unicef
Office of Research – Innocenti

unicef | for every child

THE UNICEF OFFICE OF RESEARCH – INNOCENTI

The Office of Research – Innocenti is UNICEF's dedicated research centre. It undertakes research on emerging or current issues in order to inform the strategic directions, policies and programmes of UNICEF and its partners, shape global debates on child rights and development, and inform the global research and policy agenda for all children, and particularly for the most vulnerable.

Publications produced by the Office are contributions to a global debate on children and may not necessarily reflect UNICEF policies or approaches. The views expressed are those of the authors.

The Office of Research – Innocenti receives financial support from the Government of Italy, while funding for specific projects is also provided by other governments, international institutions and private sources, including UNICEF National Committees.

For further information and to download or order this and other publications, please visit the website at: www.unicef-irc.org

INNOCENTI DISCUSSION PAPERS

UNICEF Discussion Papers are signed pieces by experts and researchers on current topics in social and economic policy and the realization of children's rights. The aim is to encourage reflection and stimulate wide-ranging discussion.

This is a peer-reviewed series.

Correspondence should be addressed to:

UNICEF Office of Research – Innocenti
Via degli Alfani, 58
50121 Florence, Italy

Tel: (+39) 055 20 330
Fax: (+39) 055 2033 220
florence@unicef.org
www.unicef-irc.org
@UNICEFInnocenti
facebook.com/UnicefInnocenti

IMPACT EVALUATION IN SETTINGS OF FRAGILITY AND HUMANITARIAN EMERGENCY

Shivit Bakrania, UNICEF Office of Research – Innocenti

Nikola Balvin, UNICEF India Country Office

Silvio Daidone, Food and Agriculture Organization of the United Nations

Jacobus de Hoop, UNICEF Office of Research – Innocenti

AUTHORSHIP

All authors made a substantial contribution to the paper and hence are placed in alphabetical order. Corresponding author: Shivit Bakrania, sbakrania@unicef.org. The findings, interpretations and conclusions expressed in this paper are those of the authors and do not necessarily reflect the policies or views of UNICEF.

ACKNOWLEDGEMENTS

CONTENTS

ABSTRACT

What are the objectives of this paper?

This paper seeks to contribute to and catalyse efforts to implement rigorous impact evaluations and other rigorous empirical research in fragile and humanitarian settings. It draws upon qualitative interviews with 10 researchers and 10 research commissioners and funders. It describes what sets apart research in fragile and humanitarian settings; identifies common challenges, opportunities, best practices, innovations and research priorities; and shares some lessons that can improve practice, research implementation and research uptake.

The concluding section provides some reflections on areas of agreement (and disagreement) between researchers and their research commissioner and funder counterparts. Recommendations are provided for increasing the quality and quantity of impact evaluations and other rigorous empirical research in fragile and humanitarian settings, focusing on and highlighting areas of agreement between researchers on the one hand, and research commissioners and funders on the other.

Why is this important?

Although starting from a low base, the quality and quantity of impact evaluations and other research in fragile and humanitarian settings is gradually increasing, demonstrating that the challenges associated with these settings can be overcome. Furthermore, there is increasing interest in generating high-quality evidence, through impact evaluations and other research, on intervention effectiveness in fragile and humanitarian settings, and translating this for the purposes of programme implementation and decision-making. This combination of experience and interest suggests that there is potential to further increase the quality and quantity of impact evaluations and other rigorous research in fragile and humanitarian contexts.

Who should read this paper?

Researchers, practitioners and research funders and commissioners engaged in undertaking, supporting or funding impact evaluations and other rigorous research in fragile and humanitarian contexts.

What are the main findings and recommendations?

Despite the numerous challenges involved, experience shows that rigorous impact evaluation and research is both possible and necessary in fragile and humanitarian settings to ensure that interventions and responses are effective.

While not identical, the views of researchers and those of research commissioners and funders largely converge. Many of the researchers interviewed stated that donors are willing to support and fund research in fragile and humanitarian settings. Interviews with commissioners and funders reinforced this willingness to support impact evaluation and other research efforts in complex contexts. There is even a sense that there are opportunities in fragile and humanitarian settings that may not exist in more stable settings, because of the level of interest and the fact that funding is flowing in and interventions are being scaled up. Therefore, there is potential to further increase the quality and quantity of impact evaluations and other research, which could be realized through:

- unifying the approach to impact evaluation and research among researchers, research commissioners and funders, and implementing partners by providing space for discussion and coordination

- enabling researchers to work in challenging and risky settings by providing timely and rapid funding processes, and including researchers in implementing agencies' risk management and security procedures

- increasing appetite for impact evaluation among implementing partners by engaging them in all stages of the evaluation/research cycle

- strengthening the capacity of local researchers and research organizations.

1. INTRODUCTION AND OBJECTIVES

Rigurous impact evaluations and empirical research studies examining the role of support programmes in fragile and humanitarian settings are few and far between. While starting from a low base, the quality and quantity of evidence has increased over time, however, and recent studies show that it is possible to generate evidence even in highly challenging contexts. This paper seeks to catalyse efforts to implement rigorous impact evaluations and other rigorous empirical research in fragile and humanitarian settings and thus improve the quality of evidence upon which decisions are based. Ultimately, the goal of the paper is to help support the delivery of appropriate, high-impact assistance to people affected by conflict, natural disasters and other humanitarian emergencies.

As such, this paper complements earlier papers with similar aims. For example, Gaarder and Annan (2013) examine the practices of impact evaluators in the peacebuilding sector to see how they address evaluation design, data collection and conflict analysis. Puri, Aladysheva, Iversen, Ghorpade and Brück (2017) explore the methodological challenges and options associated with generating high-quality evidence to inform humanitarian action. In both cases, the authors argue that it is both possible and necessary to undertake rigorous impact evaluations and other rigorous research in settings affected by fragility and by conflict and other humanitarian crises. The present paper reinforces this argument and contributes to the body of evidence by drawing upon qualitative interviews with 10 researchers who have carried out exemplary work in challenging humanitarian contexts and 10 colleagues working in organizations operating in humanitarian contexts and commissioning and/or funding research. As part of the analysis, the paper highlights areas of agreement and debate between researchers and their counterparts who commission or fund research in humanitarian and fragile settings.

Important issues addressed during the interviews included:

■ aspects that are unique to conducting impact evaluation and other empirical research in fragile and humanitarian settings

■ common challenges, opportunities and best practices for impact evaluation and other empirical research in fragile and humanitarian settings

■ key innovations that could help to overcome the challenges in conducting impact evaluation and other empirical research in fragile and humanitarian settings

■ key research questions and priorities.

Section 2 of this paper discusses key concepts and definitions used for this study, including the definitions used for 'fragile' and 'humanitarian' settings, and the distinction between rigorous 'impact evaluation' and rigorous 'research'. Section 3, the literature review, presents an overview of key debates and discussion within the literature on undertaking impact evaluation and research in fragile and humanitarian settings. Section 4 presents the methodology used for compiling this paper. Section 5, the analysis, presents the perspectives of the interviewed researchers followed by the perspectives of the research commissioners and funders working for organizations in humanitarian contexts. The concluding section reflects on areas of agreement (and disagreement) between the two groups and provides some recommendations for enabling rigorous impact evaluations and other rigorous research in fragile and humanitarian settings.

2. DEFINITIONS AND FRAMING

The boundaries between what is considered a 'fragile' or 'humanitarian' setting are often blurred, and the terms are often used interchangeably, including by the interviewees involved in this study. It is therefore useful to provide some clarity on how fragile and humanitarian settings were conceptualized for the purposes of this study.

Understandings of fragility have progressed from narrow definitions focused on the failure of so-called fragile states "to provide basic services to poor people because they are unwilling or unable to do so" to broader, multidimensional understandings (Organisation for Economic Co-operation and Development [OECD], 2006, p. 147). OECD defines fragility as "the combination of exposure to risk and insufficient coping capacity of the state, system and/or communities to manage, absorb or mitigate those risks" (OECD, 2016, p. 22). The risks in question relate to vulnerabilities in five dimensions: economic (e.g., macroeconomic shocks), environmental (e.g., natural disasters and disease epidemics), political (e.g., corruption and oppression), security (e.g., violence and crime) and societal (e.g., social and cultural cleavages).

A humanitarian crisis or setting can be defined as an "event or a series of events in a country or region that cause serious disruption to the functioning of a society, resulting in human, material, or environmental losses which exceed the ability of affected people to cope using their own resources. A crisis may be further classified according to its speed of onset (sudden or slow), its length (protracted) or cause (natural or man-made hazard or armed conflict)" (Inter-Agency Standing Committee, 2015, p. 2).

From the definitions provided, it is arguably the case that fragile settings and humanitarian crises exist on a continuum. Fragile settings are particularly susceptible to humanitarian crises and emergencies when that setting is placed under some form of acute stress. Fragile settings may also be more susceptible to prolonged humanitarian crises, because institutional arrangements in such contexts could help to preserve the conditions of a crisis (Berman et al., 2016; Di John, 2008). From the perspective of this paper, it is important to recognize that both fragile and humanitarian settings, no matter how they are defined, present challenges to data collection, research and evidence generation that are more acute than in other, more stable development settings.

Another definitional issue relates to the distinction between rigorous 'impact evaluation' and rigorous 'research' more broadly.[1] This project initially began with the intention of focusing on impact evaluations that identify quantitative impacts through the construction of an experimental or quasi-experimental comparison group (for a description of this type of impact evaluation, see Gaarder and Annan, 2013, for instance). Interview discussions often moved beyond this, however, into broader issues of conducting rigorous empirical research and evidence generation in fragile and humanitarian settings. It can be argued that the challenges to ensuring the rigour and quality of impact evaluations also apply to empirical research more broadly in fragile and humanitarian settings. Many of the findings presented here with respect to impact evaluation are thus equally valid for evidence generation more broadly, especially research comprising primary data collection.

---

1   According to UNICEF, an 'impact evaluation' is a type of evaluative research, while 'research' refers to the "systematic process of the collection and analysis of data and information, in order to generate new knowledge, to answer a specific question or to test a hypothesis" (UNICEF Office of Research – Innocenti, 2014, p. 2). Research becomes rigorous when it is underpinned by "scientific notions of proof, validity, reliability, and has minimised bias" (Blanchet et al., 2018, p. 4).

3. LITERATURE REVIEW

This brief literature review presents an overview of key debates and discussion within the literature on undertaking impact evaluation and research in fragile and humanitarian settings.

This paper is not the first to reflect on impact evaluation in these contexts, however. For example, Buttenheim (2010) synthesizes a set of guiding principles and analytic frameworks for impact evaluation in the aftermath of natural disasters, pointing out the challenges of conducting quality impact evaluations in post-disaster settings. Burde (2012) examines randomized controlled trials (RCTs) in conflict-affected settings and stresses the importance of properly combining qualitative and quantitative methods to avoid an over-reliance on one approach and to increase contextual understanding. Gaarder and Annan (2013) present some examples from a collection of impact evaluations of conflict prevention and peacebuilding interventions, and examine the practices of impact evaluators in the peacebuilding sector to explore how they address evaluation design, data collection and conflict analysis. Puri, Aladysheva, Iversen, Ghorpade and Brück (2015) explore the "methodological options and challenges associated with collecting and generating high-quality evidence needed to answer key questions about the performance of humanitarian assistance" (p. 1). With the help of six case studies, they show how to apply impact evaluation in humanitarian settings, respecting ethical concerns and improving assistance. Verwimp, Justino and Brück (2019) take stock of research on the microeconomics of violent conflict. They argue that innovative data collection methods and new technologies, together with the right partnerships, mean that information and data can be obtained from active conflict zones.

The increasing interest in this topic reflects the fact that although significant funds are being invested in interventions in fragile and humanitarian settings, rigorous evidence on their impact is limited (Ager et al., 2014; Aker, 2017; Idris, 2019). Systematic reviews of cash-based approaches and health interventions in fragile and humanitarian settings find that, in addition to the limited quantity of evidence, the quality of studies is also generally weak (Blanchet et al., 2017; 2018 Doocy & Tappis, 2016; Mishra, 2017). Indeed, while the quality and quantity of evidence has increased over time, the most rigorous studies have been conducted only from year 2000 onwards (Blanchet et al., 2017).

Strengthening the evidence base for interventions in fragile and humanitarian settings is seen by numerous authors as imperative for development organizations, especially given the growing scale and complexity of programming in such contexts (Ager et al., 2014; Berman et al., 2016; Blanchet & Roberts, 2015; Blanchet et al., 2018; Peterman, de Hoop, Cuesta & Yuster, 2018). A key issue is that fragile and humanitarian settings are characterized by specific risks, vulnerabilities and flux, and questions remain as to how evidence generated in more stable settings can be generalized to such contexts (Blanchet et al., 2017).

Experience thus far highlights a number of challenges to conducting rigorous impact evaluation and other rigorous research, many of which were identified by interviewees in this study and are explored in more detail in the following sections. An overriding challenge in these contexts is the uniqueness of interventions, which aim to respond – often rapidly – to events that have not been anticipated. This, combined with the acute vulnerability of affected populations and the commonly experienced rapid changes in events, means that research – and particularly impact evaluation, which may require a different approach to that of programme delivery – is seldom a priority in fragile and humanitarian settings (Ager et al., 2014).

This has several implications for impact evaluation and other types of rigorous research. For example, it is difficult to implement RCTs because the urgency of humanitarian assistance is often at odds with the necessity to randomly assign beneficiaries into treatment and control groups. Furthermore, withholding treatment in emergency situations could be considered unethical (Puri et al., 2017) and may exacerbate conflict between groups. Related to this, attribution of effects to any one intervention is also challenging as often an array of actors will be operating in settings that necessitate a rapid response (Ager et al., 2014). Other ethical challenges beyond those associated with RCTs include the capacity of participants to give informed consent, and the imperative to *do no harm* if the research relates to sensitive subjects (Berman et al., 2016). Another frequently mentioned challenge in such contexts is the scarcity of relevant data or the lack of access to data, including administrative data (Aurino, Brück, Daidone, Natali & Schwab, 2018; Idris, 2019). The limited quantity of evidence in fragile and humanitarian settings could also be indicative of limited technical capacity among researchers and research organizations (Blanchet et al., 2017), including local organizations.

There is some recognition in the literature that the challenges, for both researchers and programme beneficiaries, are gendered. Bahati (2019) states that female researchers often encounter enormous challenges in fragile and conflict-affected settings. These include prejudice, distrust, sexism and the threat of physical danger. Presser (2005) describes how gender can "influence esteem, trust, and rapport, which facilitate or thwart access to data" (p. 2072). Huggins and Glebbeek (2003) argue that gender plays a role in the study of male-dominated institutions, whereby one's gender can affect the negotiation and maintenance of power between the researcher and the researched. Furthermore, inter-gender dynamics can thwart some research objectives and promote others. For example, female researchers can sometimes negotiate access in settings where men would struggle, such as when discussing sensitive subjects with other women (Bahati, 2019).

Ideally, impact evaluations should be gender-sensitive (i.e., include gender-disaggregated data and analysis), however, this may not be possible in certain fragile and humanitarian settings, especially where institutions are male-dominated. For example, in an impact evaluation of aid in north-east Afghanistan, social customs and security challenges meant that men had to accompany female researchers to survey female respondents, which made it impossible for teams to work effectively. Eventually, a decision was made to include only male respondents in the survey (OECD, 2012).

While there are shortcomings in the literature on *how* to do research in fragile and humanitarian settings (Idris, 2019), recent studies do show that it is possible to generate rigorous evidence on programme impacts in these contexts. For example, a review conducted by Bailey (2013) identified 18 studies and evaluations that compared cash, vouchers and/or food aid to analyse the appropriateness and impacts on food consumption of these different tools. Seven of the interventions reviewed were designed and implemented according to a randomized design, whereby (clusters of) households were randomly assigned to different 'treatment' groups (e.g., cash, vouchers, food aid). Four were part of a multi-country study undertaken by the International Food Policy Research Institute (IFPRI) on behalf of the World Food Programme (WFP) to analyse the effectiveness and efficiency of cash, vouchers and food aid; the countries selected were Ecuador (Hidrobo, Hoddinott, Peterman, Margolies & Moreira, 2014), the Niger (Hoddinott, Sandström & Upton, 2018), Uganda (Gilligan, Margolies, Quiñones & Roy, 2013) and Yemen (Schwab, Margolies & Hoddinott, 2013). Four further studies also used RCT designs, IFPRI/WFP research in Malawi (Audsley et al., 2010 and Sri Lanka (Sandström & Tchatchua, 2010; Sharma, 2006) and Tufts University/Concern Worldwide research in the Democratic Republic of the Congo (Aker, 2017). One study, Tranchant et al. (2018) used a quasi-experimental design, combining propensity score matching with difference-in-differences, to estimate the impacts of WFP food assistance in conflict-affected areas of Mali.

In recent years, the Syrian refugee crisis has received a great deal of attention from both international donors and the research community. Lehmann and Masterson (2014) analyse the impacts of the winter cash transfer programme run by the United Nations High Commissioner for Refugees (UNHCR) with the objective of keeping people warm and dry during the cold winter months. Battistin (2016) analyses the impact of multi-purpose cash assistance for Syrian refugees still living in Lebanon during the sixth year of the crisis. Finally, a mixed methods study by de Hoop, Morey, Ring, Rothbard and Seidenfeld (2018) examines the effects of a cash transfer programme for the benefit of Syrian children in Lebanon. All three studies used a regression discontinuity design.

The evidence generated by these types of studies has the potential to help shape and enrich debates around programming and decision-making in fragile and humanitarian settings. As argued by Puri et al. (2017), rigorous "impact evaluation methods can be used successfully and in an ethical manner to distil lessons about how to deliver humanitarian assistance effectively and efficiently" (p. 520). In keeping with this sentiment, the following sections present the perspectives of researchers and research commissioners and funders with actual experience of conducting or commissioning research in fragile and humanitarian settings.

4. METHODOLOGY

Semi-structured interviews were conducted with 10 expert researchers with experience of conducting impact evaluations and research in fragile and/or humanitarian contexts, and 10 research commissioners or funders who have commissioned or funded impact evaluations in fragile and/or humanitarian settings.[2]

The 10 researchers (and their affiliation at the time of the interview in 2017) were: Jenny Aker, Tufts University; Jeannie Annan, International Rescue Committee (IRC); Tilman Brück, International Security and Development Center (ISDC) and Leibniz Institute of Vegetable and Ornamental Crops (IGZ); Dana Burde, New York University; Fotini Christia, Massachusetts Institute of Technology (MIT); John F. Hoddinott, Cornell University; Daniel Masterson, Stanford University; Daniel Maxwell, Tufts University; Amber Peterman, UNICEF Office of Research – Innocenti; and Jyotsna Puri, Green Climate Fund.

The 10 research commissioners and funders from organizations (their affiliations at the time of the interview in 2017) operating in humanitarian contexts were: Ulrich Assankpon, Food and Agriculture Organization of the United Nations (FAO); Prisca Benelli, Save the Children UK; Marie Gaarder, International Initiative for Impact Evaluation (3ie); Marie-France Guimond (IRC); Göran Holmqvist, Swedish International Development Cooperation Agency (Sida); Tara Kaul, 3ie; Massimo La Rosa, European Commission Humanitarian Aid Office (ECHO); Claire Mariani, UNICEF; Urs Nagel, UNICEF; and Jonathan Patrick, UK Department for International Development (DFID).

Details about the background of each interviewee are provided in Appendix A. The initial intention was to interview a larger number of individuals; after reaching saturation point on topics and themes, however, it was decided that 20 interviews were sufficient for the purposes of this study.

Standardized interview guides were prepared to structure the conversations: one for the interviews with researchers and another for the interviews with research commissioners/funders. These guides can be found in Appendix B. Interviewees were asked not only to respond to the questions and themes included in the appropriate guide, but also to raise any relevant additional issues not initially foreseen by the authors of this paper.

Interviews were recorded, and paraphrased transcripts were prepared based on the recordings and notes taken during the interviews. These were thematically analysed and grouped according to the objectives of this paper (e.g., to explore challenges and innovations in impact evaluations in settings of fragility and humanitarian emergency) by two independent researchers, who then discussed and reconciled their analysis and qualitative coding. The most salient findings from these thematically analysed transcripts are presented in this paper. The main themes and sub-themes from the qualitative coding are included in Appendix C.

---

2   Initially, the goal was to focus the interviews on impact evaluations of social protection programmes in fragile and humanitarian settings. The social protection focus was discarded, however, because the results of the interviews apply to impact evaluation and rigorous research more broadly in these contexts.

5. ANALYSIS OF INTERVIEWS WITH RESEARCHERS

This section presents the key findings from the interviews with researchers and the subsequent thematic analysis of the transcripts.

5.1 Challenges

Data collection in settings of fragility and emergency is often risky and challenging. The themes of access to people and locations, and of safety and security for researchers emerged in many interviews. Dana Burde stated that interviewing certain segments of a population can be challenging, because certain geographical locations are difficult to access. This can be due to a lack of transport infrastructure in remote and rural locations, or the need to hold a permit or clearance to visit some sensitive locations, whether this is a requirement of governments, employers or funders. The fluid nature of emergencies means that locations where research has been planned can suddenly become 'no-go areas'.

Therefore, there are often security concerns that one needs to consider in research planning. In some cases, it is impossible for foreign researchers to participate in the fieldwork, and local researchers need to be trained and managed remotely. For example, Daniel Maxwell mentioned his experience in Somalia in Al Shabab-controlled areas during the 2011 famine and recounted how field teams "accidentally ended up in the line of fire". He worked intensively with field teams, built up strong relationships with them and had extensive debriefing sessions. This increased the field teams' skills and the quality of fieldwork. Similar experiences were reported by Dana Burde, who relied on locally recruited staff in Afghanistan, not only to overcome the risks for international researchers, but also to build up a relationship of trust with local counterparts, which led to long-term working relationships.

Fragile and humanitarian contexts are inherently unpredictable, and research is often subject to time constraints, which presents uncertainty. Programmes often change rapidly in response to external events, implementation cycles are short and decisions need to be made quickly. It is therefore not possible to have the luxury of waiting for an optimal time for data collection. This means that one has to be prepared for the research question to change, for access to participants to change and/or for research objectives to change. Researchers stressed the need for creativity, flexibility and adaptability in such contexts, where it is essential to plan for potential risks and factor in security and infrastructure considerations. As mentioned by Jenny Aker, this can help the researcher to provide results that are timely and appropriate to the programme implementation cycle.

Further uncertainty arises because populations in humanitarian emergencies are often transient or mobile. Sampling is challenging, as is the application of traditional RCTs, which may not necessarily work in this type of setting. Several researchers mentioned the difficulty of creating a valid counterfactual for evaluating the impact of interventions used to respond to rapid onset events. One key issue is that when a control group is created, it is seldom valid. Another key issue concerns covariability: Large areas are affected during a humanitarian emergency, and it is difficult to find locations or beneficiaries similar to the affected population for the purposes of comparison.

Attribution is also a challenge, owing to the multiplicity of actors often working in the same location. Therefore, as Jyotsna Puri explained, it is "difficult to: (1) understand the frequency and the doses of a single intervention; and (2) attribute impact to a single intervention in a context of multiple interventions. Coordinating amongst these actors and their programmatic priorities makes this complicated."

Local research capacity and infrastructure was another frequently mentioned challenge. At the individual researcher and research organization level, there is sometimes a lack of knowledge and skills to undertake rigorous research and apply the appropriate research designs in complex situations. Furthermore, in contexts where there are few data collection institutes, principal researchers need to spend a significant amount of time training enumerators. A lack of knowledge and experience can also mean that new technologies are not capitalized on, for example, to construct baselines or to track populations on the move.

Some researchers stated that understanding of and capacity for impact evaluation and research in fragile and humanitarian settings is increasing, in line with the increased quantity of evaluations and research being undertaken in such contexts. As Daniel Masterson stated, "with increased understanding of impact evaluations … there is a much wider range of people aware that a proper research design and questionnaire wording is necessary for causal inference." He made a caveat to his statement, however, in recognition that more needs to be done to ensure quality, stating: "a lot of people now do research that looks like it is high quality, but may not actually be of high quality". Evidently, the building of partnerships with local research organizations and the training of local researchers is fundamental to improving capacity and skills.

Despite these challenges, some researchers argued that many of the barriers that exist in fragile and humanitarian settings are also present in more stable contexts. For example, Daniel Masterson stated that some of the "biggest challenges are probably similar across contexts"; such challenges include training enumerators and the trade-off between speed and quality. Similarly, Dana Burde argued: "many aspects between [fragile and non-fragile] contexts are similar, and many are different". This suggests that some of the issues related to undertaking impact evaluation and other rigorous research in fragile and humanitarian settings are not hugely different to those encountered by researchers in other, more stable development settings. What *is* vastly different between the settings is *how* researchers must operate.

5.2 Ethics

There are important ethical challenges to be considered when carrying out impact evaluations and other research in fragile and humanitarian settings. Populations tend to be highly vulnerable and the principle of *do no harm* must be applied in all humanitarian settings; conflict and humanitarian challenges cannot be exacerbated in any way by the research, nor can participants be put at risk. For example, asking about respondents' religious backgrounds may be sensitive in settings of sectarian conflict.

Informed consent is arguably even more critical for populations affected by conflict and also by high levels of illiteracy, where consent forms can cause suspicion and disquiet. Fotini Christia contended: "understanding what 'informed consent' means by people who are traumatized is hard. Are they really free agents? Do they feel pressured or compelled to participate? Or are they strategically giving information to access aid?"

It is important to consider the backgrounds of those who do the interviewing. Enumerators should come from a neutral group, not a group involved in the conflict. Similarly, it is important to consider the setting in which the interview takes place. For instance, displaced people living in camps may feel obligated to take part in research sanctioned by authorities.

There is also a sense that implementing RCTs has several ethical implications. Establishing a control group by deliberately excluding a subgroup from participation in a support programme may be indefensible in an emergency setting and could potentially be a source of unrest.

Despite these challenges, researchers contended that ethical challenges should not be a deterrent to rigorous research in fragile and humanitarian settings. It is important for people living in settings of conflict to have the opportunity to be heard. As mentioned by Dana Burde, "millions of people live in countries affected by conflict and go about their business every day. Not evaluating programmes in these settings is unethical." Moreover, as underlined by Tilman Brück, researchers should not be afraid to discuss the humanitarian emergency with respondents. The emergency is a reality of people's day-to-day lives; it is not necessarily a sensitive topic in and of itself and should not be avoided. Furthermore, humanitarian settings may actually provide valid opportunities to conduct RCTs; resource constraints mean that it is impossible to deliver aid to everyone at the same time, and a phased roll-out of services allows for randomized implementation.

Finally, it may be challenging to obtain ethical clearance in country. Even where institutional review boards do exist, they may lack the capacity to provide ethical approval.

5.3 Researcher skills

Interviewees were asked to comment on whether researchers in fragile and humanitarian contexts need specific skills. Jeannie Annan contended that while a different skill set is not necessarily needed in fragile and humanitarian settings, a different mindset is. Flexibility and adaptability are key, along with some higher tolerance for risks. Beyond the importance of being flexible and adaptable, an emphasis was placed on willingness to incur risk, on contextual understanding and on other non-cognitive or 'soft' skills.

There was a sense that impact evaluation and other types of research in humanitarian settings should be based on a thorough understanding of the wider social and political dynamics of the context within which the programme is being implemented. Therefore, as mentioned by Tilman Brück, researchers should not only be well grounded in their own discipline and methods, but they should also be aware of the discourse in related disciplines that approach the same and related topics from other perspectives. A wide interest in the historical, economic and social aspects of the humanitarian emergency is key.

There was also broad agreement that emotional, social and other non-cognitive skills are crucially important. Key skills include the ability to stay calm in stressful and sometimes rapidly changing circumstances, and the capacity to build strong and trusting relationships with local counterparts and recruited enumerators. This is especially important where the presence of international researchers is potentially disruptive or puts their safety at risk. Furthermore, it is local counterparts who will possess the necessary contextual and cultural knowledge, access to sensitive locations, and language abilities. It may therefore be necessary for international researchers to develop the skills to work remotely, which entails recruiting talented interviewers and training them. This process requires patience, strong interpersonal skills and an understanding of cultural differences.

5.4 Funding

The interviewed researchers offered some insights related to the availability of funding for impact evaluations and other research in fragile and humanitarian settings.

Some felt that donors are showing a willingness to support research in fragile and humanitarian settings – this is a situation that has improved over time. According to Fotini Christia, there are even opportunities in fragile contexts that may not exist in non-fragile contexts, because funding is flowing in and interventions are being scaled up. For other researchers, such as Jeannie Annan, "the challenge of getting funding remains, even though there is increasing interest in evaluation in humanitarian settings". Even so, like others interviewed, her overall experience to date has been fairly positive, with donors demonstrating flexibility and understanding of the various challenges.

Other researchers noted how research is still a hard sell in fragile and humanitarian settings, with lack of foresight precluding the integration of evaluation into programme design at the outset. The urgency of situations, in which every penny counts, often means that donors are more interested in ensuring that all available funds are directed to the response rather than dedicating a portion to costly impact evaluation or research. As Tilman Brück stated, "humanitarian funding is, like mercury, extremely volatile; funding for research, in contrast, is slow and sticky". Funding synchronization is also an issue, according to the same researcher: "when you get the approval for research funding, the funding for the humanitarian work will have disappeared". More systematic and flexible thinking about research funding is therefore key: This can allow funding decisions to happen in concert with research design, well before conflicts break out and disasters occur, because these can often be predicted to some extent.

5.5 Relevance of research to policymakers and practitioners

Researchers were asked to reflect upon how research in fragile and humanitarian settings can be made relevant for practitioners. Tensions and mismatches often exist between the expectations of researchers and implementing partners. For example, implementing partners may not want to feel like they are being evaluated; they want to feel respected and want recognition for their work in difficult circumstances. Researchers must be responsive and recognize the constraints that partners face.

Engaging with implementers throughout all stages of the research cycle is seen as important to ensuring that research is demand-driven. This begins with engaging implementing partners at the design phase and gleaning their input on the research question, objectives and the plan for analysis. During the research and implementation cycle, it is important to provide results in a timely fashion: If a partner learns that the programme is not (as) effective, finding this out early allows them time to deal with issues.

Another strand of responses focused on the importance of implementation and operational research. Jyotsna Puri and John Hoddinott both contended that there are still major evidence gaps on the mechanisms underlying successful interventions in fragile and humanitarian settings, and that process-oriented questions are very important to implementing partners.

Researchers also highlighted that there is room for an improved understanding and appreciation of evidence generation among policymakers and implementing partners. Knowledge generation takes time, energy and financial commitment. There was agreement among the researchers, however, that when such an investment is made, there is real potential and merit in research that measures the effects of humanitarian programming and shows pathways towards its improvement.

5.6 Innovations

Researchers were asked whether there are any innovative approaches – such as technological advances – which they consider promising for impact evaluation and research in fragile and humanitarian settings. Responses were varied and covered several themes. Appendix C provides a full list of the themes that emerged around innovative approaches.

Approaches that use a geographic information system (GIS) or the Global Positioning System (GPS) were frequently mentioned. These technologies can be used to randomly identify dwellings or locations for constructing sampling frames, to track enumerators, to anticipate natural disasters or to detect migration patterns.

Mobile phones were also frequently mentioned. These can be used to enable telephone interviewing or phone surveys for rapid assessments, to send Short Messaging Service (SMS) data to complement household surveys, and to receive feedback on service users' experiences of services/programmes via dedicated hotlines.

Several researchers commented on non-technological innovations related to the incorporation of different types of data, arguing that administrative data are underused. Tilman Brück suggested that more needs to be done so that this type of data can be used, by: (1) limiting any disconnect between data collection efforts by national and international agencies; (2) establishing national strategies for openly available data; (3) empowering or strengthening the ability of national (statistical) agencies to collect data; (4) providing researchers access to the administrative and monitoring data of international agencies, where possible; and (5) thinking about designing monitoring and evaluation data collection efforts such that they can also feed into a wider knowledge generation agenda. Other researchers argued in favour of 'lighter' data collection, whether by reducing survey length, using a narrower set of questions or by conducting limited surveys using mobile phones.

5.7 Research designs and methodologies

Interestingly, several researchers commented on innovations and creativity in terms of research design and methodology, and several responses include critiques of technology. For example, some interviewees were sceptical of the utility of technologies in fragile and humanitarian settings, arguing that technology cannot replace credible research designs and methods. Moreover, there is no replacement for face-to-face surveys: Well-worded questions offer more precision than other, technological data sources and it is difficult to draw inferences from highly generalized big data.

In terms of research design specifically, responses suggest that alternative approaches to traditional RCTs may be desirable in challenging contexts. Jeannie Annan argued that quasi-experimental approaches are underused. Fotini Christia offered list experiments and conjoint experiments (survey techniques to obtain truthful information on sensitive topics) as possible additional approaches. Daniel Maxwell stated that qualitative evidence on what is happening in no-go areas may be better than no evidence at all, since the difficulty of obtaining representative samples may preclude quantitative research. For him, undertaking research in such contexts involves asking what is the least bad, least risky and least time-consuming approach that will not detract from the quality of research.

6. ANALYSIS OF INTERVIEWS WITH RESEARCH COMMISSIONERS AND FUNDERS

This section presents the key findings from the interviews with research commissioners and funders and the subsequent thematic analysis of the transcripts.

6.1 Experiences of guiding and contracting out impact evaluations

Research commissioners and funders were asked to reflect on their experiences of guiding and contracting out impact evaluations in fragile and humanitarian settings. There was a degree of agreement with researchers on the key challenges, such as the challenges related to access, security and data collection. Other perspectives unique to research commissioners and funders also emerged.

Commissioners and funders concurred with researchers on key skills needed, such as the ability to be flexible and adaptable, and on the lack of research capacity among local partners. In reflecting upon their experiences of working with researchers, however, commissioners and funders revealed a few criticisms. For example, Claire Mariani described how research consultants often arrive with a predetermined approach, which remains rigid throughout the evaluation. Jonathan Patrick suggested that traditional evaluators and academics are seldom equipped to work in fragile and humanitarian contexts, even in terms of completing practical tasks such as transferring funds into the country. He added that it is vital that researchers work closely with and seek advice from local implementing partners.

In a related strand of comments, commissioners and funders also emphasized the difficulties they have faced in finding local researchers or organizations with the required capacity to undertake rigorous evaluations and research. For example, Prisca Benelli stated that while it is hard to generalize, in some fragile and humanitarian settings it is difficult to find individuals or organizations for data collection, because skilled individuals have often left the country. There was a consensus that technical capacity needs to be strengthened in such contexts.

Several commissioners and funders commented on the management of research. Additional layers of management exist where local capacity is lacking, because of the need to engage both international and local partners to undertake different roles and functions. In Ulrich Assankpon's experience, it is difficult to source one organization that can do everything and so two contracts are often required: one with a local partner with contextual knowledge for data collection, and another with a partner (typically international) experienced in conducting research for the data analysis. Several funders mentioned the challenges of actually finding skilled international consultants who are willing to undertake work in challenging contexts. Massimo La Rosa succinctly stated: "Basically, the risk to work in these countries is negatively correlated with the technical skills available." This necessitates either contracting a small number of local partners, many of whom lack the necessary skills, or hiring international partners to work and manage activities remotely.

Integration of evaluation and implementation emerged as another key issue. For Jonathan Patrick, rather than contracting large external organizations for evaluation, researchers embedded in the implementing organization are more likely to be trusted and deployed more quickly, and their results are more likely to be used. Indeed, evaluation should be seen as a cooperative learning activity, with evaluators and implementers working closely together to 'co-produce'. One way of ensuring this is to form a working and equal partnership with research organizations, rather than contract out the work. A partnership approach also helps to mitigate security and access risks. For example, Marie-France

Guimond mentioned that IRC has its own research team, which assists programme staff with designing impact evaluations. She stated that this team works "with research partners and they understand the approach of working on the evaluation. They include the questions important to both the research partner and IRC. Publications are co-authored by IRC staff and staff from the partnered research institute."

The provision of quick and timely results emerged as another priority. Jonathan Patrick argued that many evaluations are poorly designed, stating: "standard models developed in other settings are often shoehorned into humanitarian settings". For him, nimble impact evaluations focusing on smaller questions, which deliver faster results, are the way to go. Claire Mariani emphasized that approaches to evaluation should not hamper implementation. A balance needs to be found "between the urgency and demand of the situation, and the need to generate rigorous evidence". There also emerged a sense that certain humanitarian contexts lend themselves better to evaluation and research than others – for example, a protracted crisis as opposed to a rapid onset event.

6.2 Ethics

A number of ethical issues were mentioned by research commissioners and funders, including the implications of working with local partners in conflict-affected settings. Prisca Benelli stated that sometimes a local partner may be an actor in the conflict, or present barriers to collecting data, even when these data are not sensitive. It is vital to think carefully about whom to partner with; even academic institutions can have political affiliations. Local community members who are deployed as enumerators may also be affiliated with various entities with power or standing in a particular context. It is thus necessary to employ methods to ensure that responses are as objective as possible.

Ethical concerns were raised around the use of control groups in RCTs, expanding upon some of the points made by researchers. Claire Mariani argued that it should be assumed in emergency settings that there will not be a control group. In contrast, Urs Nagel reinforced a point made by some researchers that limited resources may mean that randomly selecting participants is actually the most ethical way to assign a programme or intervention. Jonathan Patrick suggested that the focus should be on measuring "the effects of an intervention on the best-known alternative" rather than making comparisons with a control group that receives no intervention.

6.3 Raising resources in fragile and humanitarian settings

Commissioners and funders were asked to reflect on raising resources for data collection and analysis in fragile and humanitarian contexts. As part of this, they were asked for their perspectives on whether donors were more or less inclined to fund impact evaluation and research, and why.

The overall sense gleaned from responses is that donors are generally open to and supportive of impact evaluation and rigorous research in fragile and humanitarian settings. This represents a shifting trend, whereby donors have moved from a position of ensuring that maximum resources are used to save lives, to recognizing the importance of achieving value for money and results. From the perspective of many interviewees, an obstacle to increasing the quality and quantity of impact evaluations and research is not donor willingness, but convincing implementing partners who may have a low appetite for such efforts.

6.4 Ability of donors and funders to use rigorous impact evaluations and other rigorous research in decision-making

Responses from commissioners and funders were varied on this topic, with some stating that they have been able to use impact evaluation and other research and others stating that they have not.

Those who responded that they have been able to use research to guide funding, policy and programming decisions made the caveat that their use of research has not always been systematic and that making decisions also entails consideration of other factors such as capacity, sustainability and political will.

Context specificity and the generalizability of findings were key issues for those who were less positive about their ability to use research. Fragile and humanitarian settings are specific contexts, and findings from one such complex context may not be generalizable to another. A further issue is the speed at which results are needed; by the time a study's findings are disseminated, the intervention may have progressed some way through its implementation cycle. Other interviewees simply felt that there is insufficient evaluative evidence from fragile and humanitarian contexts.

6.5 Key research questions and priorities

Research commissioners and funders were asked what they consider the key research questions and priorities for fragile and humanitarian contexts. The responses are very diverse, covering a wide range of themes with little consensus among interviewees. Appendix C provides a list of the key research questions for these settings.

The need for implementation research was one of the few areas where there was some consensus among commissioners and funders, further supporting comments made by researchers. Commissioners and funders were keen to know how to most effectively implement programmes and deliver services in fragile and humanitarian settings (on this topic, see also Cherrier, 2019). Another area of consensus was around how social protection programmes, especially cash transfer programmes, can be used for multiple purposes or combined with other interventions or services.

Another issue that came up was the humanitarian–development divide. Göran Holmqvist stated that this issue is at the top of his agenda. Similarly, Massimo La Rosa emphasized the need for research on the nexus between humanitarian assistance and longer-term development and how this can be strengthened to enable more sustainable action.

7. DISCUSSION AND RECOMMENDATIONS

This paper summarizes the experiences of researchers and research commissioners and funders in carrying out rigorous impact evaluations and other rigorous research in fragile and humanitarian settings. In doing so, some common challenges, opportunities and suggestions for best practice have emerged. Areas of agreement (and disagreement) between researchers and their research commissioner and funder counterparts are reflected upon below.

7.1 Key challenges

Areas of agreement: There was agreement between researchers and research commissioners and funders on some key challenges related to undertaking impact evaluation and other rigorous research in unpredictable environments where there are time constraints and where there may be transient populations. These include: (1) challenges related to access, security and data collection; (2) local partners lacking the capacity to undertake impact evaluation and other rigorous research; (3) the presence of multiple actors and organizations in humanitarian settings, which makes it difficult to attribute impact to a single intervention; and (4) the difficulties of implementing RCTs and creating a valid counterfactual. On this last point, researchers commented on the difficulties and potential alternative approaches, while commissioners and funders were arguably more sceptical about the use of RCTs in emergency contexts.

Other challenges were mentioned by both researchers and commissioners and funders, but approached from slightly different perspectives. Both sets of interviewees recognized that data collection is a challenge. Researchers focused on the process of data collection, the instruments and the skills needed for remote management of this process, and the monitoring and quality control of data collection by local partners. Meanwhile, commissioners and funders focused on different approaches to collecting data and the use of different types of data, including administrative data.

Both groups mentioned ethical issues related to RCTs and to partnering with local institutions that may be involved in unethical practices. Researchers tended to emphasize ethical issues more than research commissioners and funders, however. Ethical issues mentioned by researchers include: misperceptions that participation in research may reap some form of benefit; the importance of doing no harm to avoid putting participants at risk; issues around informed consent; and issues to do with the capacity of local institutional review boards.

The two groups differed in the extent to which they felt that evaluation and research needs to be integrated with implementation. Researchers highlighted the fact that there can be tensions between researchers and implementers, and offered increased engagement as a potential solution. Commissioners and funders went further, stressing the need for evaluation and research that is integrated with implementation.

Areas of disagreement: The most marked differences in opinion between the two sets of interviewees were observed in the following areas: flexible and adaptable approaches to impact evaluation; incentives; and research management.

Researchers stressed the need for flexibility and adaptability when conducting impact evaluation in fragile and humanitarian settings. This relates to the mindset of researchers and the need for an understanding of social and political dynamics, for social and emotional skills in stressful

circumstances, and for the ability to work with local researchers. Despite this, commissioners and funders nevertheless asserted that many academic researchers do not have the skills to work in fragile and humanitarian contexts, and argued that research consultants arrive with predetermined approaches and cannot adapt well to changing circumstances.

Researchers stressed the lack of incentives for conducting impact evaluation and other rigorous research in humanitarian settings, which present various risks. From a researcher's perspective, the anticipated pay-off of working in stable settings may be higher, as research is faster to implement and quality can be more readily controlled in such contexts. In contrast, commissioners and funders perceived a lack of willingness among researchers to work in humanitarian settings.

Research commissioners and funders placed more of an emphasis on research management. In particular, they emphasized the difficulties faced in finding research organizations with the necessary analytical skills and contextual knowledge to conduct evaluation and research in fragile and humanitarian settings. More layers of management are added where local capacity is lacking, because of the need to engage both international and local partners to undertake different roles and functions.

Gender: Gendered challenges to conducting impact evaluation and research were not mentioned in any of the interviews, in part because this topic was not sufficiently incorporated in the interview guides. This is an issue that needs to be examined in more detail in relation to both researchers and participants. For example, female researchers may be at greater risk of gender-based violence when conducting research in fragile and humanitarian contexts, or their ability to conduct the work may be questioned when working in male-dominated institutions. On the other hand, their gender may prove an advantage when researching sensitive topics with women. The gender-based challenges experienced by participants are likely to be even greater, with the risk of gender-based violence increasing conflict settings (Stark & Ager, 2011). Ensuring that impact evaluations and empirical research are gender-sensitive in design and implementation is an integral aspect of ethical research in fragile and humanitarian settings and one that deserves further investigation.

7.2 Opportunities

The above analysis reveals no huge gulf in the perspectives of researchers and research commissioners and funders. Both groups identified opportunities for future research. Many researchers stated that donors are willing to support and fund research in fragile and humanitarian settings. There is even a sense that there are opportunities in these settings that may not exist in other, more stable development settings, because funding is flowing in and interventions are being scaled up. Research commissioners and funders revealed that they are generally open to and supportive of impact evaluation and research in these settings. Overall, a picture emerged that neither funding nor donor willingness is an obstacle to increasing the quality and quantity of impact evaluations and empirical research. Rather, the challenges relate to a lack of foresight that precludes the integration of evaluation and research into programme design at the outset; resolving this short-sighted approach will entail convincing implementation partners who may have a low appetite for such efforts.

Despite the ethical challenges mentioned, it was felt that these should not deter rigorous research – at least, no more than they should in more stable contexts. Moreover, it is an ethical imperative that the voices of people living in fragile and humanitarian settings are heard, and that the effectiveness of interventions is evaluated. Many of the ethical challenges mentioned are also present in more stable development contexts. Therefore, rigorous impact evaluation and other rigorous research should also

be possible in fragile and humanitarian settings, as several existing studies demonstrate (for example, see Bailey, 2013), provided that certain enabling factors are present to increase coordination between researchers, commissioners and funders, and implementing partners.

Finally, there is an opportunity and a need for implementation and operational research on *how* to implement programmes in fragile and humanitarian contexts.

7.3 Innovations

On promising and innovative approaches to evaluation and research in fragile and humanitarian settings, many responses included a note of scepticism or critiques of technological solutions. Several respondents proposed innovations related to the incorporation of different types of data, arguing, for example, that administrative data are underused and that technologies such as GIS, GPS and mobile phones can help in certain instances. Technology is, however, neither a panacea nor a replacement for appropriate and credible research designs and methodologies, and direct data collection on the ground. In fragile and humanitarian settings, being flexible and adaptable may entail the use of the least risky and least time-consuming approach. Furthermore, alternative approaches to traditional RCTs, including quasi-experimental and qualitative designs, may be more feasible in certain circumstances.

7.4 Recommendations

Based upon a reading of the analysis conducted here, several recommendations can be made for increasing the quality and quantity of impact evaluations and other rigorous empirical research in fragile and humanitarian settings.

A unified approach
There are several areas of agreement between researchers and research commissioners and funders, and their perspectives on conducting impact evaluations and rigorous research in fragile and humanitarian contexts do not differ widely. Furthermore, there is a willingness among commissioners and funders to support this type of work. This presents an opportunity for a more unified approach between researchers, commissioners and funders, and implementing partners.

Interviewees offered no specific recommendations to demonstrate what a unified approach might look like in practice, but Peterman et al. (2018) comment on some of the enabling factors, which seem pertinent here. These relate to increasing coordination between researchers, programmers and funders, so that evidence generation strategies align with funding cycles and feed into programming. The realization of these enablers could go some way to ensuring that incentives are present for researchers and that the demand for evidence among commissioners and funders is met (Peterman et al., 2018). Enabling factors for a unified approach include the following:

- "An adequate space to present and discuss evidence" (p. 1), which can allow for evidence generation to coordinate with programming and for evidence to feed into programmatic decisions.

- Development of coherent evidence generation strategies to improve coordination within and between implementing partners and commissioners and funders.

- "Protected research funding … when programmatic funds dwindle" (p. 1).

Increasing appetite for evaluation and research among implementing partners

Another issue highlighted in the analysis is the fact that local implementing partners face huge challenges in their work, so all efforts must be made to ensure that evaluation and research helps rather than hinders implementation. Strategies for increasing appetite for evaluation and research among implementing partners are highlighted in the analysis. Implementing partners need to be engaged and consulted at all stages of the research cycle to ensure that evaluation and research are demand-driven. Results need to be provided in a timely fashion: If findings reveal that the programme is not achieving the expected impacts, learning this early on gives implementing partners time to deal with issues.

Increasing appetite for evaluation and research among researchers

Also highlighted in the analysis is the fact that researchers who take on impact evaluation and research projects in challenging contexts must make a significant investment for an uncertain pay-off. Funders can facilitate such investments by aligning funding sources and guaranteeing the availability of funding to see research through to completion. An ideal approach suggested from the researcher's perspective is to have an open challenge fund with a rapid review process and rolling submissions. Implementing organizations can further increase the incentives for evaluation and research in complex settings by sharing internal data sources and allowing findings to be openly shared so that they may feed into public debates.

Enabling researchers to work in challenging contexts

As mentioned, it may be challenging for academic researchers and other research consultants to work in humanitarian contexts. A researcher's institution or employer may be unable to provide appropriate security clearance and insurance and therefore forbid visits to certain geographical locations, even if the individual is willing and able to travel. Implementing organizations may consider enhancing researchers' ability to travel by offering coverage under their risk management and security procedures. These procedures may include, for instance, security training, security briefings and travel safety measures.

Capacity strengthening of local research organizations

Finally, while the understanding of and capacity for evaluation and research among local organizations in fragile and humanitarian settings is increasing – in line with the increased quantity of evaluations and research being undertaken in such contexts – more needs to be done to ensure that local researchers have the necessary skills to undertake rigorous impact evaluation and other rigorous research. Building partnerships with local research organizations and training local researchers are both ethical and technical imperatives. Increasing local capacity may help to reduce the research management challenges facing commissioners and funders. 'It may also help to avoid wasting time and effort in trying to source international researchers, who may be unwilling to take the risk of working in challenging and insecure environments.'

REFERENCES

Ager, A., Burnham, G., Checchi, F., Gayer, M., Grais, R. F., Henkens, M., Massaquoi, M. B. F., Nandy, R., Navarro-Colorado, C., & Spiegel, P. (2014). Strengthening the evidence base for health programming in humanitarian crises. *Science*, *345*(6202), 1290–1292.

Aker, J. C. (2017). Comparing cash and voucher transfers in a humanitarian context: Evidence from the Democratic Republic of Congo. *World Bank Economic Review*, *31*(1), 44–70.

Audsley, B., R. Halme and N. Balzer (2010) Comparing cash and food transfers: a cost-benefit analysis from Rural Malawi. In Omamo, S.W., U. Gentilini and S. Sandström (eds.) Revolution: From Food Aid to Food Assistance- Innovations in Overcoming Hunger. World Food Programme, Rome.

Aurino, E., Brück, T., Daidone, S., Natali, L., & Schwab, B. (2018, June 21). *Administrative data: Missed opportunity for learning and research in humanitarian emergencies?* (Evidence for Action blog). United Nations Children's Fund. https://blogs.unicef.org/evidence-for-action/administrative-data-missed-opportunity-for-learning-and-research-in-humanitarian-emergencies

Bahati, I. (2019, June 24). *The challenges facing female researchers in conflict settings* (From Poverty to Power blog). Oxfam GB. https://oxfamblogs.org/fp2p/the-challenges-facing-female-researchers-in-conflict-settings

Bailey, S. (2013). *The impact of cash transfers on food consumption in humanitarian settings: A review of evidence*. Canadian Foodgrains Bank, Winnipeg. Available at: https://reliefweb.int/sites/reliefweb.int/files/resources/cfgb---impact-of-cash-transfers-on-food-consumption-may-2013-final-clean.pdf

Battistin, F. (2016). *Impact evaluation of the multipurpose cash assistance programme*. Lebanon Cash Consortium, Beirut. https://resourcecentre.savethechildren.net/node/12676/pdf/lccimpactevaluationformcafebruary2016final.pdf

Berman, G., Hart, J., O'Mathúna, D., Mattellone, E., Potts, A., O'Kane, C., Shusterman, J., & Tanner, T. (2016). *What we know about ethical research involving children in humanitarian settings: An overview of principles, the literature and case studies* (Innocenti Working Paper 2016-18). United Nations Children's Fund Office of Research – Innocenti, Florence.

Blanchet, K. & Roberts, B. (2015). *An evidence review of research on health interventions in humanitarian crises.* London School of Hygiene and Tropical Medicine, London.

Blanchet, K., Ramesh, A., Frison, S., Warren, E., Hossain, M., Smith, J., Knight, A., Post, N., Lewis, C., Woodward, A., Dahab, M., Ruby, A., Sistenich, V., Pantuliano, S., & Roberts, B. (2017). Evidence on public health interventions in humanitarian crises. *Lancet*, *390*(10109), 2287–2296.

Blanchet, K., Allen, C., Breckon, J., Davies, P., Duclos, D., Jansen, J., Mthiyane, H., & Clarke, M. (2018). *Using research evidence in the humanitarian sector: A practice guide*. Evidence Aid, London School of Hygiene and Tropical Medicine, and Nesta (Alliance for Useful Evidence), London.

Burde, D. (2012). Assessing Impact and Bridging Methodological Divides: Randomized Trials in Countries Affected by Conflict. *Comparative Education Review*, 56(3), 448–473.

Buttenheim, A. (2010). Impact evaluation in the post-disaster setting: A case study of the 2005 Pakistan earthquake. *Journal of Development Effectiveness*, *2*(2), 197–227.

Cherrier, C. (2019). *Learning from and for operations: Outcomes from a workshop on evidence on social protection in contexts of fragility and forced displacement*. European Commission Guidance Package on Social Protection across the Humanitarian–Development Nexus (SPaN). https://europa.eu/capacity4dev/sp-nexus/documents/gp-think-piece-number-1-learning-and-operations-unicef-innocenti-2018

de Hoop, J., Morey, M., Ring, H., Rothbard, V., & Seidenfeld, D. (2018). *"Min Ila" cash transfer programme for displaced Syrian children in Lebanon (UNICEF and WFP). Impact evaluation endline report*. American Institutes for Research, Washington, D. C.

Di John, J. (2008). *Conceptualising the causes and consequences of failed states: A critical review of the literature*. Crisis States Research Centre, London School of Economics and Political Science, London.

Doocy, S., & Tappis, H. (2016). *Cash-based approaches in humanitarian emergencies: A systematic review* (Systematic Review 28). International Initiative for Impact Evaluation (3ie), London.

Gaarder, M., & Annan, J. (2013). *Impact evaluation of conflict prevention and peacebuilding interventions* (Policy Research Working Paper 6496). World Bank, Washington, D. C..

Gilligan, D. O., Margolies, A., Quiñones, E., & Roy, S. (2013). *Impact evaluation of cash and food transfers at early childhood development centers in Karamoja, Uganda. Final impact report*. International Food Policy Research Institute, Washington, D. C.. http://www.cashlearning.org/downloads/impact-evaluation-of-cash-and-food-transfers-at-early-childhood-dev.pdf

Hidrobo, M., Hoddinott, J., Peterman, A., Margolies, A., & Moreira, V. (2014). Cash, food, or vouchers? Evidence from a randomized experiment in northern Ecuador. *Journal of Development Economics*, *107*, 144–156.

Hoddinott, J., Sandström, S., & Upton, J. (2018). The impact of cash and food transfers: Evidence from a randomized intervention in Niger. *American Journal of Agricultural Economics*, *100*(4), 1032–1049.

Huggins, M. K., & Glebbeek, M.-L. (2003). Women studying violent male institutions: Cross-gendered dynamics in police research on secrecy and danger. *Theoretical Criminology*, *7*(3), 363–387.

Idris, I. (2019). *Doing research in fragile contexts: Literature review* (GSDRC Research Report no. 1444). Governance and Social Development Resource Centre, University of Birmingham, Birmingham, UK.

Inter-Agency Standing Committee. (2015). *Introduction to humanitarian action: A brief guide for resident coordinators*. IASC, Geneva.

Lehmann, C., & Masterson, D. (2014). *An impact evaluation of the 2013–2014 winter cash assistance program for Syrian refugees in Lebanon*. International Rescue Committee, Beirut. https://www.rescue.org/sites/default/files/document/631/emergencyeconomiesevaluationreport-lebanon2014.pdf

Maxwell, D., Majid, N., Adan, G., Abdirahman, K., & Kim, J. J. (2016). Facing famine: Somali experiences in the famine of 2011. *Food Policy*, *65*, 63–73.

Mishra, A. (2017). *Child outcomes of cash transfer programming. What works and what doesn't for children in humanitarian and development contexts* (Research brief). Save the Children UK, London.

Organisation for Economic Co-operation and Development. (2006). Applying Strategic Environmental Assessment: Good Practice Guidance for Development Co-operation (DAC Guidelines and Reference Series). OECD Publishing, Paris.

OECD (2012). *Evaluating peacebuilding activities in settings of conflict and fragility: Improving learning for results* (DAC Guidelines and References series). OECD Publishing, Paris.

OECD (2016). *States of Fragility 2016: Understanding Violence*. Paris: OECD Publishing, Paris.

Peterman, A., de Hoop, J., Cuesta, J., & Yuster, A. (2018). *Evidence on social protection in contexts of fragility and forced displacement* (Innocenti Research Brief 2018-22). United Nations Children's Fund Office of Research – Innocenti, Florence.

Presser, L. (2005). Negotiating power and narrative in research: Implications for feminist methodology. *Signs Journal of Women in Culture and Society*, *30*(4), 2067–2090.

Puri, J., Aladysheva, A., Iversen, V., Ghorpade, Y., & Brück, T. (2017). Can rigorous impact evaluations improve humanitarian assistance? *Journal of Development Effectiveness*, *9*(4), 519–542.

Puri, J., Aladysheva, A., Iversen, V., Ghorpade, Y., & Brück, T. (2015). *What methods may be used in impact evaluations of humanitarian assistance?* (IZA Discussion Paper No. 8755). Institute for the Study of Labor, Bonn.

Sandström, S., & Tchatchua, L. (2010). Do cash transfers improve food security in emergencies? Evidence from Sri Lanka. In S. W. Omamo, U. Gentilini, & S. Sandström (Eds.), *Revolution: From food aid to food assistance – Innovations in overcoming hunger* (pp. 75–88). World Food Programme, Rome.

Schwab, B., Margolies, A., & Hoddinott, J. (2013). *Impact evaluation of cash and food transfers for the seasonal emergency safety net in Hajjah and Ibb governorates. Yemen endline report*. International Food Policy Research Institute, Washington, D. C. http://www.cashlearning.org/downloads/wfp257674.pdf

Sharma, M. (2006). *An assessment of the effects of the cash transfer pilot project on household consumption patterns in tsunami-affected areas of Sri Lanka*. Report submitted to the World Food Programme, Rome, by the International Food Policy Research Institute, Washington, D. C.

Stark, L., & Ager, A. (2011). A systematic review of prevalence studies of gender-based violence in complex emergencies. *Trauma, Violence, & Abuse*, *12*(3), 127–134. https://doi.org/10.1177/1524838011404252

Tranchant, J.-P., Gelli, A., Bliznashka, L., Diallo, A. S., Sacko, M., Assima, A., Siegel, E. H., Aurino, E., & Masset, E. (2018). The impact of food assistance on food insecure populations during conflict: Evidence from a quasi-experiment in Mali. *World Development*, *119*, 185–202.

United Nations Children's Fund Office of Research – Innocenti. (2014). *Taxonomy for defining and classifying UNICEF research, evaluation and studies*. UNICEF Office of Research – Innocenti, Florence.

Verwimp, P., Justino, P., & Brück, T. (2019). The microeconomics of violent conflict. *Journal of Development Economics*, *141*, 102297. https://doi.org/10.1016/j.jdeveco.2018.10.005

APPENDIX A: BIOGRAPHIES OF INTERVIEWEES

Each biography provides the individual's affiliation at the time of the interview.

RESEARCHERS

Jenny Aker is a professor at Tufts University. Her work partly focuses on the impact of information technology on development outcomes, including in humanitarian settings in the Democratic Republic of the Congo and the Niger. Prior to completing her PhD, she worked for Catholic Relief Services in West and Central Africa.

Jeannie Annan is Senior Director of Research and Evaluation for the International Rescue Committee (IRC). Her research focuses on sexual and gender-based violence against women, children and youth in armed conflict, and psychosocial programmes for survivors of wartime and sexual violence. Her work aims to improve humanitarian policy and programmes through rigorous research, including evaluation of programmes that prevent and respond to violence. She has undertaken evaluation and research in several countries, including Burundi, the Democratic Republic of the Congo, Liberia and Uganda.

Tilman Brück is Team Leader for Economic Development and Food Security at Leibniz Institute of Vegetable and Ornamental Crops (IGZ) and the Founder and Director of the International Security and Development Center (ISDC). His research interests focus on the economics of household behaviour and well-being in conflict-affected and fragile economies, including measurement of violence and conflict using household surveys and impact evaluation of programmes in conflict-affected areas. He has been involved in the collection of panel data in fragile and conflict-affected settings, including in Angola, Kenya, Kyrgyzstan, the Niger, Nigeria, the Syrian Arab Republic, Uganda and Ukraine. His work comprised impact analysis in a subset of these countries.

Dana Burde is an associate professor at the Steinhardt School of Culture, Education, and Human Development at New York University. She has extensive experience of working in Afghanistan and Pakistan, among other countries. Before becoming a researcher, she was an aid worker, an experience that she indicates has made her sensitive to the needs of practitioners and their perceptions of researchers.

Fotini Christia is a professor of political science at the Massachusetts Institute of Technology (MIT). Her research interests deal with issues of conflict and cooperation in the Muslim world, and she has done extensive experimental, survey and ethnographic fieldwork on ethnicity, conflict and development in Afghanistan, Bosnia and Herzegovina, Iraq, the State of Palestine and Yemen.

John F. Hoddinott is the H. E. Babcock Professor of Food and Nutrition Economics and Policy at Cornell University. His research focuses on the causes and consequences of poverty, hunger and undernutrition in developing countries. John has ongoing research work in three countries: Bangladesh, Ethiopia and Guatemala. The Bangladesh and Ethiopia studies are a mix of prospective cohort and randomized controlled trials, focusing on agriculture, social protection, food security and nutrition. The Guatemala study is a follow-up to a randomized community nutrition intervention that was fielded in the early 1970s.

Daniel Masterson is a postdoctoral fellow at the Immigration Policy Lab at Stanford University. He

works in comparative politics, studying the political economy of migration and refugee crises, with a regional focus on the Middle East. He uses survey research, field experiments, natural experiments and qualitative fieldwork, working primarily in the Levant. He previously worked for the United Nations High Commissioner for Refugees (UNHCR), the World Bank, and the International Rescue Committee (IRC) in the Middle East.

Daniel Maxwell is the Henry J. Leir Professor in Food Security and Research Director at the Feinstein International Center, Friedman School of Nutrition Science and Policy, Tufts University. He works with governments, humanitarian agencies and affected communities at the grassroots, national and regional level to build the evidence base for improved humanitarian and resilience policy and programming.

Amber Peterman is a consultant to the Social and Economic Policy Unit at the UNICEF Office of Research – Innocenti and an associate adjunct professor at the University of North Carolina at Chapel Hill. She focuses on gender, violence and adolescent well-being and safe transitions to adulthood, including in her work with the Transfer Project evaluations of social protection and cash transfers in Africa. She has led research in more than a dozen countries and has significant experience in conducting large-scale surveys and impact evaluation in Africa, Latin America, the Middle East and Asia.

Jyotsna Puri is Head of the Independent Evaluation Unit of the Green Climate Fund. She has more than 20 years of experience in the evaluation of poverty alleviation policies related to agriculture, environment, health and infrastructure. She is also a board member of the international boards of the Humanitarian Quality Assurance Initiative (HQAI) and the International Centre for Evaluation and Development (ICED).

RESEARCH COMMISSIONERS AND FUNDERS

Ulrich Assankpon is a Regional Resilience Advisor for the Sahel and West Africa at the Food and Agriculture Organization of the United Nations (FAO). He provides programmatic and operational support in the development and implementation of the FAO Resilience Programme in Africa.

Prisca Benelli is the Humanitarian Research and Learning Manager at Save the Children UK. Rather than carrying out research herself as a principal investigator, her role is to support colleagues to conduct research. Her research interests include gender in humanitarian settings; forced migration; evidence-based humanitarian action; humanitarian action in urban settings; sampling rare, elusive and hard-to-reach populations; and the use of technology for humanitarian assistance.

Marie Gaarder is Director of the Evaluation Office and Global Director for Innovation and Country Engagement at the International Initiative for Impact Evaluation (3ie). She provides general leadership, strategic direction and guidance to 3ie work in evaluation, innovation and country engagement, in addition to overseeing the Evaluation Office. She is also responsible for initiating and managing research activities and producing papers that highlight 3ie's experience with impact evaluation.

Marie-France Guimond is a Research Systems Advisor at the International Rescue Committee (IRC). She has over 10 years of experience in evaluation, research and programme monitoring as well as programme management in humanitarian and conflict-affected contexts across three continents. As part of her role, she has developed IRC research standards, toolkits and management processes. Göran Holmqvist is Director of Asia, Middle East and Humanitarian Assistance at the Swedish

International Development Cooperation Agency (Sida).

Tara Kaul is an Evaluation Specialist at the International Initiative for Impact Evaluation (3ie), where she reviews research and manages 3ie-funded impact evaluation grants. Her current research focuses on the nutritional impact of food subsidies, and intra-household gender discrimination in educational expenditures in India.

Massimo La Rosa is the Global Thematic Adviser on Migration, Forced Displacement and Social Protection for the European Commission Humanitarian Aid Office (ECHO).

Claire Mariani is the Humanitarian Cash Transfers Coordinator at UNICEF. She has guided and contracted out multiple impact evaluations of cash transfer programmes in humanitarian settings.

Urs Nagel is the Regional Evaluation Advisor for the UNICEF Eastern and Southern Africa Regional Office. He was previously Chief of Social Policy at UNICEF Afghanistan and has been involved in impact evaluations in Afghanistan, India and Pakistan.

Jonathan Patrick is an Evaluation Adviser with the UK Department for International Development's Evaluation Unit. He has experience of working on humanitarian evaluations, including on the International Initiative for Impact Evaluation (3ie) Humanitarian Interventions Thematic Window.

APPENDIX B: INTERVIEW GUIDES

For researchers

1. Can you reflect on the difference between carrying out research involving data collection in fragile and non-fragile contexts?

Probing questions:

- What are the main challenges when working in fragile settings as compared to other settings?

- How do the two settings compare in terms of ethical challenges?

- Do data collectors and researchers working in fragile contexts need different skills?

- Can you reflect on raising resources for data collection and analysis in fragile contexts? Are donors more or less willing to finance this type of research effort? Why?

- What are the main challenges for data collection? (Think of security, gaining trust of interviewees, quality controls, partnerships with local institutions, etc.)

- How can we ensure that research in fragile and non-fragile contexts is relevant for practitioners and does not hamper programme implementation?

2. Can you tell us about your research agenda related to fragile settings? What, in your view, are some of the big questions related to social protection in fragile settings that you think researchers should focus on in the coming years?

3. Are there any innovative approaches (e.g., technology, big data, merging with administrative data) that you consider to be particularly promising for programme evaluations in fragile contexts?

Probing question:

- Have you employed any of these approaches and can you tell us about your experience?

For research commissioners and funders

1. Can you reflect on your experience of guiding and contracting out impact evaluations (using a counterfactual) which have examined optimal policy responses in settings of humanitarian emergency?

Probing questions:

- What are the main challenges you have encountered and how do they differ to evaluations in non-humanitarian settings?

- What are the main challenges related to data collection? (Think of security, gaining trust of interviewees, quality controls, partnerships with local and international institutions, etc.)

- What are the main ethical challenges?

- Do researchers have the right skills to guide data collection in humanitarian contexts?

2. Can you reflect on raising resources for data collection and analysis in fragile contexts? Are donors more or less willing to finance this type of research effort? Why?

Probing questions:

■   If the response is negative: Is there a mismatch between the interests of researchers and the questions practitioners seek to answer?

■   If it is positive: Please elaborate – how did you use the findings of the research?

3. Have you been able to use this type of research in programmatic decisions and policy discussions?

4. Can you tell us what you think are the key questions for researchers to focus on in fragile settings? In particular, what are some of the big questions related to social protection in fragile settings you would like to see answered?

APPENDIX C: QUALITATIVE CODING THEMES AND SUB-THEMES

For researchers

Themes (codes)
1 Difference between data collection in fragile and non-fragile contexts
1.1 Challenges
Access
Data scarcity
Incentives
Local research capacity
Multiplicity of actors
A valid counterfactual
Safety and security
Staff turnover
Time constraints
Transient populations
Unpredictability
1.2 Design and data collection challenges
Data collection and quality oversight
Instruments
Sampling
1.3 Ethics
Benefits perception
Do no harm
Informed consent
Institutional review board standards
Local actors
Randomized controlled trial ethics
1.4 Researcher skills
Adaptable, willingness to incur risk
Understanding of local context
Social and emotional awareness
Trusted by local partners
Ability to work remotely

Themes (codes)
2 Raising resources in fragile contexts
2.1 Donors are supportive, flexible
2.2 Flexible funding
2.3 Funds limited and lower priority
2.4 Limited integration of impact evaluation in design
2.5 Research independence
3 Research relevant for practitioners
3.1 Engagement with implementers and partners
3.2 Input into research question
3.3 Need for implementation research
3.4 Provides timely results
3.5 Respect for partners
4 Research agenda in fragile contexts
4.1 Adolescent well-being (multi-sectoral)
4.2 Disability
4.3 Education
4.4 Elites
4.5 Employment
4.6 Fragility
4.7 Infant mortality
4.8 Mental health
4.9 Migration and displacement
4.10 Refugee communities
4.11 Remittances
4.12 Self-reliance, local networks
4.13 Social protection
4.14 Social protection: 'Cash plus'
4.15 Social protection: Cash transfers
4.16 Social protection: Cash vs in-kind benefits
4.17 Social protection: Cost-effectiveness
4.18 Social protection: Implementation research
4.19 Social protection: Long-term impacts
4.20 Social protection: Market analysis
4.21 Unit of analysis
4.22 Violence

Themes (codes)
5 Innovative approaches
5.1 Big data
5.2 Critique of technology
5.3 Diaspora as entry point
5.4 Geographic information system (GIS) and Global Positioning System (GPS)
5.5 Hotlines
5.6 Machine learning
5.7 Metadata
5.8 Mobile phone use
5.9 Non-technological innovations around data
Administrative data
Contextual data
'Lighter' data
5.10 Alternative methodologies
5.11 Recording using tablet computers and wireless networks
5.12 Audio recordings
5.13 Staff skills and investment
5.14 Surveillance data
6 Views on research design and methods

For research commissioners and funders

Themes (codes)
1 Experience of impact evaluation in humanitarian settings
1.1 Adaptable, flexible
1.2 Appropriate data collection approaches
1.3 Confidence in partners
1.4 Context and access
1.5 Design and attribution
1.6 Ensuring impact evaluation fits programme design and context
1.7 Ethics control groups
1.8 Ethics overloading participants
1.9 Impact evaluation of responses to rapid events is challenging
1.10 Importance of embedded evaluations

Themes (codes)
1.11 Innovative methodologies
1.12 Partnership
1.13 Research gaps
1.14 Researcher skills
1.15 Security and risk
1.16 Time constraints
1.17 Transient populations
2 Raising resources for data collection in fragile settings
2.1 Different donor priorities
2.2 Donors supportive
2.3 Evaluation is not a priority
2.4 Increased funding necessary
3 Ability to use research
3.1 If NO, why not?
Context-specific
Short cycles
Too few impact evaluations
3.2 If YES, how were findings used?
Other factors influence decisions, such as capacity and political will
To develop toolkits and guidance
To inform policy and programming
To support funding decisions
More frequent use of synthesis, not primary research
4 Key research questions in fragile and humanitarian settings
4.1 Aligning with government
4.2 Cash for work
4.3 Cash multi-sectoral impacts
4.4 Cash plus
4.5 Cash vs in-kind benefits
4.6 Combating myths
4.7 Context not thematic
4.8 Cost-effectiveness
4.9 Education
4.10 Enhancing resilience

Themes (codes)
4.11 How social protection can be adapted to shocks
4.12 Implementation and service delivery
4.13 Malnutrition
4.14 Elderly people
4.15 Phasing out
4.16 Political and conflict analysis
4.17 Protection
4.18 Social protection and agriculture
4.19 Task shifting
4.20 Unintended effects
4.21 Within-household disaggregation
4.22 Humanitarian–development divide
4.23 Addressing myths
4.24 Holistic approach
4.25 Vulnerable populations
4.26 Adaptation of social protection system
4.27 Access to water

for every child, answers

unicef
Office of Research – Innocenti