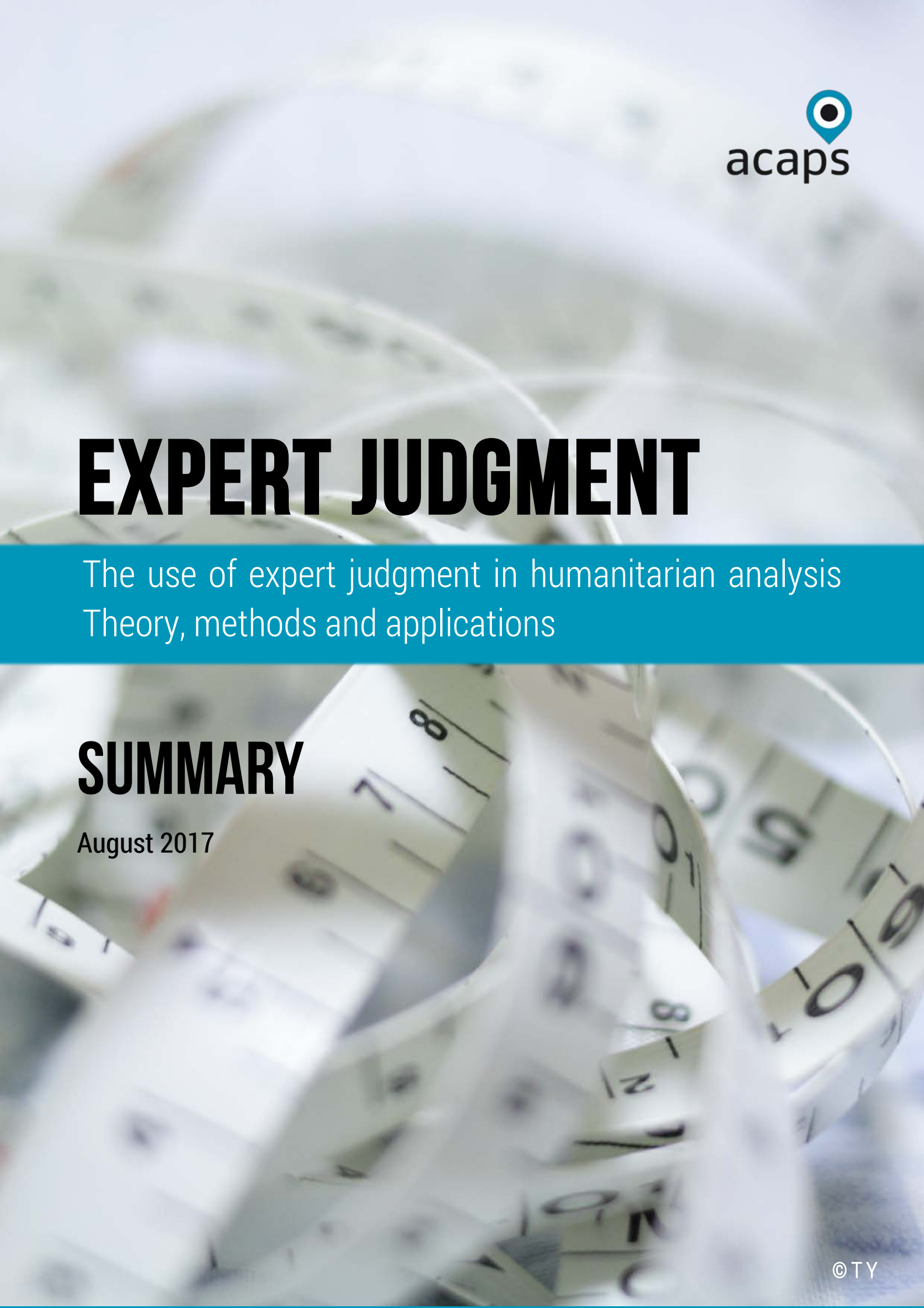


EXPERT JUDGMENT

The use of expert judgment in humanitarian analysis
Theory, methods and applications

SUMMARY

August 2017

AUTHORS

Aldo Benini

Patrice Chataigner

Nadia Noumri

Nic Parham

Jude Sweeney

Leonie Tax

CONTENT

This document is a summary of a detailed ACAPS note on the use and application of expert judgment in humanitarian analysis, which can be found on https://www.acaps.org/library

INTRODUCTION

What is expert judgment?

For the purposes of the study, the term 'humanitarian experts' covers the large and diverse group of people with specific and relevant subject-matter, technical, geographic, language, social and cultural intelligence. This includes sector and country experts, as well as humanitarian frontline workers and, in certain contexts, even the analysts and information managers. The perspectives of these experts strengthen humanitarian analysis, and if used appropriately, effectively inform humanitarian decision-making.

Experts are indispensable in modern organizations. They fill gaps in data and in the understanding of existing or missing data. They introduce, apply and teach techniques and methods, some of which staff of the experts' principals - and ultimately others - will continue to employ and disseminate. Technical experts reduce uncertainty by working out consensus opinions and probability ranges. Policy experts unravel the preferences and capacities of stakeholders; by doing so they dampen excessive certainty and may thereby increase uncertainty in strategic ways that decision makers and analysts find productive. When experts give their opinions in a context of decision-making, these become *expert judgments*.

The functional contributions of experts – data, interpretation, methods– and the professional roles that produce, process and consume judgment – expert, analyst, and decision maker - are generic and universal. They pattern the insertion and work of experts in the humanitarian sphere, too. Yet, there are some important differences vis-à-vis other institutional arenas. Typically, the environment of humanitarian action is more turbulent than those in which expert judgment methodologies have matured, such as nuclear power plant engineering. This turbulence blurs the distinctions among experts and other roles, including decision makers, analysts and key informants. It may also explain why there has been little systematic work about expert judgment methodologies for the humanitarian domain.

> When experts give their opinions in a context of decision-making, these become *expert judgments.*

Audience and objective

This note speaks primarily to *humanitarian analysts*. Analysts mediate between decision makers and experts. They facilitate the ongoing expert work, aggregate the contributions of several experts, and edit the aggregated product in the perspective and dialect to which the organization and its partners are habituated.

The goal of the note is to enhance *analyst* competence in dealing with experts and expertise. It offers insight also to *decision makers* who commission and consume expert work, and to *experts* for whom some aspects of expert judgment theory may be new. Further, it may help *stakeholders* and *academics* situate the particular dynamics of expert judgment in the humanitarian environment.

The summary and full technical brief do not intend to provide a comprehensive theory of humanitarian decision-making. Instead, tools and tasks that humanitarian analysts perform or oversee in the production of expert judgment are discussed. The major focus is on relationships among decision-makers, analysts and experts, on technical aspects of eliciting and aggregating judgments, and on demonstrating, through case studies, how reality affects even the best-laid plans to elicit and use expert judgment.

The summary focuses on the work of experts who answer specific questions at a given point in time or within a few weeks or months after recruitment at most in the context of humanitarian decision making.

THE PROCESS OF EXPERT JUDGMENT

The process of expert judgment typically is described in stages that include background and preparation; the recruitment of experts; collection ("elicitation") of experts' opinions; the combination of their contributions through quantitative aggregation and qualitative synthesis; communication of the findings to decision makers and stakeholders, and their actual use in decision-making.

Background, preparation, recruitment

Organizations feel the need to involve experts most keenly when they fail to meet objectives or when important processes are deficient. Other conditions that call for expertise are informational.

In particular, experts are more likely to be brought in when:

- normal data (e.g., from surveys) are sparse or flawed
- uncertainty is high
- experts are better, faster, or cheaper than other potential solutions
- new information requires frequent updates of assumptions and decisions
- additional validation is required
- available data are rich, yet some key parameters cannot be estimated

The situations that prompt decision-makers to hire experts are thus highly variable. Good preparation includes careful definitions at three levels:

1. the overall goal of the experts' work
2. the broad question areas they are supposed to cover
3. the specific questions they must answer in order to inform decisions meaningfully. The questions must be such that experts who can answer them can be effectively borrowed or hired from existing networks and markets.

Other key parameters to be considered prior to recruitment are:

- whether the expertise is meant to be reassuring or strategically disruptive
- whether the objective is to obtain the experts' answers or to learn their problem-solving processes
- the complexity of the areas and the questions, and the magnitude of the data required
- the number and types of experts needed, and whether they need to be recruited simultaneously, as opposed to sequential reviews and decisions to hire more or not
- the expected benefits weighed against the likely financial, opportunity and social costs of the expertise

These questions are resolved through an iterative process of refinement. This involves the sponsors (agencies funding the exercise, decision makers using the results, workers interested to acquire skills from experts), the concerned personnel (managers, analysts, data collectors) and to variable degrees, the experts themselves. It is not unusual for sponsors to ask experts to write their own terms of reference, or to begin by hiring an expert tasked with designing a multi-expert program. The degree to which prospective or confirmed experts are to work out question areas and specific questions is itself an important parameter to be determined in the preparations.

Those are all generic requirements. What is specific to humanitarian experts? In popular notions, expert authority rests on superior technical and subject matter knowledge, acquired in long years of training and experience. However, humanitarian experts function in a multi-disciplinary milieu. Most of them are valued because of particular personal mixtures of subject matter, technical, geographic, language, social and cultural intelligence. They are less clearly demarcated from other positions than experts in other domains; the boundaries with agency staff, key informants, citizen scientists and media workers are fluid and often permeated by issues of access and availability. Humanitarian experts are less often tied to any of the classic professions such as medicine, law and engineering; their knowledge is more syncretic – coherent if drawn from motley sources.

As a result, humanitarian experts are experts because in the eyes of their principals, the public and other experts, they know something that is worth knowing about humanitarian needs, risks and response.

> Humanitarian experts are experts because in the eyes of their principals, the public and other experts, they know something that is worth knowing about humanitarian needs, risks and response.

This self-referential definition is as inevitable as it is unsatisfactory. It is inevitable because the humanitarian work environment challenges professional boundaries. It is unsatisfactory because it may excuse vicious cycles of low expectations. In emergency settings, multi-sectoral gatherings and processes highly depend on agency representatives with a certain level of expertise. Notoriously, agencies often send participants whose major qualification is that they are the most expendable on that day.

The multi-disciplinary milieu and the thin boundaries with other social roles have consequences for the recruitment of experts. The decision maker decides whether the experts are primarily to reassure or to irritate, and for what purpose. The recruiters then look for an appropriate and manageable diversity of backgrounds, competencies and characters. Agency workers contribute much of the needed expertise, because they understand both the institutional agendas and relevant segments of the environment. Their use is more expedient than bringing in outsiders at additional expense and delay. The importance of local knowledge makes key informants indispensable experts.

Decision makers and analysts must then narrow down the remaining gaps with the help of external experts. They need to find experts with the right mixture of substantive and formal knowledge, at the right level of granularity. They need to decide if the primary aim is to harness certified knowledge or rather to exploit institutional affiliations. On top of that, they need to square all the desirables with budgets for consultants, agency calendars and the brittle consensus of colleagues and stakeholders.

Elicitation and recording

"Eliciting" is the technical term for the analyst's activity that causes the expert to form and express an opinion. It is a careful and carefully designed activity, down to fine detail; it adheres to *"specially designed methods of verbal and written communication".*

1. Elicitation situations can include one-on-one interviews between interviewer (analyst or data collector) and the expert, interactive expert groups, and Delphi techniques in which experts see the judgments of other experts indirectly.

2. Mode of communication, such as face-to-face, telephone and computer-aided (chiefly Web-based).

3. Elicitation techniques range from the time-consuming ethnographic approach (where the analyst rephrases the expert's responses continuously into new questions), to verbal reports (the expert thinks aloud while working towards a solution; the analyst records) to less onerous verbal probes (the analyst asks questions only after the expert reports his solution).

4. Response modes are the formats in which the experts are asked to encode their judgments, such as probabilities, ranks or ratings.

5. Aggregation is the combination of several pieces of information into one statement. This is part of the elicitation process only if it is done by the experts themselves. Aggregation is "behavioral" when it results from interaction among experts, as opposed to "mathematical" when it is the outcome of algorithms an individual expert employs.

In a classic of the expert judgment literature, Meyer and Booker (1991:100 sqq.), from which the above quote was taken, divide elicitation into five components. Every one of these should be carefully considered in the design of the elicitation process:

The combinations of the different components and refinements are almost limitless -- the elicitation chapter in Meyer and Booker runs over a hundred pages. This amount of guidance cannot and should not be absorbed until needed. Even then, one should proceed selectively and with firm resolution to adhere, and have data collectors and experts adhere, to the arrangements that are critical for the particular purpose.

Regardless of specifics, it is always helpful to evaluate elicitation options against two general insights. First, and unsurprisingly, experts are subject to the limitations of human information processing. The interaction between analyst and expert, and among experts, operates similarly to the interviewer-respondent dynamic in questionnaire surveys. Survey methodologists have broken it down to four constituent operations:

- Comprehension – respondent interprets the questions
- Retrieval – respondent recalls from memory the information needed to answer
- Judgment and estimation – respondent combines and summarizes the information they recall or substitute an estimate for a missing element
- Reporting – respondent formulates response and puts it in the required format (Groves, Fowler et al. 2004:202).

The elicitation format needs to respect the limitations of the expert at every stage, even when the work extends beyond a single encounter.

Second, the turbulence of the organizational environment limits the complexity viable in elicitation instruments. High expert turnover means more time spent on repeated briefings and loss of historic depth. Rapid changes in target populations entail higher estimation error in the updates that key informants provide. Regardless of whether the turbulence is driven from inside or outside the agency eliciting expert judgments, it limits what can be extracted and transacted. Turbulence exerts pressure for simplification – it places a premium on robustness and reliability at the expense of detail, precision, and sometimes validity.

On the analyst's side, proper arrangements to record the experts' judgments and, as much as desired, the operations that formed them must be considered in the very design of the elicitation. In addition to data and statistics, the observation bases of the individual experts need to be established and recorded. It is one thing to collect estimates about the proportion of destroyed buildings from four experts, and another to know that expert A rests his or her estimate on his visits to two villages since the earthquake, B has been to four, C to 15 and D to at least 20. This information is critical in the aggregation phase.

Aggregation and synthesis

After information has been collected from multiple experts and properly recorded by the analysts, the next step is to "aggregate" it. "Aggregation" has several meanings that need to be kept separate.

In quantitative contexts**,** an analyst reduces the estimates that the experts have produced of the same variable to one combined value. Ideally, the aggregation produces also a measure of confidence or concordance, such as the confidence interval around the estimate of a continuous variable or of a proportion. Aggregation in the quantitative context also refers to operations that combine counts and estimates made on low-level administrative or geographical units. They produce statistics of the next higher level(s), chiefly by summation and weighted averages. Importantly, the uncertainty measures are not additive, but need to be recalculated at every level.

In the qualitative domain, "aggregation" is a misnomer. Propositions can be counted from their occurrences in expert reports, but such frequencies do not constitute a summary. When analysts summarize qualitative information from multiple experts, they need to find abstractions on a higher level, perhaps borrowed from academic theories. The process is more aptly called "synthesis", the term used in this note.

We were, however, unable to find pertinent literature guiding the synthesis of qualitative expertise in the humanitarian world. Guidance can be taken from outside, from Sandelowski and Barroso's *Metasynthesis of Qualitative Findings* (2003). These authors lay out a solid, practical three-step process. In step one, the analyst collects, abstracts and orders the findings from all the contributions under review. In step two, the analyst reorganizes the findings, comparing them in multiple ways. In step three, the analyst extracts the experts' own syntheses, imports external concepts and fuses interpretations in their light. By concentrating on this one approach, a path is cut through the sprawling thicket of qualitative methods.

There are numerous methods and approaches to aggregating of quantitative data. When aggregating quantitative expert judgments, analysts frequently face a combination of the following four situations. For each situation a relevant method for aggregation is provided:

	Situation	**Possible method**
Scalars (real-valued unconstrained variables) when:	Experts state their uncertainty	Triangular probability distribution method
	The observation bases of the individual experts are known	Teta distribution method
Proportions and probabilities when:	Experts give only point estimates, but the experts are the same for all objects	Beroggi Wallace method
	The observation bases are not known	Bordley's formula

For more detail on the use, logic and caveats of these methods please refer to the full note.

Basic familiarity with Bayesian ways of thinking is essential to quantitative and qualitative judgment. Humanitarian information is frequently updated; Bayesian thinking is all about how beliefs change in the light of new evidence, such as that produced by experts.

Bayes' theorem connects the probabilities that members of a population have two attributes, A and B. Specifically, it formulates a rule about the probability that a member has attribute A if it is known that he/she has B. The theorem has much wider applications, including in expert judgment, when the meaning of A and B is generalized. Modern statistics is in the grip of a Bayesian revolution. Many humanitarian analysts will work with new methods from this wellspring at some point, providing added incentive to learn Bayesian basics early on.

> "Basic familiarity with Bayesian ways of thinking is essential to quantitative and qualitative judgment.

"Process tracing" provides an example of a Bayesian application, by updating the strength of a belief in the light of accumulating evidence. The method focuses on an observable event and traces the causal chain back in time in order to identify possible causes. It can be extremely useful for testing beliefs about causes and effects, on the cusp between the qualitative and the quantitative.

Throughout the process, continued characterizing of the uncertainty surrounding experts' estimates and interpretations is the basis of good elicitation design. Analysts should strive, as much as possible, to have experts provide such measures.

HUMANITARIAN APPLICATIONS

ACAPS' Analysis Spectrum

Ten case studies of actual humanitarian applications of expert judgment were selected. The cases are categorised according to the intended level of analysis, following the ACAPS Analysis Spectrum[1] - from exploratory analysis to descriptive, explanatory, interpretative, anticipatory, and eventually prescriptive analysis. All the cases are from ACAPS' own work, except the one on prescriptive expert judgment. The decision to develop case studies from ACAPS' experiences was primarily driven by the lack of other documented humanitarian examples – it does not suggest that ACAPS is the only actor who uses expert judgment as an informal or formal data collection method.

1. Exploratory analysis: Using experts to find information

ACAPS' "Refugees/Migrants in Europe" project ran from December 2015 to March 2016. It illustrates ways of harnessing expertise to the collection of dynamic information when other methods would be unacceptably slow, expensive or inflexible. ACAPS contacted key informants in five countries that were the most affected by the mixed migration flow in southeast Europe. It aimed at interviewing at least five individuals per affected country, and re-interviewing the same people every two to three weeks. The informants were country government, donor, Red Cross and non-governmental organization (NGO) staff, all of them closely involved in the migrant crisis. The contacts were one-on-one, and by phone, Skype and email.

The set-up proved a cost-effective, quick and useful way to identify emerging issues early on. However, ACAPS ultimately communicated with as many as 61 key informants in 44 organizations. Turnover of informants was high, either because they moved frequently, or because the motivation to be re-interviewed was weak. On the part of the ACAPS analysts, the semi-structured questionnaire proved too short and basic for the purpose. Also, the four analysts working on the project made decisions individually on how to synthesize conflicting information; a common structured approach to qualitative analysis was not in place. Up to a point, that weakness was compensated for by a strong point of the project, which was the speed of the analysis; on several occasions, ACAPS learned of and communicated about existing needs before the media noticed them.

[1] For more information on the Analysis Spectrum and the different levels of analysis see: https://www.acaps.org/sites/acaps/files/resources/files/acaps_analysis_spectrum_poster.pdf

2. Descriptive analysis: The population inside Syria, 2015

The Whole of Syria Needs Identification Framework (WoSA-NIF) was a collective effort to obtain sub-district level planning figures of the population, internally displaced persons (IDPs) and returned IDPs for the United Nations 2016 program year. ACAPS helped the United Nations Office for the Coordination of Humanitarian Affairs (OCHA) regional office in Amman, Jordan, with the design and analysis of population estimates that key informants based in the 270 sub-districts of Syria communicated, mostly by mobile phone. Ultimately, the Amman office obtained 1,323 usable estimates on 217 of the 270 sub-districts. The number of contributing key informants per sub-district varied from one to twelve.

The case is selected because it illustrates the judgment aggregation challenge and the need to make adjustments that were not foreseen in the initial design.

From the outset, it was obvious that the individual estimates would be uncertain, although nobody could anticipate the degree of uncertainty. Key informants were asked to supply three population figures: their most plausible value, and the minimum and maximum of the range in which the true value fell in their belief. This characterization of the uncertainty opened the way for a probabilistic analysis. In addition, an OCHA-led committee in Damascus supplemented the data with their own estimates, which they submitted in the same format, although with only one estimate per sub-district. A third set of estimates was obtained commercially, based on satellite imagery. It only offered point estimates, no uncertainty measures. Among them, the three sources covered all 270 sub-districts.

An ACAPS analyst performed a multi-method aggregation. Initial results revealed considerable upward bias in key informant estimates. A country level demographic model, based on consensus figures, provided a realistic figure of the national population as well as a global adjustment factor. Accordingly, the sub-district estimates were adjusted downward, with confidence intervals wrapped around them of the same relative width as in the unadjusted estimates. In a further step, a committee reviewed each of the 270 estimates and revised them, where necessary, on the basis of new and specific information. In 31 of the 270 cases, the revised figures were outside the WoSA confidence intervals.

The exercise supplied a grid of planning figures at the sub-district level, with reasonable measures of uncertainty. On the problematic side, the same adjustment factor was applied to all initial sub-district estimates. It is unrealistic to assume that the extent of bias was the same

everywhere. Two lessons to consider for the future concern design and analysis. It was helpful that during a first mission the ACAPS analyst helped work out the elicitation format and at the same time pre-tested a probabilistic aggregation algorithm. During the second mission two months later, the analysis nevertheless required frequent improvisation, in order to deal with peculiarities of the data. The use of multiple key informants per sub-district was helpful, too. Individually, the informants were too sure of their estimates; collectively, they were cautious, resulting in conservative (i.e., wider) confidence intervals.

The overall experience was such that three-valued elicitations (minimum, most plausible value, maximum) from technical experts and key informants are recommended. The three values define a suitable probability distribution. Distributions from multiple experts can be combined by a method that analysts can follow in Excel spreadsheets.

3. Explanatory analysis: Problem tree Ebola

Ebola broke out in Guinea in December 2013. It was internationally recognized as an epidemic ravaging also Liberia and Sierra Leone in the spring of 2014 and was declared closed in May 2016. At the beginning of the crisis, the international community perceived the Ebola outbreak in West Africa as a purely public health emergency. The response was oriented towards the containment of the epidemic and treatment of the sick.

However, a major lesson learned during this epidemic was the need to broaden the scope of the humanitarian response during a large-scale Ebola outbreak. The disruption of public and private services created an 'emergency within the emergency'. Entirely focused on disease control, humanitarian actors failed to activate their surge capacity and to set up emergency funding and coordination structures. It took time for the humanitarian community to recognize the complexity of the crisis and respond to secondary impacts. In hindsight, this appears common-sensical, if not always obvious. At the time, however, it took concerted efforts to change perceptions, priorities and ultimately resource mobilization. In this collective endeavour, ACAPS' analytical approach proved helpful.

Just as it was critical for the epidemiologists to isolate the Ebola strains, it was important to unravel the entire variety of secondary impacts and the derived unmet needs that together with the epidemic formed the humanitarian emergency. ACAPS analysts used a matrix structure in which humanitarian sectors were assigned to columns. Five sectors comprising health, WASH (Water, Sanitation and Hygiene), food security, livelihoods and protection, were profiled. The

matrix was a variant of a "problem tree"; such visual devices are particularly useful at the initial stage of analysis, to help the experts penetrate a broad issue by directing attention to smaller units, one causal "root" and "branch" (effect) at a time.

ACAPS' Ebola problem tree served as an advocacy document to approach donors and finance a specific project dedicated to the analysis of the Ebola crisis. It was also used as the basis for subsequent documents. Thanks to the preliminary identification of potential problems linked to the large-scale Ebola outbreak, ACAPS staff were able to accelerate the production of reports on health, food security and protection and to undertake analysis and produce reports on secondary consequences.

4. Interpretative analysis: Severity ratings in complex emergencies

In the first half of 2013, groups of humanitarian responders conducted three needs assessments in northern Syria. ACAPS personnel were involved in all three. They advised and supported the coordinating bodies in the measurement of severity and in the organization, processing and analysis of the sub-district-based field assessment data.

This continuity enabled rapid collective learning and eventually the transfer of good practices to subsequent assessments. The initial five-level severity scale, applied in five sectors (public health, food security, nutrition, shelter, water and sanitation), was found to discriminate poorly. It was replaced with a seven-level scale formulated so key informants could distinguish levels of unmet need even in regions that the conflict had already devastated.

The surveys also got increasingly better at debriefing the enumerators that returned from their assigned sub-districts with syntheses of the information that they had collected from hosts of key informants. Debriefers would discuss every sectoral severity rating with the enumerators, who had to actively justify them.

In the process, the important insight emerged that assessing data quality and drawing analytic conclusions call for different tools and competencies. The criteria for evaluating individual pieces of evidence are different from those suited for entire bodies of evidence. Furthermore, the inferential processes that result in assessment conclusions follow their own logic.

5. Interpretative analysis: Clarifying priorities

In the early stages of needs assessments and response planning, priorities may change frequently. Updated priorities are worked out by comparing the latest available information on sectors, geographical areas and social groups. In this regard, the updating mechanism used by the Nepal Assessment Unit is of special interest. The Unit was activated by the United Nations Disaster Assessment and Coordination (UNDAC) in Kathmandu days after the first of two earthquakes that struck Nepal in spring (April and May) 2015.

At its peak, the Assessment Unit employed 17 people. Its analysts were specialized by information source (social media, local newspapers, international media, government, Clusters, etc.). They would report following a firm daily rhythm - all new information reviewed by 9 am, consolidated briefs by 11 am, plenary meetings to update the ranking of sectors and affected districts at 1 pm.

The time pressure was enormous. The afternoon meetings produced priorities using a voting method, in order to avoid lengthy discussions that might or might not result in a group consensus. More time was devoted to exchanging information that had arrived from different sources. During the first two weeks, priorities shifted noticeably from districts with a wealth of information to those with wide initial information gaps that were progressively closed as more reports reached Kathmandu.

The specialization according to source type had advantages as well as drawbacks. The analysts quickly learned to focus on the most productive and reliable sources within their respective briefs. Familiarity helped spot duplicate or amplified stories. But the arrangement fostered segmented knowledge as well as blindness to information gaps. It took organized efforts, in the discipline of the tightly led meetings, to build broader views.

Besides demonstrating the advantage of having a collaborative system for multiple analysts, the case relates expert judgment production to decision-making under time pressure.

6. Interpretative analysis: Decision support in an information-poor environment

Haiti, sadly remembered for the earthquake of 2010, suffered another disaster in October 2016.

Hurricane Matthew harmed some 2.1 million people. Two ACAPS analysts were part of the UNDAC assessment team to analyse humanitarian needs. In the process, the two made a running inventory of all the assessment information arriving at Port-au-Prince. They evaluated every elemental combination of affected commune (61 in total), topic (humanitarian access and sectors; 9 in total) and information dimension (3) against a six-level scale. An aggregation formula supplied an overall information gap score for every affected commune.

The results, visualized in periodic maps, dramatically drove home the slow progress of assessments. The maps were widely noted in government and humanitarian agencies and caused the assessment activity to ramp up. The work was appreciated because the evaluation of gaps was based on a methodology that ACAPS had already developed elsewhere, and which its analysts adapted to the local conditions and applied with meticulous consistency. The case illustrates how a complex situation can be better understood by appropriately combining a large number of "small" judgments, each generated on the same criteria.

7. Anticipatory analysis: Qualitative risk analysis

On a monthly basis, ACAPS produces detailed analyses of particularly relevant or dynamic risks. The project augments the scope of forward-looking analysis and contributes to the analytical quality of humanitarian early warning and preparedness. It focuses on contexts that are deteriorating beyond their current trend and on the specific risks that are expected to materialize within a one to six-month timeframe.

The case study exemplifies the situation in which analysts, by necessity, have come to fill the role of expert, the potential but also the danger that this poses for consistent judgment.

The project relies primarily on secondary data, rather than on the opinions of external experts. A team of ten analysts works part-time on this project. They review the secondary data, and in this case are responsible for providing the expert judgment. Each analyst looks after six to ten countries and is supposed to show deep contextual understanding as well as monitor changes daily. The team follows a seven-step risk procedure that moves from initial problem analysis to risk prioritization, and subsequent continued monitoring. The project does not compute quantitative probabilities, but uses a five-level scale that combines levels of magnitude (persons requiring humanitarian assistance) and intensity (severity of the impact).

Some of the challenges transcend risk analysis. Foremost is the consistency by which the ten analysts identify risks and estimate likelihood and impacts. Information does not always arrive in time to recognize deteriorating situations early. Conversely, information to determine that a risk has materialized or has dissolved due to improved circumstances is often not conclusive. The weekly production rhythm is so short that analysts tend to focus on the present, at the expense of the broader picture. The project is striving to develop a more sophisticated methodology, with greater reliance on indicators that can be followed over longer periods of time.

8. Anticipatory analysis: Scenario building

ACAPS has several times brought together experts in scenario-building exercises. It relies on expert judgment in order to discern possible clusters of developments that could occur in crisis areas. Past exercises addressed the armed conflict in Nigeria, food security in Indonesia, the European refugee crisis and developments in Syria.

The exercises follow a "chain of plausibility" approach. The experts identify trigger events liable to bring about sets of conditions that together define the scenarios. In turn, the conditions jointly produce humanitarian impacts.

Typically, thirty or more experts would meet face to face. Not all meetings were equally productive. ACAPS found that one and the same meeting format could not support the progression from general topics to highly specific questions, but needed to change from step to step. For developing broad generalities, bilateral meetings with experts are effective. The topics can then be refined in a workshop that ideally attracts between ten and fifteen experts. To work out likelihoods of scenarios and impact, a separate phase in which the analyst works with no more than five experts has proven conducive. Care must be taken to formulate questions that are sufficiently specific to engage the experts' technical knowledge, to disentangle political positions from technical perspectives, and to collect the views of all participants.

9. Anticipatory analysis: Visualizing impact and probability

Some of the classic expert-judgment methodologists were opposed to assigning probabilities to scenarios (see scenario building, above). ACAPS lets workshop participants discuss and, if

possible, agree on probabilities. The scales are verbal, with five levels ranging from highly unlikely to highly likely. In reports, the scales are rendered as discrete color scales. Although they are not numerically defined, the equal size of the blocks suggests equally wide probability intervals. Similarly, the impacts expected of a given scenario are visualized in a five-level color scale.

Probabilities are assigned to scenarios because they are needed for contingency planning. If the risk that a scenario implies (probability X impact) is estimated, potentially costly preparations can be avoided or minimized. Scenarios with higher risk deserve higher priority for preparation.

The case study demonstrates complications with probabilities, telling the experience of two workshops held in Nigeria in August 2016. The workshops aimed to anticipate possible developments in the northeast conflict zone between October 2016 and June 2017. The participants were drawn from government agencies and international NGOs in two places: Maiduguri, the coordination center of the affected areas, and Abuja, the national capital.

The workshop dynamics were such that the ACAPS facilitators could not get the participants to work out scenario probabilities. Four scenarios were defined, by degrees of change in security situations. The discussion of trigger events and their weights and probabilities took up most of the time. Weighting and aggregating trigger events essentially remained unresolved. The facilitators later decided the scenario probabilities on the strength of bilateral discussions with experts, after the workshops.

The selection and weighting of events, and the aggregation of their probabilities in the scenario that they are expected to trigger remain challenging. Also, workshop participants tend not to be overly concerned with the definition of exhaustive and mutually exclusive scenarios. They sometimes create scenarios that are overlapping, defined by differing sets of variables. The analysts found that, contrary to our original scale that implies equal probability ranges, scales with ranges of unequal width, with a wide range in the uninformative middle, and progressively narrower ranges towards the extremes, provide better consistency and flexibility, but need more explaining.

Ultimately, the art of guiding probability discussions in scenario workshops is more effective when the facilitators emphasize plausibility over precision. Selection of participants with

adequate skills, continuity between workshops and disciplined facilitation make it likely that workshops produce plausible scenarios and meaningful probability ranges.

10. Prescriptive analysis: Experts vs. decision-makers

ACAPS does not prescribe, except through methodological advice and exemplary analysis. For a prescriptive-level case study, it was necessary to turn to outside experience. The case, a strategy change during the 2010-11 famine in Somalia, was selected because of an apparent paradox. It illustrates the possible gap between expert judgment and the decision making that it was expected to inform.

By that time, Somalia was an information-rich environment. The early-warning system functioned; the risk of impending famine was correctly assessed and repeatedly communicated. Expert judgment was good – pertinent, timely and reliable.

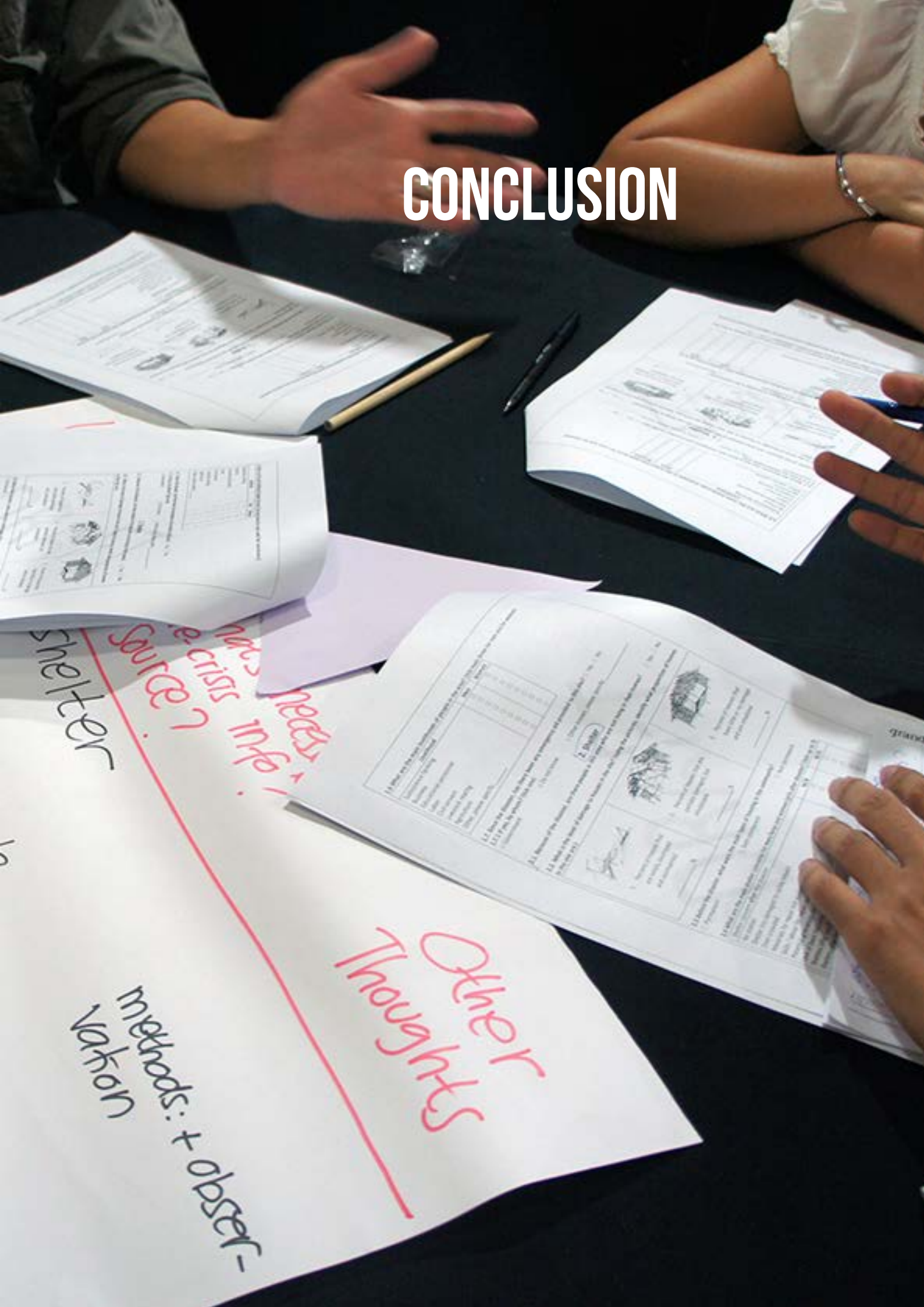
Nonetheless, the humanitarian community failed to mount a timely response.

Other factors stood in its way: uncertainty compounded by lack of access, under-subscribed funding requests, varying beliefs in the feasibility of strategies, and coordination burdens. Only when the United Nations declared a famine – by this time it had fully developed – was the response rapidly and radically transformed.

One comes away with the impression that *technical* expertise was not in short supply, but that *policy* expertise was not effective. More importantly, the belated response to the famine teaches that expert judgment does not replace leadership, nor can the best experts ensure coordinated decision-making.

The tragedy led to a lot of international soul-searching. Notably, and beyond Somalia, more attention has since been paid to response analysis and response planning, an area that itself is fertile ground for expert judgment.

CONCLUSION



"Expert opinion", in the words of a classic author in this field, *"is cheap, plentiful, and virtually inexhaustible"* (Cooke 1991:3). That is an exaggeration. At a minimum, we can say there is an established science of expert judgment. Some of its tenets transfer to the humanitarian domain almost in their entirety. The imperative to clearly define expectations at three levels – overall goal, broad question areas, specific questions – is a good example.

Other aspects are significantly modified as we move from generic expert judgment methodology to applications in the humanitarian sphere. The humanitarian environment is turbulent. The turbulence has consequences for the functioning of expert judgment. Notably:

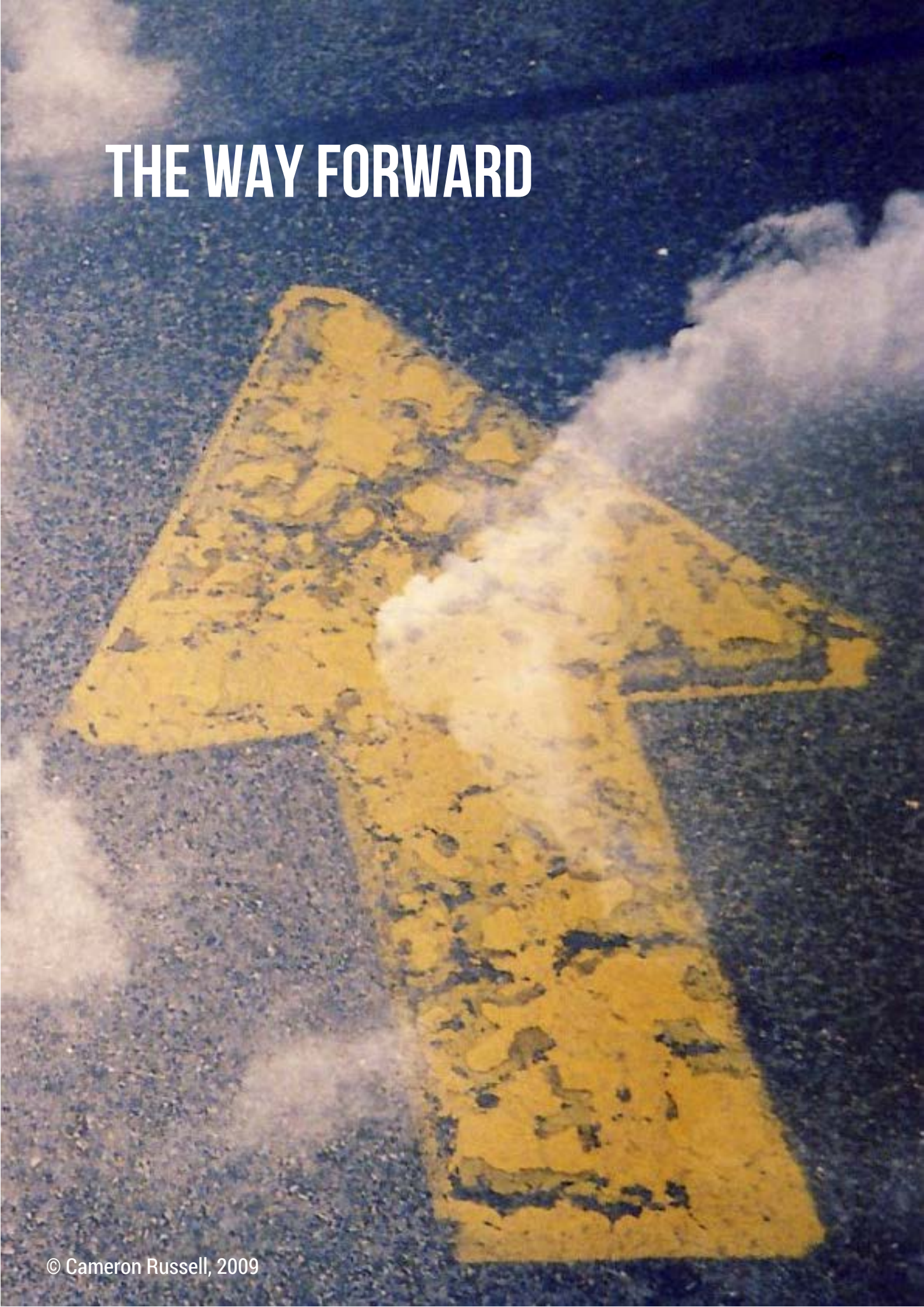
- The roles of decision maker, analyst and expert are less differentiated than in traditional expert settings.
- Local knowledge is indispensable, making the key informant as important as the technician.
- The bases of decisions grow out of sequential contributions, rather than from the consensus of experts convened simultaneously.
- The collection and aggregation of judgments have to overcome language, access, conceptual and skill barriers.
- The high turnover of personnel stunts methodological traditions; expert judgment is not yet widely recognized as an established methodology.

That is what we have seen so far. Change is pervasive, also in expert judgment. It is a safe bet that the development of Web-based applications and the Big Data revolution will unsettle the ways experts are recruited and work in the humanitarian world. Pressures for "evidence-based policies" and "value for money" will be passed on to decision-makers, and from them to analysts and experts.

It is harder to predict specifically how this will happen. Technologies that generate and process massive data of quantitative and categorical nature may be adopted with relative ease. We may also see innovations in dealing with qualitative expert judgment, such as in Web- or Intranet-based "argumentative Delphi", a technique that generates, evaluates and combines arguments and counter-arguments from a potentially large group of participants. Technique and infrastructure will connect novel coalitions of experts, key informants and even persons directly at risk in the evaluation of needs, environments and response options.

Humanitarian experts' greatest contribution, however, does not depend on technological savvy. It is intellectual and emotional, helping decision makers to confront "What's the story here?" and "Now what shall we do?" questions. It is sense-making. The kind of sense-making that, as another classic put it, "*involves turning circumstances into a situation that is comprehended explicitly in words and that serves as a springboard into action"* (Weick, Sutcliffe et al. 2005). If this study encourages readers to better prepare for this task, it will make sense, indeed.

THE WAY FORWARD



The use of structured expert judgment within humanitarian decision making is far from common practice. This note encourages its wider application, through case studies that demonstrate its scope, as well as a prescriptive chapter that provides solutions to specific problems. Above the flurry of historic and technical detail, our final paragraphs of this summary note some concerns that speak to the development of the wider business of humanitarian analysis, and the place of expert judgment within it. They take aim at longer-term attitudes, rather than immediate practices:

Shift the current focus on "data quality" to "conclusion quality"

Currently, the humanitarian community spends a lot of energy on improving data quality, and the debate continues on how best to achieve this. The focus on more and better data is incomplete, and sometimes misplaced. Perfectly valid conclusions can be derived from poor or insufficient data, and wrong ones from abundant and reliable data. If methods for assessing data quality have advanced, humanitarians have made scant progress in securing the quality of inference and interpretation. The wider adoption of analytical standards and expert judgment – both as a process and as an information source – can help to redress this imbalance.

Advocate for understanding and use

For expert judgment to turn into a mainstream method, those in a position to influence how data are generated must advocate for its use when appropriate. This includes creating an understanding of its strengths and limitations among those who commission and use humanitarian analysis. Guidance is specifically required to accompany critical humanitarian (funding) processes such as the Humanitarian Needs Overview and Humanitarian Response Plans, as well as population estimates. By the time the term "EEJ" (Eliciting Expert Judgment) has become as common as "FGD" (Focus Group Discussions), analysts' job descriptions should include familiarity with the method and techniques as an essential requirement.

Invest in wider applicability

Promotion of expert judgment should go hand in hand with an expansion of the current tools and guidance. At the time of this research, there was no common foundation of expert judgment concepts and tools in the humanitarian community. Standardized tools and guidance appropriate to this sphere are needed, specifically regarding:

- how to determine and report on levels of uncertainty, and when
- which structured analytical techniques apply to humanitarian settings, e.g. collaborative and joint analysis processes
- aggregation of expert judgment, e.g. for generating population estimates
- expert recruitment techniques and criteria, including for joint analysis workshops.

Other disciplines, particularly the decision sciences, should be queried for inspiration and guidance.

Use structured design and implementation

The note provides a set of recommendations on the actual production of expert judgment. While these skills and techniques belong chiefly to analysts, they thrive in a milieu in which leaders show an appreciation for analytic discipline. Expert judgment needs a system that is structured, planned and capable of incorporating other relevant information as required. The highly dynamic nature of a humanitarian crisis requires a stable process, with committed leaders, trained staff and rigorous methodology.

Capture and reflect uncertainty

A clear understanding of uncertainty is required for a successful elicitation process, and everyone involved plays a role in creating this comprehension. Experts are to be open about their limitations, information gaps and biases. Analysts must ensure that the uncertainty of estimates and inferences are recorded and communicated. Dissent among experts should be welcomed as a source of new insight and as corrective to false certainty; analysts should record opposing positions and their rationales, instead of papering over them with rash consensus findings. An acceptance of uncertainty, and an understanding of how to use this within decision making, is essential for all consumers of humanitarian analysis.

Recognize constraints and develop relevant approaches

The humanitarian environment makes it near-impossible to “calibrate” (assess on comparable track records or formal tests) individual experts and thus weight their opinions arithmetically. Staff turnover in emergencies is so high that it thwarts even the milder ambition of using the same experts across multiple occasions or different crises. To improve access to experts and hence the analysis of their contributions, new collaborative platforms are needed. To counteract loss of institutional memory and of coherence, arrangements are needed for analysts to extensively debrief experts and key informants. Severe limits exist on working with multiple experts simultaneously (although media like Skype have relaxed them somewhat);

strategies and methods for sequential recruitment and elicitation as well as for cumulative analysis need to be further developed. The Bayesian philosophy of updating one's beliefs in the light of new evidence should gradually filter into humanitarian analysis from disciplines where it is strongly established.

Documentation of practices

The product of the experts' work – often but not always a report – should describe the essential processes followed, the analytical techniques applied, the information sources used and the levels of uncertainty determined. Without these pieces of information, the product is diminished or even worthless to anyone outside of the persons directly involved. Documentation on expert work should at minimum state the decision-context in which the expertise was sought, how the experts were recruited, their terms of reference, and how their judgments were elicited, subsequently aggregated and finally translated into options and recommendations for the decision makers and stakeholders. Transparency on humanitarian expert judgment practices serves an additional purpose. In a setting with limited recorded practice, openness on tools and methods is essential to inspire and strengthen a budding movement that may, with good nurturing, grow into a normal practice – the practice of expert judgment.

REFERENCES

Betts, A. and L. Bloom (2014). Humanitarian innovation: The state of the art [OCHA Policy and Studies Series No. 9, November 2014], United Nations Office for the Coordination of Humanitarian Affairs (OCHA) New York, NY.

Cooke, R. M. (1991). Experts in Uncertainty. Opinion and Subjective Probability in Science. New York, Oxford, Oxford University Press: ca. 319 pp.

Groves, R. M., F. J. Fowler, M. P. Couper, J. M. Lepkowski, E. Singer and R. Tourangeau (2004). Survey methodology. Hoboken, NJ, J. Wiley.

Meyer, M. A. and J. M. Booker (1991). Eliciting and analyzing expert judgment: a practical guide. San Diego, Academic Press Limited.

Sandelowski, M. and J. Barroso (2003). "Toward a metasynthesis of qualitative findings on motherhood in HIV-positive women." Research in nursing & health 26(2): 153-170.

Weick, K. E., K. M. Sutcliffe and D. Obstfeld (2005). "Organizing and the Process of Sensemaking." Organization Science 16(4): 409-421.