Downloaded by [University of Groningen] at 05:15 06 May 2015

Humanitarian Crises, Intervention and Security

Downloaded by [University of Groningen] at 05:15 06 May 2015

This book presents a new framework of analysis to assess natural and man-made disasters and humanitarian crises, and the feasibility of interventions in these complex emergencies.

The past half-century has witnessed a dramatic increase in such crises – such as in Haiti, Iraq and Sudan – and this volume aims to pioneer a theory-based, interdisciplinary framework that can assist students and practitioners in the field to acquire the skills and expertise necessary for evidence-based decision-making and programming in humanitarian action. It has four major objectives:

- to provide a tool for diagnosing and understanding complex emergencies, and build on the concepts of state security and human security to provide a 'Snap-Shot Analysis' of the status quo;
- to provide a tool for analyzing the causes of crises as well as the related stakeholder field;
- to provide a frame to structure and analyze the information required to evaluate, monitor and/or design interventions for different actors on a project and/or program level;
- to combine concepts used in the humanitarian field with underlying theory in a practically relevant way.

The book will be of much interest to students of humanitarian intervention, human security, peacebuilding, development studies, peace studies and IR in general.

Liesbet Heyse is assistant professor in organization sociology at the Department of Sociology/ICS, University of Groningen, the Netherlands. She is author/editor of two books.

Andrej Zwitter is professor of International Relations, University of Groningen, the Netherlands. He is author/editor of three books.

Rafael Wittek is professor of sociology, Department of Sociology, University of Groningen, the Netherlands. He is author/editor of two books.

Joost Herman is professor of globalization studies and humanitarian action, University of Groningen, the Netherlands. He is author/editor of three books.

Routledge Studies in Intervention and Statebuilding
Series Editor: David Chandler

Downloaded by [University of Groningen] at 05:15 06 May 2015

Statebuilding and Intervention
Policies, practices and paradigms
Edited by David Chandler

Reintegration of Armed Groups After Conflict
Politics, violence and transition
Edited by Mats Berdal and David H. Ucko

Security, Development, and the Fragile State
Bridging the gap between theory and policy
David Carment, Stewart Prest, and Yiagadeesen Samy

Kosovo, Intervention and Statebuilding
The international community and the transition to independence
Edited by Aidan Hehir

Critical Perspectives on the Responsibility to Protect
Interrogating theory and practice
Edited by Philip Cunliffe

Statebuilding and Police Reform
The freedom of security
Barry J. Ryan

Violence in Post-Conflict Societies
Remarginalisation, remobilisers and relationships
Anders Themnér

Statebuilding in Afghanistan
Multinational contributions to reconstruction
Edited by Nik Hynek and Péter Marton

The International Community and Statebuilding
Getting its act together?
Edited by Patrice C. McMahon and Jon Western

Statebuilding and State-Formation
The political sociology of intervention
Edited by Berit Bliesemann de Guevara

Political Economy of Statebuilding
Power after peace
Edited by Mats Berdal and Dominik Zaum

New Agendas in Statebuilding
Hybridity, contingency and history
Edited by Robert Egnell and Peter Haldén

Mediation and Liberal Peacebuilding
Peace from the ashes of war?
Edited by Mikael Eriksson and Roland Kostić

Semantics of Statebuilding
Language, meanings and sovereignty
Edited by Nicolas Lemay-Hébert, Nicholas Onuf, Vojin Rakić and Petar Bojanić

Humanitarian Crises, Intervention and Security
A framework for evidence-based programming
Edited by Liesbet Heyse, Andrej Zwitter, Rafael Wittek and Joost Herman

Downloaded by [University of Groningen] at 05:15 06 May 2015

4 From theory to analysis

H-AID methodology

Rafael Wittek and Andrej Zwitter

Downloaded by [University of Groningen] at 05:15 06 May 2015

This chapter discusses some practical issues related to the collection of data in humanitarian contexts, and the assessment of its reliability and validity. Since space limitations prohibit an extensive treatment of these methodological issues in this book, this chapter can only provide a fairly brief and superficial sketch of some of the most important issues. The interested reader may consult the relevant literature for more detailed introductions. There are many excellent textbooks and guides on these topics, e.g. Bernard (2011).

The next section sketches the general purpose of our evidence-based H-AID framework, i.e. to move from an abstract and general understanding of the omnibus context to more domain or security specific theoretical constructs that can be subjected to empirical proof. The second section of the chapter describes the steps that are necessary to translate the theoretical context dimensions into valid indicators. The third section addresses the question how to identify trustworthy sources for these indicators. Finally, the fourth section discusses the question how to arrive at accurate estimates from trustworthy sources.

From omnibus context to theoretical constructs

In order to deal with information more effectively, it is good to get a general idea about context analysis. The most general form of context analysis is the omnibus context – in intelligence analysis this technique is also called star-bursting (Heuer and Pherson 2010: 102) referring to the following questions: Who? What? When? Where? (To Where? From Where?) How? Why? This technique requires sorting available information in accordance with generic criteria that are present in all contexts. Further questions more specific to humanitarian crises would be: What do we know about the causes of the crisis? What do we know about the consequences of the crisis? What do we know about relevant actors?

In addition, two other important questions to ask are: What are the known unknowns? What are the unknown unknowns? These two questions deal very specifically with present and lacking information. One can distinguish between 'known unknowns' (I know that I need to know but I have no information) and

'unknown unknowns' (I don't know that I need to know and whether I do or don't have that information). The known unknowns are quite easily responded to. If one knows that one is lacking information about malnutrition, one knows what data to collect. The unknown unknowns are a little more troublesome. Sometimes crucial information that could decide over the success or failure of a project is missing simply because of a general unawareness that a certain issue is relevant in the given context. This usually means that one knows too little about a given context in order to reasonably predict the effect of an intervention. For example, one might be well aware about all nutritional facts to implement an emergency feeding project, but one might not know enough about the cultural context that might prohibit certain feeding practices or not enough about the security situation in the region to conduct the project safely. The cultural element of feeding practices will generally be a known factor to nutrition experts; that is, it will more likely be a known unknown. However, the relevance of the security element might, due to the focus of expertise, completely escape the expert (unknown unknown).

In order to reduce these unknown unknowns, one should generally conduct a thorough analysis of the omnibus context to pick up hints of potentially relevant unknown factors. The tools presented in this book may help to realize this objective.

From theoretical constructs to valid indicators

Political, economic, social, food, health, environmental security … the starting point for a context analysis is not only *abstract* theoretical constructs, they also represent *compound* latent dimensions, i.e. they usually contain more than one facet or sub-dimension. This means that in order to make verifiable statements about the level of security threat in a domain, four operations are necessary.

Decomposition

First, one has to decide which sub-dimensions are best suited to map the general construct, i.e. one has to decompose the overarching construct into more specific categories (decomposition). For example, two possible sub-dimensions of the theoretical construct 'food security' could be to what degree (1) everyone in the population has sufficient access to food, and (2) the quality of the food is sufficient. Ideally, the number of sub-dimensions covers those sub-dimensions that are necessary while at the same time being sufficient to adequately encompass the overarching construct. Determining when the sufficiency criterion is met, remains a subjective exercise for most security domains that do not dispose of inventories of pre-defined diagnostics that have been agreed upon by a profession (e.g. like is the case for medical doctors). This judgment will also be affected by effectiveness considerations. Effective assessment requires that one should collect neither too much, nor too little information. Too much information will be hard to quickly assess and sift

Downloaded by [University of Groningen] at 05:15 06 May 2015

through. Too little information will be missing the relevant data and make the conclusions susceptible to unavoidable biases in the base data. In other words, one needs enough data to check on uncertain sources. The sources as well as the validity of the data (i.e. how well the data describes what the analyst wants to know) also affect how much (additional) data one needs. The lower the reliability and validity scores regarding particular resources, the less confidence the analyst can have in the assessment.

Operationalization

Second, one needs to find out what would be valid indicators to measure each sub-dimension (validity), and how many indicators would be needed to sufficiently cover all relevant aspects of this dimension (reliability)?

Information and evidence can be called *valid* if it is relevant and represents what one wants to know. Some indicators may simply not be valid measures for a construct. For example, if one wants to know how many people in a village are in need of what kind of nutrition, knowing the average nutritional levels of a region does not tell much. In other cases, a set of indicators may be valid, but some of them might be better and others worse in describing a certain context dimension. For example, both the average caloric intake of a population and the vicinity of food distribution points may be indicators for the sub-dimension 'access to food'. But the vicinity of food distribution points does not say much about the availability of food or the target population's ability to acquire it. Consequently, this indicator has validity problems, particularly when compared to the more direct measure 'average caloric intake'. However, this does not mean that vicinity to food distribution has to be excluded as an indicator: depending on the specific crisis situation, this indirect indicator may therefore still cover a crucial part of the 'access to food' dimension, and could be incorporated (eventually with a somewhat lower 'weight' than the more direct indicators, see below).

Operationalization of more than one indicator to measure a sub-dimension is a crucial step to ensure the *reliability* of an estimate. To take the analogue of a medical check-up: the physicians' verdict about the health condition of a patient usually is based on more than one kind of measurement (e.g. heart rate, lung function, a variety of different blood values). The reliability of the physician's statement about the health condition of a patient increases with the number of valid indicators that enter the overall assessment. Similarly, if one decides that food access is an important dimension to analyze needs of people struck by drought, then there might be different ways to measure this – e.g. one could measure household income and food prices on the market.

Finally, a note on predefined lists of indicators. One should be aware that such lists of indicators, though useful as a general guideline as argued in the following chapters, also have some major limitations. Indicator lists are never complete and not applicable to all times and places; this depends on the situation that is assessed, as well as on the availability and reliability of information.

Downloaded by [University of Groningen] at 05:15 06 May 2015

Quantification

Third, in a real life exercise, most likely the information available for each indicator will be highly diverse, varying from census data collected by national bureaus of statistics, to incidents reported in the local newspaper. Sometimes, all that may be available are subjective reports from informants. Put differently: what enters a context analysis usually is data representing at least three different measurement scales: (a) interval scales (e.g. expenditures in dollars); (b) ordinal scales (e.g. low, medium, high); (c) nominal scales (e.g. presence vs absence of something). During the quantification step, the analyst needs to translate this heterogeneously scaled data into the measurement scale of the specific context analytical tool to be used. For the Comprehensive Context Analysis (CCA), this is an ordinal scale ranging from '0' to '6' (see Chapter 5), distinguishing conditions of security from conditions with latent or manifest threat. For the stakeholder analysis, these are interval scales ranging from '0' to '100', reflecting e.g. the degree of influence of a stakeholder (see Chapter 12). It is evident that this operation – in which the analyst has to map a qualitative observation or a quantitative fact into a number – entails a lot of subjective judgment, and is of course vulnerable to measurement error. Yet it is this step that forces the analyst to explicate the underlying assumptions, and to transparently document them. Therefore, when working with this quantified information, one should always keep in mind that each number is not more than a placeholder for a qualitative and highly subjective judgment. The main purpose of this quantification step is to facilitate systematic context analysis through comparison both within and across context dimensions, time periods (i.e. pre- and post-emergency) and different crises situations.

Aggregation and weighting

Fourth, it has to be decided how these indicators are related and 'add up' to provide an overall estimate for the overarching construct, i.e. the final assessment for the level of security threat in a specific domain (aggregation). For this step to be carried out, it is necessary that all previous steps have been completed, i.e. that each indicator of a sub-dimension has been translated into a value on the respective (security or stakeholder) scale. Once this has been done, the different values can be aggregated. Though adding them up and dividing by the number of indicators is the most straightforward and most often used technique of aggregation, more refined approaches are sometimes even necessary. The reason is that in the case of simple addition, equal importance is attributed to every single indicator. Whereas this may be warranted in some cases, it may heavily distort the picture in others. In the latter case, introducing weights may be an option.

Weighting is an instrument to adjust the relative importance of an indicator in a set of other indicators measuring an overarching construct. For example, when assessing the level of political security in a region, an analyst might consider the following indicators: level of freedom of the press as measured through incidents

Downloaded by [University of Groningen] at 05:15 06 May 2015

restricting freedom of speech; level of human rights as measured by the number of documented tortures; level of trust in institutions as measured through survey data. The analyst may consider all three dimensions as equally relevant and as valid sub-dimensions and therefore may decide to add up the values of the three. But the analyst may also attach different degrees of relative importance to each of them. For example, he or she may consider the presence of human rights violations as more important than infractions of the freedom of speech and institutional trust. In this case, the analyst may decide to weigh the human rights indicator twice as heavy as the other two, e.g. by multiplying this indicator by 50 percent and each of the other two by 25 percent.

Downloaded by [University of Groningen] at 05:15 06 May 2015

From valid indicators to trustworthy sources

Once it has been decided which indicators are valid, one can look for potential data sources that may generate the desired empirical evidence. Trustworthy sources and reliable data are a major precondition for a good context analysis. A source is 'trustworthy' if the information it provides is not biased towards particularistic interests or is deliberately misleading. The information a source provides is 'reliable' if it accurately describes the respective facts. Note that a source can be trustworthy but at times deliver unreliable information, for example because it may only have partial information on a particular issue. Similarly, there may be cases in which sources that are considered as not trustworthy may provide reliable data, for example because it may be in their interest to do so in specific situations. This implies that the analyst needs to carry out independent checks, one on the general trustworthiness of the source (e.g. based on the source's reputation), and one on the reliability of the information provided by a source. We address trustworthiness issues in this section, and data reliability issues in the next section.

There are a variety of different data sources, ranging from informants, written media (e.g. government reports, official statistics, newspapers), to the various sources on the internet (e.g. blogs, interviews, wikis). Faulty or misleading information can jeopardize the success of a project and the reputation of an organization. Sometimes, deception is exactly the aim of the source, sometimes wrong information is simply the result of human error and at other times it is the result of a bias in the source. How to use the available data in such a way that a reliable and valid analysis can be made? Here we present some first steps that facilitate the correct selection of information to be used for assessing levels and degrees of comprehensive security.

After having made an inventory of potential information sources, one has to do a quick assessment about their trustworthiness. The following questions could help the analyst in this regard:

- Is the source generally trustworthy? For example, are the authors transparent in how they collected the information (method), what their sources were, who were their respondents?

- Is the source independent from partial stakeholders?
- Is the data and/or intelligence presented coherently with other assessments?
- Is the way the data is presented emotionalized or in any other ways biased?

The aim of this assessment is to categorize your sources from most trustworthy to least trustworthy. For example, data coming from an authoritarian government denying the outbreak of cholera in its own country (maybe to avoid an influx of foreign agencies) is of much less value than data presented by an IGO or independent academic institute that claims that there is an outbreak of cholera. Categorizing your sources in this way helps you to deal with such inconsistent or conflicting information. In the ideal situation one only works with the most trustworthy sources; however, if certain information is needed, then one could work with less reliable sources. However, one then needs to carefully assess the potential bias in the information presented and to account for this bias in the analysis. The following verification techniques can help the analyst to classify the reliability of sources:

- background check of the data provider;
- cross-checking the information with the same data of a different source;
- cross-checking the information with different data of a different source.

Tables 4.1 and 4.2 give a more detailed overview of key elements to check when assessing the reliability of sources and the accuracy of data.

The criteria specified in the tables can be used to provide an overall assessment of data quality. 'A1' ratings would reflect the highest quality, since it comes from a highly reliable source and reports confirmed information. Conversely, data with an 'E5' or 'F6' status would represent the other extreme: the source is unreliable (E) or its reliability cannot be judged (F), and it is either highly improbable that the information provided is accurate (5), or the accuracy cannot be judged (6).

These kinds of qualifications can provide important background information when reporting on the results on the different dimensions constituting a context or stakeholder analysis.

From trustworthy sources to reliable estimates

Once trustworthy sources have been identified, the analyst can start with working towards eliciting reliable estimates. The following two techniques are useful to do so.

Triangulation

A key principle of empirical work – whether it is done by an investigative journalist, a judge, a scientist or a humanitarian worker – consists of the independent verification of facts. Facts that have been confirmed by more than one trustworthy

Downloaded by [University of Groningen] at 05:15 06 May 2015

Downloaded by [University of Groningen] at 05:15 06 May 2015

Table 4.1 Source reliability rating matrix

Source rating scale	*Source trustworthiness rating*	*History of trustworthiness*	*Source authenticity*	*Source objectivity*	*Source access to information*	*Source not vulnerable to manipulation*	*Meets number of criteria*
A	Reliable	Yes	Yes	Yes	Yes	Yes	All 5
B	Usually reliable	Yes	Yes or No	Yes or No	Yes or No	Yes or No	History, plus 3
C	Fairly reliable	Yes	Yes or No	Yes or No	Yes or No	Yes or No	History, plus 2
D	Not usually reliable	Yes or No	Yes or No	Yes or No	Yes or No	Yes or No	2 of 5
E	Unreliable	No	Yes or No	Yes or No	Yes or No	Yes or No	n.a.
F	Cannot be judged	No basis for evaluating the reliability of the source					

Source: based on work of the Canadian Intelligence as described by Hibbs-Pherson and Pherson (2013: 100).

Notes
To be rated 'A', the source authenticity must be verifiable, and the source's history of reliability must be verifiable post-factum or by independent verifiable means. To be rated 'A' or 'B', the source's reporting must always be reliable with no significant errors. To be rated 'C', the majority of the source's reporting must be accurate and actionable. To be rated 'D', a minority of the source's reporting must still be accurate and actionable.

Downloaded by [University of Groningen] at 05:15 06 May 2015

Table 4.2 Data accuracy rating matrix

Data rating scale	*Information accuracy rating*	*Independent verifiable means*	*Source subject competency*	*Logical*	*Practical and plausible*	*Consistent*	*Meets number of criteria*
1	Confirmed	Yes	Yes	Yes	Yes	Yes	All 5
2	Probably true	No	Yes	Yes	Yes	Yes	4
3	Possibly true	No	Yes or No	Yes or No	Yes or No	Yes or No	3
4	Doubtfully true	No	Yes or No	Yes or No	Yes or No	Yes or No	2
5	Improbable	No	Yes or No	Yes or No	Yes or No	Yes or No	0 or 1
6	Cannot be judged	No basis for evaluating the validity of the information					

Source: based on work of the Canadian Intelligence as described by Hibbs-Pherson and Pherson (2013: 100).

source are more credible and reliable than facts based on only one source. Establishing the reliability of an estimate by cross-checking it through different sources is called triangulation. Triangulation is even stronger if there is also variation in the type of sources. For example, evidence for the severity of a food crisis would be very strong if it is independently confirmed, e.g. by (a) doctors reporting systematic underfeeding of a large proportion of their patients, (b) a household survey including assessments of caloric intake, conducted by a medical anthropologist working in the field, (c) reports in the local newspaper.

Triangulation attempts usually face two major challenges. First, it will not always be possible to verify the reliability of an estimate. In this case, the analyst needs to decide whether or not the estimate should be included. This may be warranted if the source has a high level of trustworthiness, and there are no alternative estimates available for other indicators covering the sub-dimension. Second, different sources may yield different – and sometimes even inconsistent or contradictory – estimates. In this case, the analyst should first try to resolve the inconsistencies by investigating what might have caused them. For example, the estimate of source A may refer to a slightly different time period or geographical level of analysis than the estimate of source B. For example, does the information refer to a nation, a region or a sub-group in a city? Similarly, estimates based on statistical evidence may be based on different algorithms, or on different definitions of the analytical categories. For example, who classifies as 'unemployed' in official statistics can vary across countries and through time. If this step does not resolve the inconsistency, and the sources are all considered to be equally trustworthy, the analyst may decide to use both of them, e.g. by using their average, and letting the highest and the lowest estimate define the upper and the lower boundary of the confidence interval (see below).

Confidence intervals

Confidence intervals provide interval estimates by defining an upper and a lower boundary for an estimate. The interval spans the range of values that is likely to contain the unknown 'true' value. Interval estimates differ from point estimates, which consist of a single specific value. The size of the confidence interval indicates the reliability of the estimate: the smaller the range, the more reliable the estimate is believed to be. Working with confidence intervals is particularly important in situations where estimates are likely to be inaccurate. This is likely to be the case for most humanitarian settings. We will introduce a method of how to use confidence intervals in measuring and visualization in Chapter 5.

Conclusion

The quality of any context analysis stands and falls with the quality of the evidence on which it is based. This chapter described the most fundamental steps that have to be taken to ensure a decent quality of collected evidence and the estimates that follow from it. Collecting valid and reliable data on emergency

Downloaded by [University of Groningen] at 05:15 06 May 2015

settings of course is not an easy endeavor: there are usually never enough time, staff and other resources available for inquiry and fact checking. However, as this chapter has tried to make clear, this does not mean that some minimum standards of data collection and management have to be violated.

The steps outlined in this chapter force the analyst to be explicit not only about whether or not the evidence meets these minimal criteria, but also – in case the analyst decides to proceed – to provide good reasons why such a violation is considered acceptable, and which precautions are taken to prevent mis- or over-interpretation of the evidence. Finally, it is important to realize that – even in a situation of trustworthy sources and good data quality – the numerical estimates may create the illusion of a precision of measurement that does not correspond with the subjective nature of most judgments that have to be made when conducting a context analysis of emergency settings.

Downloaded by [University of Groningen] at 05:15 06 May 2015