisis in East Africa—Ethiopia Real-Time Evaluation Report. Oxford, Valid International.

Food and Agriculture Organization (FAO). (2008). The Integrated Phase and Humanitarian Phase Classification Technical Manual. Version 1.1. Nairobi, Kenya, Food and Agriculture Organization.

Garfield, R., C. Blake et al. (2011). Common Needs Assessments and Humanitarian Action. Humanitarian Network Paper Number 69. London, Overseas Development Institute.

Gilovich, T., D. Griffen, and D. Kahneman. (2002). *Heuristics and Biases: The Psychology of Intuitive Judgement*. Cambridge University Press.

Golden, M., M. Brennan et al. (2006). Measuring Mortality, Nutritional Status, and Food Security in Crisis Situations: SMART Methodology Version 1, SMART.

Haan, N., N. Majid et al. (2005). A Review of Emergency Food Security Assessment Practice in Ethiopia. London, Overseas Development Institute.

Hillier, D. and B. Dempsey. (2012). A Dangerous Delay: The Cost of Late Response to Early Warnings in the 2011 Drought in the Horn of Africa. UK, Save the Children and Oxfam.

IASC. (2012). Real-time Evaluation of the Humanitarian Response to the Horn of Africa Drought Crisis: Somalia 2011–12.

Inter-Agency Standing Committee (IASC). (2009). Initial Rapid Assessment (IRA): Guidance Notes.

International Federation of Red Cross and Red Crescent Societies (IFRC). (2011). World Disasters Report 2011: Focus on Hunger and Malnutrition.

Kahneman, D. and A. Tversky. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica* XLVII: 263–291.

Levine, S. and C. Chastre (2004). Missing the Point: An Analysis of Food Security Interventions in the Great Lakes. HPG Network Paper 47. London, Overseas Development Institute, Humanitarian Policy Group.

Levine, S., A. Crosskey et al. (2011). System Failure: Revisiting the Problems of Timely Response to Crises in the Horn of Africa. London, Overseas Development Institute.

Maxwell, D., J. Parker, and H. Stobaugh. (2012). What Drives Program Choice in Food Security Crises? Examining the "Response Analysis" Question. *World Development*, Special Edition on Impacts of Innovative Food Assistance Instruments, forthcoming.

Maxwell, D. and B. Watkins. (2003). Humanitarian Information Systems and Emergencies in the Greater Horn of Africa: Logical Components and Logical Linkages. *Disasters* 27(1): 72–90.

McClelland, D. (1967). *The Achieving Society*. The Free Press.

Mock, N. and R. Garfield. (2007). Health Tracking for Improved Humanitarian Performance. *Prehospital and Disaster Medicine* 22(5): 377–383.

Office for the Coordination of Humanitarian Affairs (OCHA). (2009). Assessment and Classification of Emergencies (ACE) Project: Mapping of Key Emergency Needs Assessment and Analysis Initiatives.

Overseas Development Institute (ODI). (2009). Humanitarian Diagnostics: The Use of Information and Analysis in Crisis Response Decisions. Paper prepared for FAO. London, HPG/ODI.

Oxfam. (2012). Crisis in a New World Order: Challenging the Humanitarian Project, Oxfam Briefing Paper 158.

Stern, E. et al. (2012). Broadening the Range of Designs and Methods for Evaluations: Report of a Study Commissioned by the Department for International Development. DFID Working Paper 38. London, DFID.

Teger, A. (1980). *Too Much Invested to Quit*. New York: Pergamon.

Thomson, A., P. Jenden et al. (1998). Information, Risk and Disaster Preparedness: Responses to the 1997 El Nino Event. Research Report DFID ESCOR No. AG1215. London, SOS Sahel.

UNICEF Internal Report. (2011). Analytical Review of Current Humanitarian Advocacy Practices in UNICEF—Synthesis Report.

United Nations High Commissioner for Refugees, World Food Programme et al. (2000). Food Security Assessments, Self-Reliance, Targeting and Phasing Out in Ongoing Refugee Situations. Rome, WFP.

Walker, P. (1989). Famine Warning Systems, Victims and Destitution. London, Earthscan Publications.

World Food Programme (WFP). (2009a). Comprehensive Food Security and Vulnerability Analysis. Rome, WFP.

World Food Programme (WFP). (2009b). Emergency Food Security Assessment Handbook. Rome, WFP.

World Health Organization (WHO). (2009). Health and Nutrition Tracking Service (HNTS) Consultancy Report and Background Document: Priority Indicators in Complex Emergencies.

World Health Organization, World Food Programme, United Nations System Standing Committee on Nutrition, and United Nations Children's Fund. (2007). Community-based Management of Sever Acute Malnutrition: A Joint Statement by the World Health Organization, the World Food Programme, the United Nations System Standing Committee on Nutrition, and the United Nations Children's Fund.

6.1 Research methods

A combination of both secondary and primary data was collected for this study, primarily in the form of qualitative research methods. The following steps were conducted in carrying out this study:

- **Literature review**. First, an in-depth review of secondary documentation was conducted on both published and gray literature regarding decision making, and the use of assessments, other evidence, and analysis in humanitarian practice. This provided in-depth background knowledge to identify specific topics to further research within the following data collection methods.
- **Case Study Interviews**. A series of semi-structured, open-ended interviews were conducted with nearly 60 experts, practitioners, and agency representatives within the humanitarian community. Most interviews were conducted in-person, with the exception of a few where it was only feasible to conduct the interview via phone or Skype. Individuals interviewed for this research worked for a variety of national and international NGOs, UN agencies, government bodies, and donor agencies. These and other key informants had experience in a multitude of emergency responses, with experience in rapid-onset natural disasters, protracted conflict, and drought, among others.
- **Case Study Documentation Review**. A documentation review of secondary data in each of the country case studies was conducted to supplement and verify information collected in interviews. These documents included agency and government strategies, funding policies and procedures, example needs assessments, decision making tools, and others.
- **Final Analysis**. The final collection and analysis of data for this study took place in November/December of 2012.

Researchers chose to approach the study by analyzing the research questions on three levels: (1) globally; (2) country level; (3) program level. At the global level, approximately half of humanitarian funding is provided by three key donors: US, UK, and EU. Also, many smaller donors tend to follow the larger donors' lead. Thus, understanding how these big donors function would provide insight into and possible ways of affecting how the entire system works. Therefore, these were the main donors analyzed in the study.

Country-level case studies were purposefully selected to capture a variety of regions, contexts, governing bodies, and types of disasters, including rapid-onset, protracted, and slow-onset crises. The first case study, Ethiopia, represents longstanding food insecurity, with slow-onset and recurrent crises of drought. The second, Democratic Republic of Congo (DRC), represents a situation of protracted instability, conflict (or "post-conflict" and "transition"), with an unstable and uninvolved governing body. The third case study, Philippines, represents a context with recurrent sudden-onset crises in the form of flooding, with a very stable and active government.

In choosing which operational agencies to interview, a cross section of major agencies, including UN agencies, large NGOs, small NGOs, International Red Cross bodies (ICRC or IFRC), and national NGOs were targeted. Researchers attempted to gather primary information from a variety of both US and European donors. Members from the Clusters System, as an influential coordination mechanism for individual and collective decision making, were also targeted for interviews. On a program level, it was chosen to focus on food security, nutrition, health, WASH, and protection as the primary life-saving sectors during an emergency.

6.2 Individuals interviewed

Individuals from the following agencies were interviewed and/or participated in round table discussions for this study:

List of Agencies from which individuals were interviewed

Ethiopia:
- CARE
- Catholic Relief Services
- Disaster Risk Management and Food Security Sector (DRMFSS)
- Emergency Nutrition Coordinator Unit (ENCU) within Disaster Risk Management and Food Security Sector (DRMFSS)
- Famine Early Warning Systems Network (FEWS NET)
- Food and Agriculture Organization (FAO)
- Médecins Sans Frontières (MSF) Belgium
- Médecins Sans Frontières (MSF) Spain
- Office for the Coordination of Humanitarian Affairs (OCHA)
- Office for the Coordination of Humanitarian Affairs (OCHA)–Humanitarian Relief Fund
- Relief Society of Tigray (REST)
- Save the Children
- UK Department for International Development (DFID)
- US Agency for International Development (USAID)
- World Food Programme (WFP)

DRC:
- European Community Humanitarian Office (ECHO)
- Food and Agriculture Organization (FAO)
- International Committee of the Red Cross (ICRC)
- Office for the Coordination of Humanitarian Affairs (OCHA)
- Oxfam US
- Swedish International Development Cooperation Agency (SIDA)
- UK Department for International Development (DFID)
- UN High Commissioner for Refugees (UNHCR)
- United Nations Children's Fund (UNICEF)

Philippines:
- Action Contra la Faim
- Philippines Department of Health
- International Federation of Red Cross and Red Crescent Societies (IFRC)
- Office for the Coordination of Humanitarian Affairs (OCHA)
- Office of Civil Defense (National Disaster Risk Reduction and Management Council)
- Oxfam UK
- Plant International
- United Nations Populations Fund (UNFPA)
- United Nations Children's Fund (UNICEF)
- World Food Programme (WFP)

Other Key Information (Global or HQ):
- UN High Commissioner for Refugees (UNHCR)
- International Federation of Red Cross and Red Crescent Societies (IFRC)
- International Council of Voluntary Agencies (ICVA)

Feinstein International Center

Feinstein International Center
Tufts University
114 Curtis Street
Somerville, MA 02144
USA
tel: +1 617.627.3423
fax: +1 617.627.3428
fic.tufts.edu