



Evaluating protection in humanitarian action: Issues and challenges

Francesca Bonino

ALNAP is a unique system-wide network dedicated to improving the performance of humanitarian action through shared learning. www.alnap.org

An electronic copy of the study and other related resources are available on the ALNAP website at www.alnap.org/evaluating-protection.

About the author:

Francesca Bonino is a Research Fellow with the ALNAP Secretariat based in London.

Suggested citation

Bonino, F. (2014) *Evaluating protection in humanitarian action: Issues and challenges*. ALNAP Working Paper. London: ALNAP/ODI.

© ALNAP/ODI 2014. This work is licensed under a Creative Commons Attribution-Non Commercial Licence (CC BY-NC 3.0).

ISBN 978-1-910454-14-5

Publication and communications managed by Maria Gili

Edited by Amanda Morgan

Copy edited by Alex Potter

Design by Maria Gili and Deborah Durojaiye

Acknowledgements

The author would like to thank Caitlin Procter for her assistance with research during the early stages of work on this paper, and Paul Knox Clarke, John Mitchell, and Luz Saavedra from the ALNAP Secretariat for their support and comments. Many thanks go also to the interviewees who so kindly gave some of their time to discuss protection and evaluation of protection issues, offered their reflections on the subject, and suggested additional contacts and literature to review.

This paper greatly benefitted from the comments and inputs received from external reviewers who offered thorough and insightful comments on an early draft of this paper. Many thanks to Louise Aubin (UNHCR/GPC), Rachel Hastie (Oxfam GB), Jenny McAvoy (InterAction), Velina Mikova (ICRC), Kathrine Starup (Danish Refugee Council) and Arne Treves (UNHCR). Their input and suggestions have been extremely useful in helping focus and sharpen the content of the paper. Any responsibility for shortcomings, factual errors or misinterpretation rests solely with the author.

Contents

Abbreviations and acronyms	5
1. Introduction	6
1.1 About the paper	6
1.2 The gaps	8
1.3 The aim of this initiative	10
2. Four common issues that complicate the evaluation of protection	11
2.1 Defining protection in humanitarian action	12
2.2 Identifying results of interest, defining 'success' and framing the evaluation	16
2.3 Collecting and analysing data	24
2.4 Understanding cause-and-effect issues	30
3. Next steps	35
Annexes	38
Annex A. The concept of protection in humanitarian evaluation guidance	39
Annex B. Methodology note	44
References	50
Further readings	61

Boxes and figures

Figure 1. Using the BetterEvaluation framework to guide the structure of the paper	7
Figure 2. The egg protection framework	14
Box 1. The global protection architecture and its evolution	17
Box 2. Where does protection mainstreaming start?	21

Abbreviations and acronyms

ALNAP	Active Learning Network for Accountability and Performance in Humanitarian Action
DG ECHO	European Commission's Directorate-General for Humanitarian Aid
EHA	Evaluation of humanitarian action
GPC	Global Protection Cluster
ICRC	International Committee of the Red Cross
IRC	International Rescue Committee
NGO	Non-governmental organisation
OECD-DAC	Organisation for Economic Co-operation and Development's Development Assistance Committee
SGBV	Sexual and gender-based violence
UN	United Nations
UNHCR	United Nations High Commissioner for Refugees
UNICEF	United Nations Children's Fund

1. Introduction

The importance of evaluating protection results mirrors that of evaluating the results of other aspects of humanitarian action. Evaluation can generate evidence to support decision-making, substantiate accountability claims and support learning (UNICEF, 2013a: 2). But there is a perception among both programme staff and evaluators that protection is harder to evaluate than other aspects of humanitarian action, that it is less tangible than other areas and that the quality of evidence from existing evaluations of protection is less than satisfactory (Reichhold et al., 2013). This paper identifies and describes the challenges affecting the evaluation of protection in humanitarian action with a view to suggesting options to start addressing them.

1.1 About the paper

This paper attempts to identify the issues and challenges relating to the evaluation of protection work carried out by humanitarian actors, including those both with and without a specific protection mandate. It excludes literature and practice on the Responsibility to Protect doctrine and work carried out specifically in the context of human rights, legal or prosecutorial actions (including the work of the International Criminal Court), and security and military actions (such as peacekeeping missions) with protection components.

The paper is written primarily for staff in evaluation, protection, and programme-advising roles whose effectiveness in commissioning and using evaluations is essential to improving the quality and use of evaluative evidence. Such staff include evaluators; members of operational agencies and donor offices who commission, lead and support evaluations; and evaluation researchers and consultants. The paper is also relevant to staff working in protection programming, advisory, and policy-making positions in both headquarters and field offices, who are often called on to comment on evaluation terms of reference and inception reports and to host evaluation missions.

“

There is a perception among both programme staff and evaluators that protection is harder to evaluate than other aspects of humanitarian action. This paper identifies and describes the challenges affecting the evaluation of protection with a view to suggesting options to start addressing them.

”

The BetterEvaluation framework (Better Evaluation, 2014) given in Figure 1 was used to organise and structure the contents of the paper and make it accessible to a wide and diverse audience. This framework was not developed specifically from a humanitarian evaluation perspective.¹ Nevertheless, it is used here because it offers a comprehensive and flexible way of organising vast amounts of information relating to practical evaluation issues by grouping them into a series of broad process steps that are common to virtually every type of evaluation exercise in most contexts – including those where humanitarian actors operate. These broad process steps are (1) manage an evaluation or evaluation system; (2) *define what is to be evaluated*; (3) *identify the results of interest for the evaluation and framing the evaluation*; (4) *collect and analyse data (about context, activities and results)*; (5) *understand the causes of results*; (6) synthesise data from one or more evaluation; and (7) report and support the use of evaluation findings.

The framework can be used in various ways (see BetterEvaluation, 2014b) and the paper will make partial use of it by focusing on four of the seven process steps (italicised in the list given above). This is because the paper aims to single out the specific challenges that impact on the evaluation of protection, in addition to what is generally found to affect humanitarian evaluation practice.

Figure 1. Using the BetterEvaluation framework to guide the structure of the paper



Source: BetterEvaluation (2014:1).

¹ This paper uses the definition of humanitarian evaluation presented in the ALNAP guide on evaluating humanitarian action: 'the systematic and objective examination of humanitarian action, intended to draw lessons to improve policy and practice and enhance accountability. ... Systematic implies a planned and consistent approach. Objective implies stepping back from the immediacy of the humanitarian action and getting some perspective on it. Examination implies exploration or analysis to determine the worth or significance of the action' (Cosgrave and Buchanan-Smith, 2013: 15).

The remaining three categories in the BetterEvaluation framework (managing an evaluation, synthesising evaluative evidence and reporting and supporting the use of findings) are not discussed in this paper as the background research did not point towards protection-specific evaluation challenges regarding these.

For consistency, the BetterEvaluation framework will also be used to guide the development of the next publication in this ALNAP series looking at the evaluation of protection, where we will move from identifying issues and challenges in evaluation to proposing approaches and options to address them.

Section 2 of the paper reviews key issues and challenges affecting the evaluation of protection in humanitarian action, in line with some of the categories suggested by the BetterEvaluation framework. **Section 3** lays out a plan for further research, and for the development and piloting of evaluation guidance to begin to address the challenges identified in Section 2. Additional information is contained in two annexes: **Annex A** gives an overview of how humanitarian inter-agency evaluation guidance has dealt with protection thus far, while **Annex B** explains the methodology used to develop this paper.

1.2 The gaps

Despite the urgency attached to the topic (including in the light of ongoing crises in the Central African Republic, Iraq and Syria), the evaluation of protection in humanitarian action is lagging behind other areas of inquiry in the evaluation of humanitarian action (EHA). The depth and extent of evaluative analysis appear to be unsatisfactory, as was recently discussed, for example, by Reichhold et al. (2013: 16), who highlight the scarcity of humanitarian evaluation work attempting to make the shift from outputs and the activity/process level to the outcomes and impacts level of analysis.

Protection issues do not appear to be adequately covered in EHA-specific guidance. Annex A contains a brief review of inter-agency humanitarian evaluation guidance which indicates that references to protection are somewhat scanty and, where present, suggest rather different – and often inconsistent – ways to treat protection in humanitarian evaluations.

“

The literature reviewed revealed some gaps that have also been noticed by many ALNAP Members, including the lack of agreement on how to conceptualise and frame protection in humanitarian evaluation and the lack of a taxonomy for it.

”

For example, in some evaluation guidance (e.g. IASC, 2011a) protection is treated as a cross-cutting issue; in others (e.g. Hallam, 1998) as an overarching theme that all humanitarian evaluators should consider to some extent in their EHA work. In other guidance examples (e.g. IASC-IAHE Group, 2014a) protection is referred to as one of the possible evaluation criteria, together with the other more frequently used criteria of relevance, effectiveness, efficiency and so on of the Organisation for Economic Co-operation and Development's Development Assistance Committee (OECD-DAC) (OECD-DAC, 2010).

Overall, the humanitarian evaluation literature reviewed revealed some gaps that have also been noticed by many ALNAP Members, including (1) the lack of agreement on how to conceptualise and frame protection in humanitarian evaluation – which to some extent reflects the complexity of and challenges in protection programming (discussed in Section 2); and (2) the lack of a taxonomy for evaluators that spells out how different understandings of protection (programming and results) may call for different treatments in evaluations. The present ALNAP initiative aims to contribute to addressing these two gaps.

In addition to reviewing evaluation-specific literature, we examined the literature and guidance that deals with protection from a broader programming angle to see whether and how they discuss evaluation.

Here, the humanitarian literature concentrates on several protection thematic and context-specific areas such as: protection in situations of natural disasters; protection in situations of armed conflict; protection from sexual exploitation and abuse and from sexual and gender-based violence (SGBV); the protection of children; the protection of refugees; protection in situations of displacement; protection from landmines and explosive remnants of war; the protection of the elderly; and protection from mental and psychosocial harm in humanitarian crisis contexts.

In terms of these protection thematic areas/protection sub-themes some dedicated efforts have been made to improve overall programming and in some cases to improve evaluation by developing specific and thematic guidance, e.g. in the thematic areas of protection from mental and psychosocial harm (see Boothby et al., 2006; Ager et al. 2011) and child protection (see de las Kropiwnicki, 2012; Save the Children, 2012b; 2013). However, as some interviewees noted, the sector-, agency- and mandate-specific lenses through which this programming guidance is written have somewhat inhibited the work of capturing, synthesising, and validating good evaluation practices and evaluation lessons worth sharing across the protection programming and evaluation communities.

Nonetheless, overall there are signs of increased attention to the subject of the evaluation of protection in humanitarian work. For example, the International Committee of the Red Cross (ICRC) dedicated more space to evaluation in its 2013 Professional standards for protection work (ICRC, 2013: 5) than in the previous edition (ICRC, 2009). The Global Protection Cluster (GPC) has also recently commissioned a whole-of-system review of protection in humanitarian action to ascertain the extent to which the issue enjoys a central role (IASC Principles, 2013; GPC, 2014).

“

The sector-, agency- and mandate-specific lenses through which protection programming guidance has been written have somewhat inhibited the work of capturing, synthesising, and validating good evaluation practices and evaluation lessons worth sharing.

”

1.3 The aim of this initiative

This ALNAP initiative addresses the gaps outlined above in two ways: (1) by identifying the main challenges facing the evaluation of humanitarian protection and potential strategies to overcome them; and (2) by exploring and piloting some of these evaluation approaches and related tools. The first step constitutes the focus of the present paper; the second step will be covered in a subsequent paper featuring more detailed evaluation guidance. Results will ultimately be incorporated into the ALNAP guide on *Evaluation of humanitarian action*, which at the time of writing is being piloted by numerous Members of the ALNAP Network and is available in three languages (English, French and Spanish) at www.alnap.org/eha.

2. Four common issues that complicate the evaluation of protection

Below is a list of recurring questions that emerged from both the literature and the interviews conducted for this paper that in a nutshell capture the challenges that make the evaluation of protection exceptionally complex:

- As evaluators and evaluation commissioning staff, how can we develop a solid understanding of what we are attempting to achieve when we are evaluating protection in humanitarian action?
- What does successful protection results look like? How can we articulate and measure programme performance at different levels of results in terms of protection for both programming and evaluation purposes? How is this different for mainstreamed, integrated protection and dedicated/specialist protection programming?
- How should we customise data collection and analysis strategies when we deal with sensitive information?
- How can we understand cause-and-effect issues in protection in humanitarian contexts?

Apart from the question on cause-and-effect issues – a typical one for evaluators that defines their whole professional category and endeavour – we are not suggesting that these questions preoccupy evaluators only. For example, they also are of concern to staff working in programme design and management, and other operational and advisory positions.²

Before we can decide what is feasible and appropriate and how best to evaluate any undertaking – be it a project, programme, policy or other form of intervention in a given context – *we need first to understand what we are looking at.*

“

What are we attempting to achieve when we are evaluating protection in humanitarian action?

What does successful protection results look like? How should we customise data collection and analysis when we deal with sensitive information?

”

² Other initiatives, such as the one lead by InterAction on Result-Based Protection are also tackling some of those questions from a policy and programming angle. See InterAction, 2012.

Both earlier writers on protection and interviewees for this paper emphasised the need for clarity on what is being evaluated, what constitutes results in protection and what constitutes success in protection efforts (e.g. Reichhold et al., 2013; Reichhold and Binder, 2014). Moreover, virtually all interviewees agreed that clarity on the nature of protection-related concepts and applications among programme and protection staff, advisors, and evaluators is critical to improving the quality of both humanitarian programming, monitoring and evaluation.

So, what exactly are we dealing with when we evaluate protection in humanitarian action?

This question forms a point of origin of many of the challenges confronting evaluators examining protection in humanitarian action because *it deals with the WHAT of an evaluation as it relates to the WHAT of an intervention*. The BetterEvaluation framework (2014b: 4-5) suggests approaching the question by breaking it into two parts:

1. The first involves defining *what is to be evaluated* based on how the intervention (activity, programme, service, policy, etc.) is understood to work. This is discussed in section 2.1.
2. The second involves *identifying the results of interest that the evaluation will look at and framing the evaluation accordingly*. Discussed in section 2.2.

2.1 Defining protection in humanitarian action

It is useful to start by introducing a definition of protection in humanitarian action to help describe and delimit the subject we are looking at. Protection has been defined as comprising ‘all activities aimed at obtaining full respect for the rights of the individual in accordance with the letter and the spirit of the relevant bodies of law’ (IASC, 2011b: 5). This definition leads on from an earlier one proposed by the ICRC following a series of workshops in the late 1990s that brought together over 50 humanitarian and human rights organisations. (Giassi Caverzasio, 2001).

In humanitarian literature protection has also been described as ‘the other side of assistance’ (IASC, 2002: 5) and, together with assistance, as one of the two main pillars of humanitarian action (Sphere Project, 2011: 26).

“

Protection has been defined as comprising ‘all activities aimed at obtaining full respect for the rights of the individual in accordance with the letter and the spirit of the relevant bodies of law’.

Protection has also been described as ‘the other side of assistance’.

”

“

Protection defies neat labelling because it is at the same time the goal underlying the whole humanitarian response, an approach or lens on the humanitarian response, and a more narrowly-defined family of activities.

”

Murray and Landry,
2013

This language is broad and rather all-encompassing in nature – too much so, some authors would say (see Hastie et al., 2007; DuBois, 2010; Niland, 2013). An earlier ICRC analysis on this issue (Giossi Caverzasio, 2001: 18-20) suggested that such broad language was needed to accommodate and make space for different and complementary perspectives on protection and the different registers used when describing it:

For some, a protection activity [is] an activity aimed at implementing international law. This means that organizations have to move from a traditional view of victims and beneficiaries to one which views them as possessors of rights with legitimate claims under international law. For others, a protection activity is an activity in defence of human rights, or which documents human rights abuses. ... For a third category, a protection activity is any humanitarian activity, given that its ultimate goal is to protect people. According to this view, supplying food to starving people should be considered protection. In short, assistance is also protection. (Giossi Caverzasio, 2001: 20)

In a recent study commissioned for the GPC, Murray and Landry (2013) aptly capture this multifaceted nature of protection, which ‘defies neat labelling’, as they put it:

Protection defies neat labelling because it is at the same time the *goal* underlying the whole humanitarian response (the reason for humanitarian action), an *approach* or lens on the humanitarian response (a way of understanding all dimensions of humanitarian endeavour), and a more narrowly-defined family of **activities** that aim to prevent and mitigate threats to vulnerable persons. (Murray and Landry, 2013: 4; original emphasis)

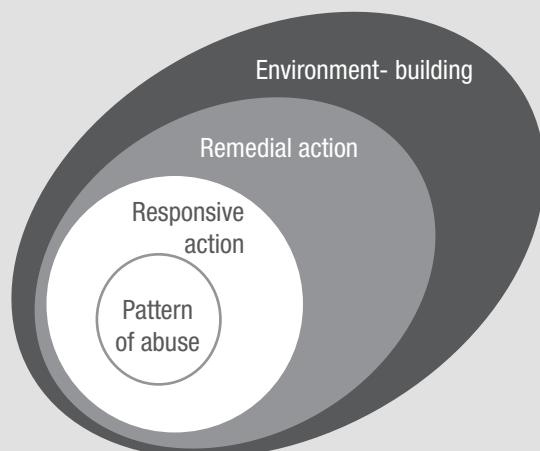
Acknowledging this diversity of perspectives, the ICRC proposed the so-called egg protection framework (Figure 2), which identifies three types of non-hierarchical and interdependent protection actions that can be carried out simultaneously (Giossi Caverzasio, 2001: 21-24):

- **Responsive actions** aim to stop, prevent the recurrence of, or alleviate the immediate effects of an emerging or established pattern of abuse.
- **Remedial actions** are undertaken after abuse has occurred to restore people's dignity and ensure adequate living conditions.
- **Environment-building actions** aim to foster a political, social, cultural, institutional, and legislative environment that enables or encourages national authorities to fulfil their obligations and respect individual rights.

Ideally, 'good [protection] strategies will address these three levels by making use of a wide range of very different activities' (ICRC, 2012: 29). Although the egg framework emerged from an exchange among agencies with protection mandates, it is now widely used by non-mandated agencies as well (e.g. Allaire, 2013).

The ICRC has identified five ways in which these three types of protection action can be carried out – persuasion, support, substitution, mobilisation and

Figure 2. The egg protection framework



Source: Giossi Caverzasio (2001: 21).

“

We do have a shared definition ... the real problem is that the definition is broad and therefore is understood differently. Our energies should go into finding ways to deal in a practical sense with multiple interpretations of a very broad concept.

”

Interview with advisory staff member

denunciation – depending on the context, the actors concerned and their mandate (ICRC, 2005: 395-398; 2012: 29). Various authors have proposed criteria for choosing the most appropriate type of protection action in a given context.³

These criteria include:

- the willingness and capacity of the state and the authorities to respond
- the capacity of civilian communities to help themselves
- the agency's capacity to respond
- the risk the action would create for the civilian population's security
- the political risk it would create for the agency's security and access
- the duration of the action
- the agency's experience with similar actions in a given setting
- what other actors are choosing to do.

We took note of seminal efforts to clarify foundational issues relating to protection in the different contexts in which humanitarian actors operate (see IASC, 2011b; Sutton et al., 2012b; IASC Principals, 2013; ICRC, 2013). Nevertheless, a perception of fragmentation remains among different voices in the protection arena.

For example, there still seems to be some pockets of discussions among practitioners about defining protection and agreeing on what it means in practice in different humanitarian contexts. Others, however, argue that the real outstanding issue is framing protection in a way that allows for collaboration and the accommodation of multiple perspectives and mandates. One interviewee offered an insight that captures and powerfully summarises these points:

I would challenge whether it is the lack of a shared definition [that is] the real problem, or whether it is the too simplistic framing. We do have a shared definition ... the real problem is slightly different – that the definition is broad and therefore is understood differently, and largely through the lens of the specific mandate or interest of each agency or actor. ... Perhaps this is an area where there cannot be agreed definitions that go into great depth or detail. ... Our energies should go into finding ways to deal in a practical sense with multiple interpretations of a very broad concept. (Interview with advisory staff member)

³ See Slim and Bonwick (2005: 81-83) for a concise introduction to this topic and Bonard (1999) for more details.

Why does this become an issue when evaluating protection?

We discussed above that *different perspectives on and meanings of protection coexist in humanitarian action*, and that the concept itself is multifaceted. This complicates the task of articulating policies and translating strategies into actionable programming and interventions. These different perspectives and meanings may form part of the same intervention and may feed – often implicitly – into the analysis and formulation of assumptions and strategies that drive programming.

From an evaluation point of view, this is challenging because it requires evaluators firstly to be alert to these different perspectives on protection and, secondly, to disentangle their various related actions and expected results so that they can define these actions' relevance, value and relative contribution to a given result. This last point takes us to the next question, which is *how to identify the expected results linked to different protection actions and frame an evaluation accordingly*.

“

One of the main issues raised by interviewees was the challenge of identifying relevant results in an evaluation and what 'success' looks like in terms of protection, in order to be able to establish how, why, for whom and in which context 'success' came about.

”

2.2 Identifying results of interest, defining 'success' and framing the evaluation

One of the main issues raised by interviewees was the challenge of identifying results and what 'success' looks like in terms of protection in humanitarian action, in order to be able to establish *how, why, for whom and in which context 'success' came about* (see sec. 2.4).

The challenge here is that *the different types of protection actions outlined above tend to aim for different results*. Reflecting on this, we considered that in the broadest sense one can identify *three main programming scenarios that correspond to three different evaluation scenarios*.

Scenario 1: When protection is provided through specialised/dedicated actions

Dedicated and specialised protection actions (also described in the literature as vertical or stand-alone protection actions) are the traditional remit of the work of protection actors with a specific mandate spelled out in international legal instruments (see Box 1 for more details). Murray and Landry (2013: 5) reminds us that this type of protection action is the one usually featured in the ‘protection chapter’ of some key humanitarian funding tools such as the consolidated appeal process.

Ideally, in this programming scenario the desired outcomes of the intervention should explicitly articulate and speak to protection issues (e.g. Davies and Ngendakuriyo, 2009; de sas Kropiwnicki, 2012). This programming scenario results in an evaluation where protection can be – and has often been – treated as a main line of evaluative inquiry. Some examples come from the evaluation of the relevance, quality, and results of specialist protection actions and services such as legal referral for refugees, displaced and stateless persons, or the evaluation of specialist child protection work, specialist psychosocial support services or dedicated services provided to victims of SGBV.

Box 1. The global protection architecture and its evolution

When discussing protection, it is important to emphasise upfront that primary responsibility for assistance and protection in the contexts of conflicts and natural disasters lies with states and national authorities (e.g. IASC, 2011b; ICRC, 2013). States are legally obliged to protect, respect and fulfil the human rights of all persons – displaced or not – within their jurisdiction in accordance with the standards of national law and international humanitarian law, and human rights and refugee law.¹

In situations of armed conflict, responsibility for protection extends under international humanitarian law to all parties to the conflict, including armed non-state actors (ICRC, 2013: 45).

¹ See Haider (2013) for a recent overview and Bouchet-Saulnier (2013) for details and an explanation of international legal terminology.



When national authorities lack the capacity or will to ensure the protection of people under their jurisdiction – or worse, themselves violate the population’s human rights – other actors (national and international) play important roles in protection (ICRC, 2013: 45; see also Forsythe, 2001). These actors constitute the backbone of the global protection architecture. They include international justice institutions and security, human rights and aid actors, including humanitarian actors.

States have conferred specific protection mandates on a number of international humanitarian and human rights organisations, including the ICRC, the United Nations (UN) Office of the High Commissioner for Human Rights, the UN High Commissioner for Refugees (UNHCR) and the UN Children’s Fund (UNICEF). Their mandates derive from a variety of sources, including international treaties, statutes of the Red Cross/Red Crescent Movement, and UN General Assembly resolutions (ICRC, 2013: 45; see also Mackintosh, 2000).

Affected communities themselves also play a role in reducing threats and augmenting their own security (Reichhold et al., 2013: 18); but as one interviewee noted, depending on the circumstances, they could also aggravate their own risk of and exposure to harm (see South et al., 2012).

Several observers have noted that it is only in the light of these foundational elements in the global protection architecture that one should calibrate the overall ambitions and aspirations of what humanitarian actors can aim to achieve in the face of some of the gravest issues – such as coercion and extreme violence – that drive the protection needs confronting aid actors on the ground.²

² For more discussion on this point see, for instance, Ferris (2011), DuBois and Barnett (2007) and Niland et al. (2014).





The global protection architecture – at least in the segment that relates to the sphere of humanitarian action – has evolved since the late 1990s. Protection work in humanitarian contexts used to be limited to two types of actors: those with the specific mandates mentioned above, and those without an explicit mandate, but with specific protection expertise (such as the Danish Refugee Council and Norwegian Refugee Council). However, beginning in the mid-1990s, changes in the environment for protection work (ICRC, 2013a) have led an increasingly diverse set of actors to design and carry out programmes in humanitarian contexts with explicit protection components, e.g. Oxfam GB (see McAvoy, 2005), World Vision, Mercy Corps and the International Rescue Committee (IRC). Without holding a specific mandate, a growing number of agencies have started internal and inter-agency discussions on to what extent and how to integrate protection principles – such as not causing harm and ensuring equality and meaningful access to impartial assistance – into their areas and sector of work.³ Other guidance and examples on integrating and mainstreaming protection concerns in various sectors can be found in a dedicated set of inter-agency guidelines (Sutton et al., 2012a). Agency-specific examples are given, for instance, in Walden et al. (2007), Leguéné et al. (2012), Save the Children (2012b) and Crawford and Pattugalan (2013).

3 An overview of the principles is featured in Sphere Project (2011: 28-30) and is also provided by the GPC (n.d.: 1).



Scenario 2: When protection is integrated into other sectoral and multi-sectoral interventions

In this scenario protection is integrated into other humanitarian sectoral and multi-sectoral programmes (e.g. health, nutrition, livelihood, education, shelter, camp coordination and camp management interventions, etc.). Ideally, in this programming scenario the desired outcomes of the intervention should be articulated in sector-specific terms and should include some reference to protection, which may then be treated in an evaluation as a secondary or auxiliary line of inquiry.

For example, in the series of mixed-method impact evaluations of the contribution of food assistance to durable solutions in protracted refugee situations commissioned jointly by UNHCR and the World Food Programme (Canteli et al., 2013) protection has been looked at – together with gender – as one of the five result areas (which included food security, nutrition, livelihoods, and effects of food assistance on relations between refugees and host populations) that the evaluation studied. Other examples are found in Bonard et al. (2010), Leguéné et al. (2012) and Bennet et al. (2012).

“

In the broadest sense one can identify three main programming scenarios that correspond to three different evaluation scenarios: When protection is provided through specialised/dedicated actions (scenario 1), when it is integrated into other interventions (scenario 2) or when reference to it is implicit (scenario 3).

”

Scenario 3: When reference to protection is implicit and derives from do-no-harm and safe, accessible and dignified programming

In this scenario we find that reference to and the translation of protection in a given humanitarian programme or intervention is implicit and derives from the application of various principles and frameworks – e.g. relating to *safe programming* and the *do no harm* principle.

It is worth noting that programming guidance and various agency literature are inconsistent when labelling this scenario, which is described by some as *protection mainstreaming* and by others as *safe, dignified and accessible programming*, or just *safe programming*. Additional details on this issue are presented in Box 2; however, the point is that in this scenario programme documents and related programme outcomes may not include any direct reference to protection. There may be references to general agency guidelines and policy commitments relating to protection, but how they are translated into programming actions (including resources- and funding-wise) may not be explicitly articulated.

In this scenario, an evaluator may not be asked to explicitly look at protection. Nevertheless, the evaluative analysis can touch on protection issues framed, for instance, in terms of ‘do no harm’, safety, access and dignity. The scope and level of ambition in terms of protection-specific analysis may be different compared to the first two scenarios outlined above.

Box 2. Where does protection mainstreaming start?

There are different descriptions of and perspectives on protection mainstreaming.

The latest ICRC professional standards for protection state that all humanitarian actors should integrate protection concerns into their practice. This can be done via protection activities as such, as part of their protection mainstreaming efforts, or as part of their application of the ‘do no harm’ principle or good quality programming (ICRC, 2013: 14). The GPC (n.d.: 1) sees protection mainstreaming as the process of incorporating protection principles and promoting meaningful access, safety and dignity in humanitarian aid. To give another example, the IRC (2013: 5) defines protection mainstreaming as ‘the process through which fundamental human rights principles, including non-discrimination, meaningful access and safety and dignity are recognized and realized in program design and implementation’.

Some observers have argued that the actions described as protection mainstreaming – that may include actions such as ensuring that latrines have lights and locks – are more a matter of ‘good programming’ rather than protection as such. They suggest that a helpful way of identifying where good quality programming ends and more protection-specific work starts is by identifying whether an intervention’s desired outcome is expressed in sectoral terms (such as improved public health or education outcomes) or in protection-specific terms (InterAction, 2013b).

An annotated reference list on protection mainstreaming compiled by the GPC is available at www.globalprotectioncluster.org/en/areas-of-responsibility/protection-mainstreaming.html, while various points of view on this issue can be found in Dolan and Hovil (2006), DuBois and Barnett (2007), Soussan (2008), Ferris and Starck (2010) and Ferris (2011).

Challenges common to the three scenarios

The first challenge appears to be that evaluators looking at protection struggle to reconstruct or infer how an intervention was understood to work, what results it was intended to achieve and on which understanding of protection it was based. One interviewee made the following observation:

I think the core foundational challenge in evaluating humanitarian protection work is definitional. ... Perhaps it is less of a problem for certain thematic areas of protection work, such as child protection, where actors are beginning to converge around meaning and language. When it comes to protection, everybody thinks they are doing it, but what are the expected outcomes we want to reach? For example, is it violence reduction or also family strengthening? ... Where do we draw a line? This echoes the other questions of whether humanitarian actors are only aiming at saving lives or also at protecting dignity and rights. (Interview with a researcher and evaluator)

This reflection was echoed by other interviewees and underscores the difficulty of translating protection concerns, and protection analysis and strategies into programmes and actions that are measurable and link into monitoring and evaluation systems.

A second set of challenges that seem to be common to protection actions in the three scenarios outlined above relate to the absence or poor quality of programme management and planning tools and to 'the absence of detailed indicators of process, results and impact' (de sas Kropiwnicki, 2012: ix). These challenges relate to the fact that protection objectives tend to be overly ambitious and unrealistic, unfocused, and/or fail to relate to specified outputs. Risks and assumptions are often vague and not accompanied by risk mitigation strategies (de sas Kropiwnicki, 2012: x).

“

A challenge to all three protection scenarios is when evaluators struggle to reconstruct how an intervention was understood to work, what results it intended to achieve and on which understanding of protection it was based.

”

Implications for evaluation

The challenges and shortcomings outlined above have repercussions for evaluation, for instance:

Clearly articulating the expected results of an intervention affects whether a programme can be evaluated, because it determines the understanding of what to evaluate and where to look for sources of data to establish evidence of its results, whether positive or negative, intended or unintended (e.g. Bamberger et al., 2012: 181-205).

Understanding causation (as discussed in more detail in section 2.4) is greatly complicated when the expected outputs, outcomes, and impacts are not spelled out in strategies and programme documents (e.g. Davidson, 2005: 67-84).

Moreover, the description of the three programming and evaluation scenarios above reinforces the point made earlier on the multifaceted nature of protection. From an evaluation perspective, each scenario should call for employing appropriate tactics to:

- single out relevant protection results at different levels (output, process, outcome, impact) and at different time points during the life of an intervention and beyond
- understand the relevant 'programme mechanisms' (Westhorp, 2014: 5) that are likely to operate in each scenario and influence how results are achieved
- calibrate stakeholder expectations about the depth and comprehensiveness of the evaluative analysis that are feasible and useful for each scenario.

2.3 Collecting and analysing data

A necessary step in each evaluation exercise is that of gathering data to answer descriptive questions about the activities of the project/programme/policy, their possible results and the context in which they were implemented (BetterEvaluation, 2014: 6). In this context an evaluation of protection encounters two main problems:

1. The first relates to *accessing and managing sensitive data in preparation for and during an evaluation of protection.*
2. The second involves *dealing with the less tangible components of protection work in humanitarian action – including, most notably, advocacy work.*

These two issues are discussed below.

Accessing and managing sensitive data

The sensitivity of protection information makes evaluating humanitarian protection exceptionally complex, and there is a heightened need for confidentiality.

Several interviewees reflected on the point that in certain contexts, just the act of gathering protection-related data and the very presence of evaluation staff can inadvertently increase communities and individual exposure to risk. (See ICRC, 2013 chapter 6, pp.77-102). Moreover, the way in which organisations deal with sensitive data throughout the life of a programme or intervention and then during an evaluation can also compromise its access to affected populations.

Constraints on data gathering are found throughout the protection programme cycle, but may be amplified during evaluation. People who have suffered trauma, such as physical or psychological abuse, may not want to share their experience and may not feel comfortable being asked about it (see Stark et al., 2014). Sharing information about their experiences could also put them at risk of further harm – from perpetrators, from the stigma attached to their experience or from reliving their suffering:

“

The sensitivity of protection information makes evaluating humanitarian protection exceptionally complex, and there is a heightened need for confidentiality.

”

There can be more risk than benefit in involving those who are at risk or have suffered harm and violence. ... This is about people talking about their lives, people talking about having suffered extortion from their own leaders, people who have been forced to live as displaced, children who have suffered abuses. ... When you gather data for protection work, there are lives at stake in a way that is different than when you gather data for health and nutrition programmes, for instance. (Interview with a researcher/evaluator)

The risk-benefit ratio of gathering data for a protection evaluation may be too high:

It is very difficult to determine the extent of sexual violence given prevailing cultural factors which make it difficult for victims to talk about such crimes. Direct information is particularly hard to come by, especially during a short stay of a few days. As in many parts of the work, reticence to discuss SGBV reflects concerns about stigma and possible recriminations. In fact, the team was informed that Sudanese refugees who speak out about problems, and particularly SGBV issues, risk being considered as 'traitors' in Yida ... there were recriminations against staff of one NGO attempting to highlight SGBV and promote reproductive health (the latter is sometimes seen as an inducement to 'promiscuity' and a threat to men's control over fertility). There have been difficulties even in getting female translators to translate information relating to cases of SGBV. (Ambroso et al., 2013: 22-23)

The evaluator may need to make more restrictive choices about gathering data from primary stakeholders and triangulating evaluative data than would be the case in other types of evaluations:

I am not sure evaluators need to talk to the victims in all cases. There seems to be an assumption that one can only conduct evaluations if you can gather information first hand. The type of information you need in an evaluation depends on the type of activities you are evaluating. You may not need in all cases to gather individual-level data. (Interview with advisory staff member)

When evaluators approach the task of looking at protection they almost have to rely on secondary information, as in a meta-evaluation. There are major limitations because you often cannot ask ... affected people [directly], and during focus group discussions what we do is inferring. ... This is different compared to evaluating food consumption, for instance, where you can look at sources such as [the Famine Early Warning System and Emergency Food Security Assessment]. This is not the case for protection, where we often have no other literature or studies. (Interview with advisory staff member)

Information collected both during an evaluation and throughout a protection programme includes sensitive details about people's traumatic experiences and the support offered to them, including legal and administrative support, psychosocial support, and counselling. Evaluators need to decide whether and how to gather and use this type of personal information and how to ensure that confidential information stays confidential. (See for instance Myer et al., 2009)

The issue of informed consent needs to be emphasised here. Ensuring informed consent is about ensuring that people understand why they are being asked questions and how their answers will be used (see ICRC, 2013: 93). Programme participants and other stakeholders approached and contacted as part of an evaluation should freely consent to participate in the exercise without being unduly pressured to do so. Informed consent also includes reassurance that declining to participate will not affect, for instance, any services provided to those who prefer not to participate (Brikci and Green, 2007: 5).

Important questions for evaluators to address are whether it is always necessary to gather sensitive individual-level information, and what the trade-offs would be if an evaluation instead emphasised data about broader protection trends in a more aggregate manner as opposed to more granular data.

“

Evaluators need to decide whether and how to gather and use sensitive personal information and how to ensure that confidential information stays confidential.

”

Sensitive protection-related data can have legal/ prosecutorial and disciplinary implications. For example, there could be cases in which the data gathered as part of an evaluation points towards criminal activities, violent acts liable to prosecution under national legislation, or malpractice and abuse (including sexual abuse) on the part of agency staff or partners. These circumstances can raise issues related to access to, disclosure and reporting of sensitive information on the part of those who commission an evaluation, those who receive an evaluation mission during fieldwork and the evaluation team carrying out the exercise.⁴ It remains to be seen whether humanitarian evaluation policies are currently adequately equipped to deal with those types of circumstances.

Analysis, findings and recommendations based on protection-sensitive information can also be politically sensitive, which can affect the ability and willingness of evaluators and agencies that commission evaluations to publicly disclose evaluation reports. The following reflection on the activities of the ICRC also applies to other actors carrying out protection work:

The ICRC constantly needs to balance its need to respect confidentiality with being accountable to its stakeholders. This has implications for results-based reporting and is one of the main reasons why it is easier for the ICRC to report outcomes for their assistance activities than for protection, even though protection is often a priority need. (Baker et al., 2013: 43)

A recent evaluation reflected on the risks of disseminating sensitive information:

Study findings include the sensitive issues of large-scale influx of settlers to the host area, large numbers of people from the host community being registered as refugees, and the vibrant economic interactions that the camps have stimulated. Disseminating these unprecedented findings risks creating a more difficult relationship between [the Government of Kenya], humanitarian agencies and host communities, as they shed new light on the significant benefits of the camps for the hosting area, as well as their negative impacts, and put into perspective the frequently-raised issue of compensation to host communities for the hosting of refugees. (Enghoff, 2010: 19)

⁴ In an example from the ICRC, Rule 73 of the Rules of Procedure and Evidence of the International Criminal Court requires confidential ICRC information to be treated as privileged, and thus not subject to disclosure, unless the ICRC consents to disclosure or the information is public (Jeannet, 2000).

Another evaluation addressed the difficulty of substantiating conclusions based on highly sensitive information:

Due to its sensitive nature, some of the material collected for this review has not been included in this report. The conclusions and recommendations aim to elicit important learning from the full range of experience including that which is not documented here, and the reader may find some disjunction in making direct links from case studies to some conclusions and recommendations. (Oxfam, 2011: 37)

Less tangible components of protection

We consider two main issues here. The first is the fact that the perceptual and psychosocial dimensions of protection – e.g. thoughts, emotions, behaviours, perceptions and social experiences – add complexity to protection work in general, and by extension to its evaluation:

Exposure to the disruption, loss, and violence associated with humanitarian crises places significant psychological and social strain on children and adults, their families and communities. ... It is essential that social and psychological issues are not ignored while homes are rebuilt, social services re-established and livelihoods recommenced. (Ager et al., 2011: 19)

The concepts of dignity and protection are important values, although they are hard to translate into ‘measurable outcomes’. Whilst technical knowledge has significantly progressed in the traditional assistance sectors of food, health, shelter, water and sanitation, there remains relatively little experience in estimating the impact of interventions on the protection of civilians. (Hofmann et al., 2004: 20)⁵

Protection programmes may identify improved safety and well-being as some of the outcomes of interest they are working towards. In these circumstances the issue confronting evaluators is above all one of *measurement*. The challenges include: (1) describing these types of concepts and results in ways that take into account their perceptual and socially constructed dimensions; and (2) identifying and tracking over time composite indicators that can help measure them.

“
The perceptual and psychosocial dimensions of protection add complexity to protection work in general, and by extension to its evaluation.

”

⁵ See also Boothby et al. (2006); Devereux et al. (2013).

“

The process of accessing and gathering, managing and analysing sensitive information poses challenges such as restricting the options available when collecting information, and requiring evaluators to assess the suitability of measures, metrics and indicators during an evaluation.

”

A second issue is that of advocacy, which was mentioned by several interviewees as a key element of humanitarian assistance in general and protection work in particular. It is an elusive concept to frame, measure and evaluate (see Tsui et al., 2014) and, despite its central role, has seldom been the subject of conventional EHA practice and guidance.⁶

Evaluating advocacy is particularly challenging for several reasons, including the short timelines of some humanitarian interventions; the multiple actors involved in or ‘touched’ by advocacy work; and the difficulty of identifying and tracking over time outcomes relating to attitude, behaviour change and, in some cases, even changes in the law. In such circumstances the main issue confronting evaluators is that of dealing with the attribution and contribution of results and framing the evaluation in ways that allow it to capture relevant short- and potentially longer-term/emerging results.

This section has discussed how the process of accessing and gathering, managing and analysing sensitive information adds an extra layer of complexity when evaluating protection and poses challenges on several fronts:

- it greatly restricts the range of options available to evaluators when collecting information from individuals and groups in humanitarian contexts
- it requires evaluators to carefully assess the suitability of measures, metrics and indicators used during an evaluation
- it may require evaluators to develop new or customise existing indicators for evaluations that are adjusted to the various programming and evaluation scenarios outlined above (sec. 2.2).

⁶ We found little guidance on how to evaluate advocacy in humanitarian contexts, but multi-mandate organisations working across emergency and development contexts have commissioned some recent work (Reisman et al., 2007; Cohen et al, 2010; Coe and Majot, 2013).

2.4 Understanding cause-and-effect issues

Causation questions in evaluation relate to results (outcomes or impacts)⁷ and what brought them about. Not all evaluations will go – and are required to go – as far as tackling causation questions. Indeed, the bulk of EHA work goes to address more descriptive questions and questions that ask what happened during an intervention.

However, evaluation questions that ask about cause-and-effect need to be answered in order to *differentiate between the effects of a programme/intervention and other changes observed in a given context* that have not been caused or influenced by an intervention. Evaluators should be able to show the relationship between a specific intervention or series of interventions and the situation described (Knox Clarke and Darcy, 2014: 40). Evaluations that provide credible answers to causation questions should also be able to tell whether changes (positive or negative) in a given context and population have been brought about or impacted by an intervention (activity, programme, policy, etc.), as opposed to being the result of coincidental or concurrent interactions (Davidson, 2000: 17).

Any attempt to understand the causes and effects of any type of protection-related work in humanitarian contexts is plagued by the same challenges that notoriously affect any attempt to establish causation in other domains of humanitarian action and in the aid sector more in general.⁸ One interviewee captured this point:

Embarking on protection evaluation work is daunting ... because of the challenges of dealing with causality and attribution, and because designing evaluations that can deal with causation issues are daunting in general for all types of interventions in humanitarian settings. (Interview with staff in advisory position)

Leading evaluation author and practitioner Jane Davidson offers some reflections that indicate the pervasiveness of the causation challenge across the whole evaluation field:

“

Evaluations that provide credible answers to causation questions should also be able to tell whether changes in a given context and population have been brought about or impacted by an intervention, as opposed to being the result of coincidental or concurrent interactions.

”

⁷ In programming and evaluation literature the two terms are used inconsistently. What some agencies describe as ‘longer-term outcomes’ are called ‘impacts’ in other aid actor/funding agency programming documents.

⁸ For an introduction to the issue of causation in general evaluation theory and practice, two solid starting points are Scriven (1991) and Rogers (2001). For an introduction to causation in evaluation written from the perspective of aid and development cooperation, see Stern et al. (2012).

“

We [evaluators] often either oversimplify causation or overcomplicate it!... Demonstrating causal links can seem like an impossible task... many people either abandon the issue altogether either by tacking a bunch of disclaimers to their evaluations or by downplaying its importance.

”

Davidson, 2005 and 2013

As a profession, we [evaluators] often either oversimplify causation or we overcomplicate it! (Davidson, 2013: 3)

Although the causation issue is incredibly important, demonstrating causal links can seem like an impossible task, especially for evaluators with limited time and resources (most of us likely fall into that category). For this reason, many people abandon the issue altogether, either by tacking on a bunch of disclaimers to their evaluations or by downplaying the importance of causal analysis. (Davidson, 2005: 68)

Several protection-specific features add to the complexity of addressing causation in evaluation, but it is worth noting here the ‘classic’ issues that *all* evaluators encounter in virtually *all* more or less unstable and fragile programme environments – including the one where humanitarian actors operate:

Results are unlikely to be caused by any single factor (Proudlock et al., 2009: 6), and causal links are often ‘messy, unpredictable and iterative’ (Proudlock et al., 2009: 16-17). It follows that any result (outcome or impact, whether positive or negative) that can be observed in a given context will be a result of interactions within and across systems and is not simply the result of a single programme, policy or other form of intervention (Westhorp, 2014: 4-5).

A related point is that *evaluation designs must be robust enough to allow other possible explanations for observed changes to be ruled out* (Morra Imas and Rist, 2009: 252-256). This is one method that evaluators can use to establish that an observed result is not the consequence of coincidental and concurrent changes (Davidson, 2000: 17).

Specific to humanitarian action are the challenges related to the *availability, accessibility, and quality of data and evidence generated from assessment and monitoring systems* (Knox Clarke and Darcy, 2014).⁹ This specific issue of data availability and quality affects humanitarian evaluation practice in general, but is particularly acute when evaluators attempt to answer causal questions. For example, this is because monitoring systems and other sources of periodic data gathered on programme performance often struggle to move beyond output and process indicators to capture early or emerging outcomes. Without monitoring or other types of periodic

⁹ See Knox Clarke and Darcy (2014) for an overview of this issue and Walden (2013) for a specific discussion on how fragile and conflict contexts affect monitoring, evaluation, accountability and learning systems.

sources of information that are regularly updated to capture results and changes, as opposed to just outputs and processes, evaluators' attempts to establish causation become exponentially more complex. They also become more cumbersome in terms of costs; time; and the skills, size, and composition of the evaluation team needed to gather new data, and often reconstruct baseline data (Bamberger, 2010).

Another issue specific to humanitarian evaluation practice – which often attracts the polarised views of proponents and opponents¹⁰ is that of the *desirability and feasibility of resorting to evaluation approaches that use a counterfactual, or other forms of control groups, as the main tactic to establish causation*. (See Puri et al., 2014).

Resorting to this sort of evaluation approaches to conduct a causal analysis in the case of protection work is likely to attract even more polarised views and frictions. Depending on the specific evaluation design chosen and the way in which a comparison group is identified, ethical issues can arise to compound the practical and feasibility issues confronting evaluators who resort to these approaches. There are practical ways to overcome some of them, however, including in humanitarian settings (see Buttenheim, 2009). To date, the evaluation of protection work does not appear to have crossed paths with evaluation approaches that use counterfactuals, but various options can be pursued to explore possible alternative explanations of results. The next ALNAP study will feature some guidance looking specifically at those evaluation design issues.

All the interviewees seemed to agree that understanding causation in protection evaluation is particularly difficult because it requires examining longer time frames, distinguishing between different levels of results and trying to focus further down the results chain. The further down the results chain we wish to focus an evaluation, the more complex the evaluators' job, because more influences, interactions and actors come into play. This is even more the case in fluid and unstable humanitarian contexts. A further problem arises when protection programmes attempt to work with or influence the behaviour of actors ranging from local authorities, to non-state armed groups and militias, to traditional chiefs and local leadership figures in a community. They all embody individual and collective beliefs reflecting social norms that one intervention or programme alone would be hardly likely to tackle and change.

“

All the interviewees seemed to agree that understanding causation in protection evaluation is particularly difficult because it requires examining longer time frames, distinguishing between different levels of results and trying to focus further down the results chain.

”

10 Refer to Stern et al. (2012) for an overview, in particular chap. 3, pp. 14-35.

Several interviewees also reflected on the challenge of customising an evaluation approach to take into account the various ways in which protection is understood to work in the different programming scenarios outlined above. This makes it challenging to (1) establish the *contribution of specific programme actions to different levels of results*; and (2) identify which programme actions and interactions with people and contexts bring about and sustain results over time. For example, the process of integrating protection into other work complicates the task of establishing the relative value and weight of different protection components and how they contribute to specific results. An evaluation of the actions of the European Commission's Directorate-General for Humanitarian Aid (DG ECHO) in the Democratic Republic of the Congo touched on this specific challenge:

Effectiveness can only be measured in projects which include clearly identifiable results, components or activities. A review of the project portfolio shows that, except for the funding of the protection mandated agencies, DG ECHO does not appear to fund enough projects which include protection results or clearly identified protection components. Therefore, the evaluation will be limited to a detailed review of the principles of these projects. (Bonard et al., 2010: 37)

So where does this leave us? At this stage there are more questions than answers. This reflection on causation, and on the desire to understand the contribution that a given intervention or package of interventions bring to achieve protection results, leads us to pose three final questions to evaluators and staff who commission evaluations:

1. Is it feasible and always desirable to jump straight into asking causation questions when evaluating protection?
2. How do we get ready to tackle cause-and-effect questions in protection? Which types of evaluative questions should we pose and which evaluative exercises should we consider conducting to better prepare us to address causation questions in protection?
3. Is it viable – and perhaps more useful – to treat protection not only as the subject – the WHAT – of an evaluation, but also as a separate evaluation criterion to use for humanitarian programmes and interventions that do not have an explicit protection focus or explicit protection outcomes?

The final section of this paper suggests the next steps we conceived to answer these questions and deal with the other questions raised in this scoping paper.

3. Next steps

This paper has identified key challenges in the evaluation of protection that add to and amplify the complexity of EHA work.

During this preliminary research and scoping work it emerged that the evaluation of protection in humanitarian action is impacted by a similar set of complications and challenges that affect other branches of EHA discussed elsewhere in humanitarian literature but with an extra layer of complexity.¹¹ This is worth noting because we do not wish to give the impression that the case of the evaluation of protection is unique. Put differently, other branches of humanitarian evaluation – and more generally of aid evaluation – are also confronted with complex programming and implementation scenarios that are in many ways comparable to that of protection. The evaluation of advocacy work and peace-building are two examples that come to mind and to which we will return in search of suggestions and good practices that could assist humanitarian evaluators of protection.

The next goal in this research and guidance development process is to explore ways to address the challenges that have been identified and to offer practical guidance that can be tested in real-world evaluation exercises with the support of ALNAP Members who are interested in joining this work.

The pilot guidance document is expected to be ready in mid-2015. Its development will include identifying evaluation design options that address protection issues; exploring their requirements, advantages and limits; and presenting practical suggestions for organising evaluative data and analysing findings. This will be carried out in three steps:

1. The first step will be to compile all interviewee suggestions regarding evaluation tools and approaches. These leads will be investigated further, relevant evaluation literature will be reviewed, and the approaches that are most relevant to the protection evaluation challenges identified in the present paper will be chosen.
2. This material will then be complemented with evaluation guidance from outside the humanitarian evaluation field.
3. The third step will be to compile the best guidance obtained from these two sources.

“

The case of evaluation of protection in humanitarian action is not unique, as it is impacted by a similar set of complications and challenges that affect other branches of EHA but with an extra layer of complexity.

”

¹¹ See, for instance, Wood et al. (2001); Hofman et al. (2004); Guerrero et al. (2013); Knox Clarke and Darcy (2014, secs. 3.1 and 4.4).

Many interviewees made excellent suggestions on lines of inquiry to follow and specific evaluation tools to review. These included guidance emerging from work on evaluating mental health and psychosocial support in emergency settings, which may be particularly relevant in its approach to dealing with less tangible outcomes and more perceptual issues (e.g. IASC, 2007; Ager et al., 2014). Others suggested work in the area of child protection, particularly on developing a menu of outcome indicators for different types of programming in different contexts (e.g. Save the Children, 2012b). Interviewees also mentioned recent work on safety audits in fragile contexts (e.g. Moser, 2012) and further exploration of the use of logic models, including theories of change (e.g. Babbitt et al., 2013; Vogel, 2012; Woodrow and Oatley, 2013). One interviewee commented in this regard:

We don't currently use theories of change in our programmes, but one idea worth exploring would be to use them to measure changes in a situation without getting caught into looking only at possible causality of only our actions and programmes, but mapping other possible elements and plausible explanations of what causes change. (Interview with programme and advisory staff member)

Guidance will be explored from outside the EHA sphere of practice looking at evaluation of advocacy, policy influence and peace-building (e.g. OECD, 2012; Woodrow and Oatley, 2013; Chigas et al., 2014; Tsui et al., 2014). These sources could yield valuable advice on how to handle complex tasks such as evaluating programmes established to achieve incremental change over longer time frames. They could also be helpful for programmes designed to influence the behaviours and attitudes of diverse stakeholders beyond a particular agency's sphere of control or influence, and collaborative interventions in dynamic and rapidly changing environments.

This undertaking will only have value if the results have the potential to improve the quality and robustness of the evidence generated by evaluations of protection. The following criteria drawn from those proposed by Knox-Clarke and Darcy (2014) and Spencer et al. (2003) will be used to guide this work:

- **Transparency and traceability:** Research outputs and guidance document issued will be made accessible to the broad ALNAP Network and beyond. They will be thoroughly referenced and include a method note that enables other researchers examining the same subject to trace the steps taken in this research and test the conclusions reached.
- **Accuracy** will be improved by conducting interviews with practitioners in different fields (evaluation, research and protection programming) to cross-check points emerging from other interviews.
- **Identifying and minimising bias:** Clear inclusion and exclusion criteria will be identified and adhered to when searching for, retrieving and selecting documents for review.
- **Robustness of resources selected:** Greater weight and preference will be given to evaluative literature – including grey literature, evaluation guidance from other fields and agency-specific documents – containing some details about contexts and methods used to generate a particular document or a specific piece of research or guidance.
- **Transferability:** The unit of analysis, level, scope and focus of each approach suggested in the evaluation guidance will be made explicit so that users can assess whether it can be adapted to their specific needs.

Additional notes on the methodology used for this preliminary study are presented in Annex B.

ALNAP invites Members who are working on evaluation of protection to explore the possibility of collaboration. Correspondence can be addressed to Francesca Bonino at f.bonino@alnap.org.

Annexes



Annex A. The concept of protection in humanitarian evaluation guidance

The concept of protection has been discussed in a number of EHA guidance documents. This annex focuses mainly on those produced at the inter-agency level. This is because they were not written from an agency-specific or sector-specific perspective, are addressed to a broad audience in both operational and donor agencies, and are therefore more likely to be broadly relevant for both single- and inter-agency evaluations.

Hallam (1998) Evaluating Humanitarian Assistance Programmes in Complex Emergencies. Relief and Rehabilitation Network Good Practice Review.

This review discusses protection as an important part of a humanitarian evaluation:

Evaluations of humanitarian assistance programmes should include an assessment of measures to provide protection to the affected population. In many emergency situations, the first needs of a population under threat may be protection from murder and harassment, as well as from discrimination that can lead to exclusion from basic services. The evaluation should also assess what measures have been taken to mitigate potential negative consequences of the humanitarian programme. (Hallam, 1998: 13)

The discussion of good practice does not include a definition of humanitarian protection, although it is implicit to a degree in the quotation given above. It does offer some clear language on the need to consider protection in all evaluations of humanitarian action:

Given the context of conflict and insecurity, protection issues are critical to the effectiveness of humanitarian action. ... Assessment of the level of security and protection in the project or programme area and, where relevant, the steps taken to improve them, should be part of all humanitarian assistance evaluations. In the humanitarian assistance evaluations undertaken to date, such issues have often been left out or inadequately covered. (Hallam, 1998: 54)

The review also suggests equating the importance of and need for humanitarian agencies to evaluate protection with the importance of and need for assessing responses to other issues or concerns (Hallam, 1998: 34-35). The document discussed next was informed by this review.

OECD-DAC (2001) Guidance for Evaluating Humanitarian Assistance in Complex Emergencies

The OECD-DAC (2001) describes the record of humanitarian evaluation in assessing protection as 'very patchy':

In large part this has stemmed from such issues not being regarded as being a major concern of evaluation and the work of evaluators. Evaluations of humanitarian assistance programmes should include an assessment of the humanitarian space available, the level of human rights abuses and the measures taken to provide protection to the affected population. It will be important to explore the gender dimension of the abuses and the extent to which the agencies have taken account of this in the design of their programmes. (OECD-DAC, 2001: 15)

This document suggests that EHA should also assess the level of human rights abuses, which is advice not repeated in later evaluation guidance. Not only does this area of inquiry require specific investigative skills, but it also creates a responsibility to use the resulting information in either public advocacy or discussions with authorities and other duty bearers. This is not typical EHA territory.

The OECD-DAC proposes that in the case of humanitarian assistance, especially in complex emergencies, additional evaluation criteria – including protection – should be added to the standard OECD-DAC list (efficiency, effectiveness, impact, sustainability and relevance).

Given the context of conflict and insecurity, protection issues are also critical to the effectiveness of humanitarian action. Where levels of protection are poor it is feasible that the target population of an otherwise effective project distributing relief assistance are being killed by armed elements operating within the project area or even within the displaced persons/refugee camp. Assessment of the levels of security and protection in the area of the project or programme and, where relevant, the steps taken to improve them should be part of all humanitarian assistance evaluations. In those humanitarian assistance evaluations undertaken to date, such issues have often been left out of the study or not adequately covered. (OECD-DAC, 2001: 23)

The document points to two possible reasons for the inadequate coverage of protection in EHA: evaluation managers' lack of familiarity with protection issues, and the dearth of frameworks and norms against which humanitarian agencies' protection work can be assessed (OECD-DAC, 2001: 24).

Beck (2006) Evaluating humanitarian action using the OECD-DAC criteria: An ALNAP guide for humanitarian agencies

This document discusses protection in terms of two EHA criteria: coherence and coverage. It defines the evaluation of coherence as the assessment of 'security, developmental, trade and military policies, as well as humanitarian policies, to ensure that there is consistency and, in particular, that all policies take into account humanitarian and human-rights considerations' (Beck, 2006: 33). While arguing that protection should be taken into account under the coherence criterion, Beck (2006: 34) does not go into greater detail on the subject. However, he concludes by noting that coherence-related issues in an evaluation will likely be very political and sensitive in nature, and advises evaluation managers to ensure that their teams have adequate capacities and resources to cover these issues in an evaluation.

The second protection-related criterion, coverage, is defined as the capacity to reach 'major population groups facing life-threatening risk wherever they are' (Beck, 2006: 38), and as including both assistance and protection. The author suggests that EHA should examine the reasons why an intervention provided or failed to provide affected populations with assistance *and* protection proportionate to their needs:

Evaluating whether protection needs have been met is a key element in the assessment of coverage. Evaluators should ask whether those who needed protection have received it. Even where protection issues do not form a major part of the intervention, evaluators should still assess whether protection issues should have been integrated into planning. (Beck, 2006: 39)

This advice is noteworthy because it suggests that protection should be evaluated regardless of whether it was an explicit primary focus of an intervention or not.

All inter-agency EHA guidance published after 2006, including the ALNAP guide on real time evaluation (Cosgrave et al., 2009) and the ALNAP pilot guide on Evaluating humanitarian action (Cosgrave and Buchanan-Smith, 2012), refers to Beck's seminal work, but does not elaborate on the issues he raises.

IASC (2011a) Inter Agency Real Time Evaluation Procedures and Methodologies

This document mentions protection briefly on two occasions:

1. Firstly, it recommends treating protection as one of the cross-cutting elements against which the overall effectiveness of inter-cluster coordination should be evaluated (IASC, 2011a: 33-35).
2. Secondly, protection-specific areas of concern such as those related to protection from sexual exploitation and abuse and SGBV are mentioned as topics of inquiry that evaluators should cover when assessing the degree to which a response was needs based and whether risks of abuse were adequately identified and addressed (IASC, 2011a: 34).

IASC-IAHE Group (2014a) Inter-Agency Humanitarian Evaluations of Large Scale System-Wide Emergencies

This document is the most recent example of inter-agency guidance that examines humanitarian issues. It mentions protection as one possible evaluation criterion in a list that includes the other 'conventional' OECD-DAC criteria of relevance, coherence, coverage, connectedness, efficiency, effectiveness, impact, sustainability and coordination (OECD-DAC, 2010).

The guidance features an impact pathway in the form of a logic model describing the ideal way in which different levels of results are linked, from outputs through outcomes to longer-term impacts. It frames protection as (1) an expected early impact of collective and coordinated humanitarian action; and (2) as part of the longer-term impact of improved protection and well-being, and greater capacity to withstand, cope with and adapt to shocks.

Concluding observations

The set of evaluation guidance reviewed above shows that protection in humanitarian action has received some level of attention, but that the concept is neither defined in these documents nor framed in an actionable way. Only a few documents offer examples of evaluation questions that address protection. When the guidance document features a more detailed discussion on protection, it is usually in conjunction with issues of access and security, which are also not clearly defined, and which do not – or only partially – equate with protection in humanitarian contexts.

The documents treat protection in different ways: as a cross-cutting issue, as an all-encompassing and overarching theme that all EHA exercises should consider, or as a possible criterion to add to the conventional OECD-DAC list.

This brief review of humanitarian evaluation guidance identifies two emerging conclusions: firstly, that there is lack of agreement on how to conceptualise and frame protection in EHA; and, secondly, that no taxonomy is available to evaluators that helps to spell out how different understandings of protection (programming and results) may call for different treatments of evaluations. These are two of the perceived gaps that the present ALNAP initiative aims to address.

Annex B. Methodology note

This annex presents the research questions that guided the development of the present paper, and the criteria used when searching for and selecting the relevant literature and identifying the key informants to be interviewed at this scoping/initial stage of the research. It also outlines the steps followed to analyse the literature and interviewee responses.

The **main questions** that guided this scoping paper were:

- Does humanitarian protection work pose specific and unique challenges to evaluation compared to other aspects of EHA for which more comprehensive evaluation guidance is available?
- If this is the case, what are these challenges and how do they affect evaluation?
- What steps can be taken to address these challenges?

The research began by identifying and describing the features of humanitarian protection discussed in the literature and identified by practitioners and evaluators as posing challenges to the evaluation of protection.

The **literature reviewed** fell into four categories:

1. inter-agency and agency-specific documents that introduce and define the concept of protection in humanitarian action and the key references in this field
2. evaluation reports that feature protection as one of their lines of inquiry or explore it as a cross-cutting issue
3. humanitarian evaluation guidance that is applicable to a broad audience rather than limited to one sector or technical area of intervention – hence the decision to review inter-agency evaluation guidance rather than single-agency and sector-specific work
4. protection programming guidance that featured a substantial discussion of evaluation.

Specifically, the desk-based research for this paper consisted of the following steps:

- a review of evaluative literature on protection in humanitarian action that drew partly on the evaluation section of the ALNAP Humanitarian Evaluation and Learning Portal (www.alnap.org/resources/results.aspx?type=22); partly on interviewees' responses to a request for suggestions of key documents; and partly on institutional website searches, e.g. of the ICRC, Danish Refugee Council, UNHCR, GPC, Oxfam, IRC and InterAction websites
- a review of humanitarian evaluation guidance documents to ascertain whether and how they cover protection, with an emphasis on those developed at the inter-agency level
- a light review of (humanitarian) agency-specific guidance on protection programming and related design, implementation and monitoring issues, provided they made some less than cursory mention of evaluation

- six interviews with staff in evaluation positions, including evaluation managers, and staff working in evaluation-commissioning roles in donor agencies
- twenty-two interviews with agency staff working in operational or protection advisory positions, mainly at headquarters, including in donor agencies, of whom two are based in field offices and seven work in surge capacity functions and travel regularly to support field-level programmes and operations
- five interviews with individuals working as independent researchers or consultants, or who are affiliated with academic institutions, and who often work across humanitarian, development, human rights, and peace and conflict disciplines.

Inclusion and exclusion criteria and possible bias

There was a potential for bias to affect the choice of literature for review, as well as the weighting of different types of literature and different protection themes and sub-themes. To protect against this as far as possible, inclusion and exclusion criteria were identified to guide the selection of documents.

Agency-specific literature and guidance on protection were reviewed if they dealt even secondarily with assessment, monitoring and evaluation issues. Literature from both protection-mandated and non-mandated agencies was included, and a balance was sought among different thematic areas pertaining to protection, including, for instance, the protection of children, the protection of civilians and refugees, protection in situations of displacement, and protection from sexual exploitation and abuse.

Documents were excluded if they dealt with protection solely in security, military or legal terms. Literature on social protection and protection programming in the global North was excluded at this stage, but will be considered in the next stage of this research project when we look for good-practice examples from non-humanitarian evaluation practice.

Bias could have affected the process used to identify key informants, which started with contacting some participants from a practitioners' workshop on results-based protection held at InterAction in 2013 (www.interaction.org/work/results-based-protection). The initial list of interviewees mainly included evaluators and agency staff based at headquarters, where many staff with a protection advisory role work. Nevertheless, use of the snowballing technique, in which initial informants are asked to provide the names of other potential informants to expand the web of contacts and potential interviewees and literature suggestions (Hagen-Zanker and Mallett, 2013), made it possible to contact several people based in country offices and staff who work in protection surge capacity functions and thus spend extensive periods of time in the field.

Using the BetterEvaluation framework to structure and analyse information collected

The BetterEvaluation framework (BetterEvaluation, 2014; see Box 1) was invaluable to this scoping paper because it supports our aim of organising and offering a coherent presentation of the issues and challenges confronting evaluators of protection in humanitarian contexts. The framework is built on a sequencing of seven broad process steps or a cluster of issues that need to be addressed in any type of evaluation exercise – regardless of the specific subject matter, the method orientation or the programming context in which the evaluation takes place.

The framework can be used in different ways (see BetterEvaluation, 2014), and in this paper we made partial use of it by focusing only on four of the seven proposed categories (i.e. defining, framing, describing and understanding causes), because they capture the bulk of insights gathered from the literature and the interviews with practitioners.

The BetterEvaluation framework is also useful for analytical purposes because it is somewhat ‘agnostic’ when proposing a particular vision of evaluation or a particular approach or evaluation method orientation. From experimental to quasi- and non-experimental evaluation approaches, the framework is able to embrace a range of options. It thus offers a levelled playing ground to all evaluators and evaluation-commissioning agencies approaching a specific evaluation issue by presenting them with actionable options and signalling trade-offs. This is exactly what we aim to achieve in the next stage of this research and evaluation guidance development initiative – hence the good fit between the framework and the research questions that drive this ALNAP work.

Interview questions

The semi-structured interviews touched on the following questions:

1. What type of (humanitarian) protection-related work does your organisation do?
2. How does your organisation monitor and evaluate protection-related work?
3. What type of information is currently being collected by evaluations looking at protection?
Are you getting what you need from those evaluations? If not, what would you like to see in humanitarian evaluations of protection?
4. What do you think is challenging about evaluating protection?
5. Have any steps been taken to address these evaluation challenges either in your organisation or elsewhere?
6. Have you seen any good examples of evaluations of humanitarian protection?
7. Are you aware of any evaluation approaches, designs and/or tools that you have not (yet) directly used, but consider promising and would like to know more about?
8. What would you like ALNAP’s work on protection evaluation to produce?

Interviewees

The following individuals were interviewed and provided information used in this paper:

Aubin, Louise	UNHCR and Global Protection Cluster
Baglolle, Deborah	UK Department for International Development
Blake, Courtney	Office of US Foreign Disaster Assistance, US Agency for International Development
Bradley, Lian	Norwegian Refugee Council
Canavera, Mark	Columbia University
Cosgrave, John	Independent consultant
Ghorkhmazyan, Meri	Save the Children
Hastie, Rachel	Oxfam GB
Jean, Isabella	CDA Collaborative Learning Projects
Kahn, Clea	UK Department for International Development
Mahony, Liam	Fieldview Solutions
McAvoy, Jenny	InterAction
Metzler, Janna	Columbia University
Michelsen, Jette	Danish Ministry of Foreign Affairs
Mikova, Velina	ICRC
Morand, MaryBeth	UNHCR
Morris, Helen	UNHCR
Niland, Norah	Independent consultant
O'Neil, Glenn	Independent consultant
Pasztor, Gergey	IRC and Global Protection Cluster
Pattugalan, Gina	World Food Programme
Polastro, Riccardo	IOD PARC
Ravier, Guilhem	ICRC
Reichhold, Urban	Global Public Policy Institute
Saavedra, Luz	ALNAP, formerly of Oxfam Intermón
Siegrist, Saudamini	UNICEF
Sooma, Patrick	World Vision
Starup, Kathrine	Danish Refugee Council
Svoboda, Eva	Humanitarian Policy Group at the Overseas Development Institute
Swamy Meier-Ewert, Gita	UNHCR
Thomsen, Thomas	Danish Ministry of Foreign Affairs
Treves, Arne	UNHCR
Winter, Cara	Norwegian Refugee Council

References

The following publications can also be accessed via the Humanitarian Evaluation and Learning Portal (HELP): <http://www.alnap.org/evaluating-protection>.

Ager, A. and Metzler, J. (2012) *Child friendly spaces. A structured review of the current evidence-base*. New York: Columbia University.

Ager, A., Ager, W., Stavrou, V. and Boothby, N. (2011) *Inter-agency guide to the evaluation of psychosocial programming in emergencies*. New York: UNICEF.

Ager, A., Robinson, S. and Metzler, J. (2014) *Methodologies and tools for measuring mental health and psychosocial wellbeing of children in humanitarian contexts*. New York: Columbia University.

Allaire, A. (2013) *Protection interventions for older people in emergencies*. London: HelpAge International.

Ambroso, G., Janz, J., Lee, V. and Salomons, M. (2013) *Flooding across the border: a review of UNHCR's response to the Sudanese refugee emergency in South Sudan*. Geneva: UNHCR/PDES.

Babbit, E., Chigas, D. and Wilkinson, R. (2013) *Theories and indicators of change briefing paper: concepts and primers for conflict management and mitigation*. Washington, DC: USAID.

Baker, J., Dross, E., Shah, V. and Polastro, R. (2013) *Study: how to define and measure value for money in the humanitarian sector*. Sida Decentralised Evaluation, Report 2013:29. Stockholm: SIDA.

Bamberger, M. (2010) *Reconstructing baseline data for impact evaluation and results measurement*. Prem Notes 4. Washington, DC: World Bank.

Bamberger, M., Ruch, J. and Mabry, L. (2012) *RealWorld evaluation. Working under budget, time, data, and political constraints*. London: Sage.

Beck, T. (2006) *Evaluating humanitarian action using the OECD-DAC criteria: an ALNAP guide for humanitarian agencies*. ALNAP Guide. London: ALNAP.

Bennet, J., Betts, J., Gayfer, J. and Dinsmore, H. (2012) *Afghanistan: an evaluation of WFP's portfolio*. Country Portfolio Evaluation. Rome: WFP.

BetterEvaluation. (2014) *BetterEvaluation Rainbow Framework*. Melbourne: BetterEvaluation.

Bonard, P. (1999) *Modes of action used by humanitarian players: criteria for operational complementarity*. Geneva: ICRC.

Bonard, P., Solé R., Hidalgo, S. and Posada, S. (2010) *Evaluation of DG ECHO's actions in the Democratic Republic of Congo*. Madrid: DARA.

Boothby, N., Newman, J., Tanabe, M., Prowitt-Smith, L., Ager, A. and Wessells, M. (2006) *Assessment and evaluation of psychosocial programming for crisis-affected children: A good practice initiative*. New York: UNICEF.

Bouchet-Saulnier, F. (2013) *Dictionnaire pratique du droit humanitaire*. Paris: Éditions La Découverte.

Brikci, N. and Green, J. (2007) *A guide to using qualitative research methodology*. London: MSF and London School of Hygiene and Tropical Medicine.

Buttenheim, A. (2009) *Impact evaluation in the post-disaster setting: a conceptual discussion in the context of the 2005 Pakistan earthquake*. 3ie Working Paper No. 5. New Delhi: 3ie.

Canavera, M. (2011) *We cannot wait for others to come and protect us: lessons emerging from Oxfam GB's community based civilian protection programme in the Eastern Democratic Republic of Congo*. Oxford: Oxfam GB.

Canteli, C., Morris, T. and Steen, N. (2012) *Synthesis of mixed method impact evaluations of the contribution of food assistance to durable solutions in protracted refugee situations*. Rome and Geneva: WFP and UNHCR.

Chigas, D., Church, M. and Corlazzoli, V. (2014) *Evaluating Impacts of Peacebuilding Interventions*. London: DFID.

Coe, J. and Majot, J. (2013) *Monitoring, evaluation and learning in NGO advocacy - findings from comparative policy advocacy MEL review project*. Boston: Oxfam America.

Cohen, D. et al. (2010) *Advocacy toolkit: A guide to influencing decisions that improve children's lives*. New York: UNICEF.

Corlazzoli, V. and White, J. (2013) *Back to basics - a compilation of best practices in design, monitoring & evaluation in fragile and conflict-affected environments*. London: DFID.

Cosgrave, J. and Buchanan-Smith, M. (2012) *Evaluation of humanitarian action*. ALNAP Guide. London: ALNAP/ODI.

Cosgrave, J., Ramalingam B. and Beck, T. (2009) *Real time evaluations of humanitarian action*. ALNAP Guide. London: ALNAP.

Crawford, N. and Pattugalan, G. (eds) (2013) *Protection in practice: food assistance with safety and dignity*. Rome: WFP.

Davidson, J. (2000) 'Ascertaining causality in theory-based evaluation', in 'Program theory in evaluation: challenges and opportunities' [special issue]. *New Directions for Evaluation*, 2000(87): 17–26.

Davidson, J. (2005) *Evaluation methodology basics: the nuts and bolts of sound evaluation*. London: Sage.

Davidson, J. (2013) *Understand causes of outcomes and impacts*. AEA/BetterEvaluation Part 5. Washington, DC: AEA.

Davies, A. and Ngendakuriyo, A. (2008) *Protection and Assistance to the Congolese refugees in Burundi*. Norwegian Refugee Council Evaluation Report. Oslo: NRC.

De Sas Kropiwnicki, Z. (2012) *Protecting children from violence: a synthesis of evaluation findings*. New York: UNICEF.

Devereux, S., Roelen, K., Béné, C., Chopra, D., Leavy, J. and McGregor, J.A. (2012) *Evaluating outside the box: an alternative framework for analysing social protection programmes*. IDS Working Paper No. 431, CSP Working Paper No. 10. Brighton, UK: CPS/IDS.

Dolan, C. and Hovil, L. (2006) *Humanitarian protection in Uganda: a trojan horse?* HPG Background Paper. London: HPG/ODI.

Donini, A. (ed) (2013) *The golden fleece: manipulation and independence in humanitarian action*. Boulder, CO: Kumarian Press.

DuBois, M. (2010) 'Protection: fig-leaves and other delusions'. *Humanitarian Exchange* 46. London: HPG/ODI.

DuBois, M. and Barnett, K. (2007) 'Protection'. *Dialogues* 4. London: MSF UK.

Egeland, J., Harmer, A., Stoddard, A. (2011) *To stay and deliver: good practice for humanitarians in complex security environments*. OCHA Policy and Studies Series. New York: UNOCHA.

Enghoff, M., Hansen, B., Umar, A., Gildestad, B., Owen, M. and Obara, A. (2010) *In search of protection and livelihood: socio-economic and environmental impacts of Dadaab refugee camps on host communities*. Nairobi: Royal Danish Embassy, Government of Kenya and Norwegian Embassy.

Ferris, E. (2011) *The politics of protection: the limits of humanitarian action*. Washington, DC: Brookings Press.

Ferris, E. and Stark, C. (2010) *Incorporating protection into humanitarian action: approaches and limits*. Washington, DC: Brookings Institution.

Forsythe, D. (2001) 'Humanitarian protection: The International Committee of the Red Cross and the United Nations High Commissioner for Refugees'. *International Review of the Red Cross*, 83(843): 675-697.

GHD. (2003) *Principles and good practice of humanitarian donorship*. Stockholm: GHD.

Giossi Caverzasio, S. (ed.) (2001) *Strengthening protection in war: a search for professional standards*. Geneva: ICRC.

GPC. (2010) *Handbook for the Protection of Internally Displaced Persons*. Geneva: GPC.

GPC. (2012a) *Protection cluster coordinator model terms of reference*. Geneva: GPC.

GPC. (2012b) *Strategic framework 2012-2014*. Geneva: GPC.

GPC. (2013) *Diagnostic tool and guidance on the interaction between field protection clusters and UN missions*. Draft July 2013. Geneva: GPC.

GPC. (2014) *Independent “whole of system” review of protection in humanitarian crises – terms of reference*. Geneva: GPC.

GPC. (n.d.) *Annotated reference list on protection mainstreaming*. Pilot version. Geneva: GPC.

Guerrero, S., Woodhead, S. and Hounjet, M. (2013) *On the right track? A brief review of monitoring and evaluation in the humanitarian sector*. London: ACF UK.

Hagen-Zanker, R. and Mallett, R. (2013) *How to do a rigorous, evidence-focused literature review in international development*. ODI Working Paper. London: ODI.

Haider, H. (2013) *International legal frameworks for humanitarian action: topic guide*. Birmingham: GSDRC.

Hallam, H. (1998) *Evaluating humanitarian assistance programmes in complex emergencies*. Relief and Rehabilitation Network Good Practice Review No. 7. London: ODI.

Harvey, P., Stoddard, A., Harmer, A. and Taylor, G. (2010) *State of the humanitarian system*. ALNAP Pilot Study. London: ALNAP/ODI.

Hastie, R., Nunn, R., Martlew, N. (2007) *Final draft Oxfam GB global protection review 2007 - part 1*. Oxford: Oxfam GB.

Hastie, R., Swithern, S., Bhreathnach, A. and Bonwick, A. (2009) *Improving the safety of civilians: a protection training pack*. Oxford: Oxfam GB.

Hofmann, C.A., Roberts, L., Shoham, J. and Harvey, P. (2004) *Measuring the impact of humanitarian aid. A review of current practice*. HPG Report No. 17. London: HPG/ODI.

HPN. (2010) *Humanitarian Protection*. Humanitarian Exchange No. 46. London: HPG/ODI.

IASC Principals. (2013) *The centrality of protection in humanitarian action – statement by the Inter-Agency Standing Committee (IASC) Principals*. Geneva: IASC.

IASC. (1999) *Protection of internally displaced persons*. Policy Paper Series No. 2. New York: IASC.

IASC. (2002) *Growing the sheltering tree: protecting rights through humanitarian action*. Geneva: IASC.

IASC. (2006) *Women, girls, boys and men: different needs – equal opportunities. IASC gender handbook in humanitarian action*. Geneva: IASC.

IASC. (2007a) *IASC guidelines on mental health and psychosocial support in emergency settings*. Geneva: IASC.

IASC. (2007b) *Protection cluster working group: mission statement and terms of reference*. Geneva: IASC.

IASC. (2010) *IASC framework on durable solutions for internally displaced persons*. Geneva: IASC.

IASC. (2011a) *Inter-agency real time evaluation procedures and methodologies*. Geneva: IASC.

IASC. (2011b) *IASC operational guidelines on the protection of persons in situations of natural disasters*. Geneva: IASC.

IASC-IAHE Group. (2014a) *Inter-agency humanitarian evaluations of large scale system-wide emergencies (IAHEs): guidelines*. New York: IASC/IAHE.

ICRC. (2005) 'Action by the International Committee of the Red Cross in the event of violations of international humanitarian law or of other fundamental rules protecting persons in situations of violence'. *International Review of the Red Cross*, 87(858): 393-400.

ICRC. (2008) 'ICRC protection policy: institutional Policy'. *International Review of the Red Cross*, 90(7): 751-775.

ICRC. (2009) *Professional standards for protection work carried out by humanitarian and human rights actors in armed conflict and other situations of violence*. 1st Ed. Geneva: ICRC.

ICRC. (2012) *Enhancing protection for civilians in armed conflict and other situations of violence*. Geneva: ICRC.

ICRC. (2013) *Professional standards for protection work carried out by humanitarian and human rights actors in armed conflict and other situations of violence*. 2nd Ed. Geneva: ICRC.

InterAction. (2012) *Project brief: results-based protection*. Washington, DC: InterAction.

InterAction. (2013a) *NGO and donor consultations on results-based protection*. Washington, DC: InterAction.

InterAction. (2013b) *Protection mainstreaming learning event: meeting note*. Washington, DC: InterAction.

IRC. (2013) *Protection mainstreaming training facilitator's guide*. Washington, DC: IRC.

Jeannet, S. (2000) 'Recognition of the ICRC's long-standing rule of confidentiality - an important decision by the International Criminal Tribunal for the former Yugoslavia'. *International Review of the Red Cross*, (838): 403-425.

Kemp, E. (2012) *Lessons from Oxfam's protection programme in Province Orientale and parts of South Kivu, DRC*. Oxford: Oxfam GB.

Knox Clarke, P. and Darcy, J. (2014) *Insufficient evidence? The quality and use of evidence in humanitarian action*. ALNAP Study. London: ALNAP/ODI.

Landis, D., Stark, L., Mansourian, H. and Ager, A. (2013) *Examining child protection rapid assessment: a structured review of field learning from the Child Protection Rapid Assessment (CPRA) toolkit*. New York: CPC Network.

Landis, D., Williamson, K., Fry, D. and Stark, L. (2013) *Measuring violence against children in humanitarian settings: a scoping exercise of methods and tools*. New York and London: CPC Network and Save the Children UK.

Leguéné, P., Canteli, C., Morris, T., Posada, S., Dionadji, B. and Ververs M. (2012) *Food assistance in protracted refugee situations in Chad (2003-2011): a joint mixed method impact evaluation*. Rome: WFP.

Mackintosh, K. (2000) *The principles of humanitarian action in international humanitarian Law. Study 4 in the politics of principle: the principles of humanitarian action in practice*. London: ODI/HPG.

Mahony, L. (2006) *Proactive presence - field strategies for civilian protection*. Geneva: Centre for Humanitarian Dialogue.

Mahony, L. (2013) *Non-military strategies for civilian protection in the DRC*. Brewster, MA: Fieldview Solutions.

Majoor, H., Fisher, M. and Rhametalla, A.H. (2013) *Sudan: an evaluation of WFP's portfolio 2010–2012*. Country Portfolio Evaluation. Rome: WFP.

McAvoy, J. (2005) *East Asia's experience in putting protection into practice*. Oxford: Oxfam GB.

Metzler, J., Kaijuka, R., Vojta, M., Savage, K., Yamano, M., Schafer, A., Yu, G., Ebulu, G. and Ager, A. (2013) *Evaluation of child friendly spaces: Uganda field study report*. New York: Colombia University, Mailman School of Public Health.

Metzler, J., Savage, K., Vojta, M., Yamano, M., Shafer, A. and Ager A. (2013) *Evaluation of child friendly spaces: Ethiopia field study report*. New York: Colombia University, Mailman School of Public Health.

Morras Imas, L. and Rist, R. (2009) *The road to results: designing and conducting effective development evaluations*. Washington, DC: The World Bank.

Moser, C. (2012) 'Mainstreaming women's safety in cities into gender-based policy and programmes'. *Gender & Development* 20(3).

Murray, J. and Landry, J. (2013) Placing protection at the centre of humanitarian action - study on protection funding in complex humanitarian emergencies. Geneva: GPC.

Myer, K., Potts, A. and Roberts, L. (2009) *Grave violations of children's rights and mortality in the Central African Republic – results of a nationwide survey*. Draft. New York: Columbia University, Mailman School of Public Health.

Niland, N. (2013) 'Protection and Instrumentalization: The Contemporary Solferino?', in Donini, A. (ed), *The golden fleece: manipulation and independence in humanitarian action*. Boulder, CO: Kumarian Press.

Niland, N., Holmes, J. and Bradley, M. (2014) 'Policy debate - humanitarian protection in the midst of civil war: lessons from Sri Lanka'. *International Development Policy - Articles and Debates* 5.2. Geneva: The Graduate Institute.

OECD. (2012) *Evaluating peacebuilding activities in settings of conflict and fragility: improving learning for results*. DAC Guidelines and References Series. Paris: OECD.

OECD/DAC (2010) *Evaluating development co-operation: summary of key norms and standards*. 2nd Ed. Paris: OECD/DAC.

OECD/DAC. (1999) *Guidance for evaluating humanitarian assistance in complex emergencies*. Evaluation and Aid Effectiveness No. 1. Paris: OECD/DAC.

OHCHR and UNHCR. (2013) *The protection of human rights in humanitarian crises*. Joint Background Paper. Geneva: IASC.

Oxfam GB. (2004) *Oxfam GB's final report for an ECHO funded project on promoting protection and informed decision making for East Timorese refugees in NTT Province, Indonesia*. Oxfam GB: Oxford.

Oxfam. (2011) *Global Protection Review*. Oxford: Oxfam GB.

Paddon, E. (2011) *Emergency Preparedness and Response in South Central Somalia: Project Review and Analysis*. Copenhagen: DRC.

Paul, D. (1999) *Protection in practice: field-level strategies for protecting civilians from deliberate harm*. RRN Network Paper 30. London: ODI.

Proudlock, K. and Ramalingam, B. with Sandison, P. (2009) 'Improving humanitarian impact assessment: bridging theory and practice'. *8th ALNAP Review of Humanitarian Action*. ALNAP Review. London: ALNAP/ODI.

Puri, J., Aladysheva, A., Iversen, V., Ghorpade, Y. and Brück, T. (2014) What methods may be used in impact evaluations of humanitarian assistance?, *3ie Working Paper 23*. New Delhi: International Initiative for Impact Evaluation (3ie) (*forthcoming*)

Ravallion, M. (2009) 'Evaluating three stylized interventions'. *Designing impact evaluations: different perspectives*. Working Paper 4. New Delhi: 3ie.

Reichhold, U. and Binder, A. (2014) 'The metrics and ethics of protecting civilians', in *Global Responsibility to Protect* [special Issue], 6(2): 206-226.

Reichhold, U., Binder, A. and Niland, N. (2013) *Scoping study: what works in protection and how do we know?* Berlin: GPPi.

Reisman, J., Gienapp, A. and Stachowiak, S. (2007) *A guide to measuring advocacy and policy*. Baltimore: Annie E. Casey Foundation.

Roger, P. (2012) *Introduction to Impact Evaluation*. Impact Evaluation Notes No. 1. Washington, DC: InterAction.

Rogers, P. (2000) 'Causal models in program theory evaluation', in 'Program theory in evaluation: challenges and opportunities' [special issue]. *New Directions for Evaluation*, 2000(87): 47-55.

Rogers, P., Petrosino, A., Heubner, T. and Hacs, T. (2000) 'Program theory evaluation: practice, promise, and problems', in 'Program theory in evaluation: challenges and opportunities' [special issue]. *New Directions for Evaluation*, 2000(87): 5-13.

Rothkegal, S., Poluda, J., Wonani, C., Papy, J., Engelhardt-Wendt, E., Weyermann, B. and Henning, R. (2008) *Evaluation of UNHCR's efforts to prevent and respond to sexual and gender-based violence in situations of forced displacement*. Geneva: UNHCR.

Ruaudel, H. (2011) *Support to emergency, protection and livelihoods in Sri Lanka: project review and analysis*. Copenhagen: DRC.

Sandison, P. (2006) 'The utilisation of evaluations'. *ALNAP Review of Humanitarian Action*. ALNAP Review. London: ALNAP/ODI.

Save the Children. (2012a) *Evaluation of humanitarian action: Somalia crisis response April 2011-April 2012*. London: Save the Children.

Save the Children. (2012b) *Child protection outcome indicators*. Stockholm: Save the Children/Child Protection Initiative.

Save the Children. (2013) *Quality of child protection services: global outcome indicators handbook*. Stockholm: Save the Children/Child Protection Initiative.

Scriven, M. (1991) *Evaluation Thesaurus*. 4th ed. Newbury Park, CA: Sage.

Sheeran, A. (2008) *UNICEF child protection meta-evaluation*. New York: UNICEF.

Slim, H. and Bonwich, A. (2005) *Protection: an ALNAP guide for humanitarian agencies*. London: ALNAP.

Soussan, J. (2008) *MSF and protection: pending or closed?* Paris: MSF.

South, A. and Harragin, S., with Corbett, J., Horsey, R., Kempel, S., Fröjmark, H. and Carstensen, N. (2012) *Local to global protection in Myanmar (Burma), Sudan, South Sudan and Zimbabwe*. HPN Network Paper No. 72. London: HPG/ODI.

Spencer, L., Ritchie, J., Lewis, J. and Dillon, L. (2003) *Quality in qualitative evaluation: a framework for assessing research evidence*. London: Cabinet Office.

Sphere Project. (2011) *Humanitarian charter and minimum standards in humanitarian response: protection principles*. Geneva: The Sphere Project.

Sphere Project. (2012) *2011 edition of the Sphere handbook: What is new?* Version 2. Geneva: The Sphere Project.

Stark, L., Rubenstein, B.L., Muldoon, K. and Roberts, L. (2014) *Guidelines for implementing a national strategy to determine the magnitude and distribution of children outside of family care*. Washington, DC: USAID Center for Excellence on Children in Adversity.

States Parties to the Protocol. (1977) *Protocol additional to the Geneva Conventions of 12 August 1949 and relating to the protection of victims in international armed conflict (Protocol I)*. Geneva.

Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R. and Befani, B. (2012) *Broadening the range of designs and methods for impact evaluations*. Working Paper 38. London: DFID.

Sutton, K., Searle, L. and Lubrano, P. (2012a) *Minimum standards for protection mainstreaming*. Milton Keynes: World Vision UK.

Sutton, K., Searle, L. and Lubrano, P. (2012b) *Minimum inter-agency standards for protection mainstreaming*. Milton Keynes: World Vision UK.

Tsui, T., Hearn, S. and Young, J. (2014) *Monitoring and evaluation of policy influence and advocacy*. Working Paper 395. London: ODI.

UN Secretary General. (2014) *Summary of rights up front*. New York: UN Secretariat.

UNHCR. (2005) *An introduction to international protection. Protecting persons of concern to UNHCR*. Self-Study Module 1. Geneva: UNHCR.

UNHCR. (2006a) *Measuring protection by numbers*. Geneva: UNHCR.

UNHCR. (2006b) *Practical guide to the systematic use of standards & indicators in UNHCR operations*. 2nd ed. Geneva: UNHCR.

UNHCR. (2007) *Handbook for emergencies*. Geneva: UNHCR.

UNHCR. (2010) *High Commissioner's dialogue on protection challenges: protection gaps and responses*. Background Paper. Geneva: UNHCR.

UNHCR. (2013a) *UNHCR global appeal 2014-15 – ensuring protection for people of concern*. Geneva: UNHCR.

UNHCR. (2013b) *Community-based protection. Executive committee of the High Commissioner's programme*. Standing Committee 57th Meeting. Geneva: UNHCR.

UNHCR. (2014) 'Annex A - RFP/2014/594'. Terms of reference for the independent programme evaluation of UNHCR's response to the refugee influx in Lebanon and Jordan. Geneva: UNHCR/PDES.

UNICEF. (2008) *Child Protection Strategy*. New York: ECOSOC.

UNICEF. (2013a) *Global Evaluation Reports Oversight System (GEROS)*. New York: UNICEF.

UNICEF. (2013b) *Information management and innovation to protect children in emergencies*. Project Brief. New York: UNICEF.

Valencia-Ospina, E. (2014) *Seventh report on the protection of persons in the event of disasters*. International Law Commission, Sixty-sixth session. Geneva: International Law Commission.

Vogel, I. (2012) *Review of the use of 'theory of change' in International development*. London: DFID.

Walden, V. (2013) *A quick guide to monitoring, evaluation, accountability and learning in fragile contexts*. Oxford: Oxfam GB.

Walden, V., O'Reilly, M. and Yetter, M. (2007) *Humanitarian programmes and HIV and AIDS*. Oxford: Oxfam GB.

Westhorp, G. (2014) *Realist impact evaluation – an introduction*. London: ODI / Methods Lab.

WFP. (2011) *Mainstreaming protection tools*. Draft version. Rome: WFP.

WFP. (2012) *Integrating protection in WFP operations and programmes - guidance for WFP and partner staff*. Draft version. Rome: WFP.

Wood, A., Apthorpe, R. and Borton, J. (eds) (2001) *Evaluating international humanitarian action: reflections from practitioners*. London: Zed Books.

Woodrow, P. and Oatley, N. (2013) *Practical approaches to theories of change in conflict, security & justice programmes – part I*. London: DFID.

WV. (2011) *World Vision global protection strategy: FY12 – FY15*. Geneva: WV.

Further readings

The following publications can also be accessed via the Humanitarian Evaluation and Learning Portal (HELP): <http://www.alnap.org/evaluating-protection>.

ACT Alliance. (2010) *Humanitarian protection policy for the Act Alliance*. Geneva: ACT Alliance.

AusAID. (2013) *Protection in Humanitarian Action Framework*. Canberra: AusAID.

Brown, M., and Perschler, V. (2013) *Evaluation of UNICEF programmes to protect children in emergencies*. Synthesis Report. New York: UNICEF.

Cohen, R. and Bradley, M. (2010) 'Disasters and displacement: gaps in protection'. *Journal of International Humanitarian Legal Studies*, Vol. 1.

Contat Hickel, M. (2001) 'Protection of internally displaced persons affected by armed conflict: concept and challenges'. *International Review of the Red Cross*, 83(843): 699-711.

Darcy, J. and Hofmann, C.A. (2003) *According to need? Needs assessment and decision-making in the humanitarian sector*. HPG Report No. 15. London: HPG/ODI.

Deschamp, B., Azorbo, M. and Lohse, S. (2010) *Earth, wind & fire: a review of UNHCR's role in recent natural disasters*. Geneva: UNHCR/PDES.

ECHO. (2009) Humanitarian protection: DG ECHO's funding guidelines. Brussels: ECHO.

Hughes, S. (2012) *Human rights in situations of conflict: supporting the capacity of civil society in the Democratic Republic of the Congo*. GB Oxford: Oxfam.

IASC. (2005a) *Cluster working group on protection progress report, annex 2: responsibility-sharing for the enhancement of protection in humanitarian emergencies*. Geneva: IASC.

IASC. (2005b) *Guidelines for gender based violence in humanitarian settings: focusing on prevention and response to sexual violence in emergencies*. Geneva: IASC.

Jaspars S., O'Callaghan, S. and Stites, E. (2007) *Linking livelihoods and protection: A preliminary analysis based on a review of the literature and agency practice*. HPG Working Paper. London: HPG/ODI.

Jaspars, S. (2009) *Livelihoods and protection in Sri Lanka: a review of the Danish Refugee Council's integrated livelihoods, protection and emergency response programme in Sri Lanka*. HPG Report. London: HPG/ODI.

Jaspars, S. and O'Callaghan, S. (2008) *Challenging choices: protection and livelihoods in Darfur: a review of DRC's programme in Eastern West Darfur*. HPG Working Paper. London: HPG/ODI.

Mahony, L. and Nash, R. (2010) *Strengthening the protection impact of human rights field presences – concept paper for the OHCHR annual meeting of field presences*. Brewster, MA: Fieldview Solutions.

Mahony, L. and Nash, R. (2012) *Influence on the ground: understanding and strengthening the protection impact of United Nations human rights field presences*. Brewster, MA: Fieldview Solutions.

O'Callaghan, S., Jaspars, S. and Pavanello, S. (2009) *Losing ground: protection and livelihoods in the Occupied Palestinian Territory*. HPG Working Paper. London: HPG/ODI.

OCHA. (2009) *OCHA on message: protection*. Vol. 1. New York: OCHA.

Save the Children. (2008) *A common responsibility: The role of community-based child protection groups in protecting children from sexual abuse and exploitation*. London: Save the Children.

Swales, D. and McMillan, N. (2009) *Manual for the measurement of indicators for children in formal care*. New York: UNICEF/BCN.

UN Secretary General. (2013a) 'Renewing our commitment to the peoples and purposes of the United Nations'. UN SG Speech, 22 November, New York.

UNHCR. (2013) *Note on the mandate of the High Commissioner for refugees and his office*. Geneva: UNHCR - Division of International Protection.

Wulczyn F., Daro, D., Fluke, J., Feldman, S. Glodek, C. and Lifanda, K. (2010) *Adapting a systems approach to child protection: key concepts and considerations*. Working Paper. New York: UNICEF.



ALNAP would like to acknowledge the financial support of Irish Aid and USAID in carrying out this initiative.



Related ALNAP publications

Previous ALNAP reports on protection

Protection - An ALNAP guide for humanitarian agencies

Other ALNAP publications on evaluation

Evaluating humanitarian action

Evaluation de l'action humanitaire

Evaluación de la acción humanitaria

www.alnap.org



ALNAP
Overseas Development Institute
203 Blackfriars Road
London SE1 8NJ
United Kingdom
T + 44 (0)20 7922 0388
F + 44 (0)20 7922 0399
E alnap@alnap.org