

**Evaluating Humanitarian Protection: A Protection-Specific Evaluation Quality Assessment
Framework**

by

Matthew Jacob Pribis

Bachelor of Science, Old Dominion University, 2015

Master of International Development, University of Pittsburgh, 2021

Submitted to the Graduate Faculty of the
Graduate School of Public and International Affairs in partial fulfillment
of the requirements for the degree of
Master of International Development

University of Pittsburgh

2021

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC AND INTERNATIONAL AFFAIRS

This thesis was presented

by

Matthew Jacob Pribis

It was defended on

April 26, 2021

and approved by

Dr. Taylor Seybolt, Associate Professor

Dr. Gemma Marolda, Adjunct Instructor

Dr. Michael Poznansky, Assistant Professor

Thesis Advisor: Dr. Taylor Seybolt, Associate Professor

Copyright © by Matthew Jacob Pribis

2021

Evaluating Humanitarian Protection: A Protection-Specific Evaluation Quality Assessment Framework

Matthew Jacob Pribis, MID

University of Pittsburgh, 2021

In 2021, over 80 million people are forcibly displaced from their homes.¹ 177 million people need US\$ 28.8 billion worth of humanitarian assistance to meet their basic needs.² The international humanitarian system (IHS) exists to respond to this massive need when states or other duty-bearers are unable and/or unwilling. The gargantuan task of responding to human suffering on a global scale is one that the IHS struggles valiantly to conduct, even though the demand for aid outweighs its supply. Donor fatigue, increased needs, and the COVID-19 pandemic have created a mismatch between the availability of resources and the funding requirements of the IHS. This scarcity has forced the IHS to investigate how funds can be allocated most efficiently to make the greatest impact. Thus, the IHS is undergoing a movement of accountability reforms involving staff professionalization and the use of evidence-based practice.³ An important aspect of this movement is the application of evaluation to make informed judgements about the value of interventions and their impact, as well as how to improve them.⁴ Evaluation of humanitarian

¹ United Nations High Commissioner for Refugees, “Refugee Statistics,” accessed April 19, 2021, <https://www.unhcr.org/refugee-statistics/>.

² “Global Humanitarian Overview 2020,” unocha.org (OCHA, December 3, 2019), <https://www.unocha.org/global-humanitarian-overview-2020>.

³ See Jyotsna Puri et al., 2015, where humanitarian performance is divided into three main categories: **accuracy** (whether assistance is reaching the right people at the right time), **effectiveness** (whether assistance is bringing about the desired changes in the target population’s lives) and **efficiency** (whether the assistance is being delivered in the right doses and right ways, and with manageable costs).

⁴ J Cosgrave, M Buchanan-Smith, and A Warner, “Evaluation of Humanitarian Action Guide,” ALNAP, October 10, 2016, <https://www.alnap.org/help-library/evaluation-of-humanitarian-action-guide>.

programming contributes to a body of evidence that establishes “what works” and what does not.⁵ This research reviews current humanitarian protection literature to investigate what criteria should be used when assessing the quality of protection-specific evaluations and then applies a novel protection-specific evaluation quality assessment framework to ten protection-specific evaluation reports, finding that only five out of ten selected reports had satisfactory quality based on Global Evaluation Report System (GEROS) scoring metrics. As a result, decision makers may not know what works and what does not in humanitarian protection and should be cautious when using evaluation findings.

⁵ Cosgrave et al., “EHA Guide” 2016.

Table of Contents

1.0 Background	1
1.1 The International Humanitarian System	1
1.2 Humanitarian Protection.....	6
1.3 Evaluation of Humanitarian Protection	12
2.0 Research Approach and Methodology.....	16
2.1 Purpose and Focus.....	16
2.2 Research Questions	16
2.3 Process	17
2.3.1 Information Collection	17
2.3.2 Creating a Protection-Specific Evaluation Quality Assessment Framework	20
2.3.3 Quality Assessment of the Evaluation Reports	26
2.3.4 Calculating a Quality Meta-Score for the Selected Reports	27
2.4 Limitations	29
3.0 Findings.....	31
3.1 Quality Level of the Selected Evaluation Reports	31
3.2 Theories of Change Found in Selected Reports.....	33
3.3 Attribution of Protection Outcomes	36
4.0 Conclusion	37
5.0 Recommendations	40
Appendix A Evaluation Report Selection	44

Appendix B Figure References	46
Bibliography	47

List of Tables

Table 1: Evaluation Matrix.....	19
Table 2: Protection-Specific Evaluation Quality Assessment Framework.....	26
Table 3: The GEROS Quality Rating Scale	27
Table 4: Evaluation Report Scores.....	32
Table 5: Protection Theories of Change Identified in Selection	34

List of Figures

Figure 1: The IHS in Numbers by Personnel, Spending and Agency Type, 2017.	3
Figure 2: IHS Coordination Mechanisms.....	5
Figure 3: The Cluster Approach	6
Figure 4: The Global Humanitarian Architecture	11
Figure 5: OECD Evaluation Criteria	13
Figure 6: Quality Scores of Selected Reports	33

Acronyms and Abbreviations

3ie: International Initiative for Impact Evaluation

ALNAP: Active Learning Network for Accountability and Performance

COOPI: Cooperazione Internazionale

EHA: Evaluation of Humanitarian Action

EPHA: Evaluation of Protection in Humanitarian Action

FAO: Food and Agricultural Organization of the United Nations

GEROS: Global Evaluation Report Oversight System

GPA: Global Protection Architecture

GPC: Global Protection Cluster

HELP: Humanitarian Evaluation Learning and Performance

HPS: Humanitarian Protection Sector

IASC: Inter Agency Standing Committee

ICRC: International Committee of the Red Cross

IDP: Internally Displaced Person

IFRC: International Federation of the Red Cross

IHS: International Humanitarian System

IOM: International Organization for Migration

IRC: International Rescue Committee

IRC: International Rescue Committee

JPAL: Abdul Latif Jameel Poverty Action Lab

NGO: Non-Governmental Organization

NRC: Norwegian Refugee Council

PREDES: Centro de Estudios y Prevención de Desastres.

PSEQAF: Protection-Specific Evaluation Quality Assessment Framework

RPWG: Refugee Protection Working Group

UN: United Nations

UNDP: United Nations Development Program

UNEG: United Nations Evaluation Group

UNFPA: United Nations Population Fund

UNHCR: United Nations High Commission for Refugees, or the UN Refugee Council

UNICEF: United Nations International Children's Emergency Fund

USAID: United States Agency for International Development

WFP: World Food Program

WHO: World Health Organization

1.0 Background

1.1 The International Humanitarian System

Humanitarian action comprises a wide range of interventions whose objectives are to save lives, alleviate suffering, and maintain human dignity during and after emergencies such as war, natural disaster, pandemic or some combination of the aforementioned.⁶ It is guided by the principles of humanity, impartiality, neutrality and independence, although these principals arguably exist more in terms of intention than operational reality, especially in the context of armed conflict.⁷ The international humanitarian system (IHS), as it exists today, “originates from the European experience of war and natural disaster”, although the humanitarian imperative of acting to alleviate the suffering of others is a timeless and global gesture.⁸ The birth of the modern IHS stems from 1863, when Henry Dunant, a Swiss businessman, witnessed the terrible carnage of suffering and wounded soldiers in the aftermath of the battle of Solferino, Italy.⁹ Dunant was so moved by what he saw that he created the Red Cross, an international system for nations to organize relief societies that provide care for persons wounded in war. Dunant later facilitated the first of the Geneva Conventions, international treaties that limited the savagery and cruelty of war

⁶ “GHD Principles and Good Practice,” Good Humanitarian Donorship, accessed July 28, 2020, <https://www.ghdinitiative.org/ghd/gns/principles-good-practice-of-ghd/principles-good-practice-ghd.html>.

⁷ Taylor Seybolt, “The Myth of Neutrality,” *Peace Review* 8, no. 4 (December 1996), <https://doi.org/10.1080/10402659608426005>.

⁸ Eleanor Davey, John Borton, and Matthew Foley, “A History of the Humanitarian System: Western Origins and Foundations,” Humanitarian Policy Group (Overseas Development Institute, June 1, 2013), <https://www.odi.org/publications/7535-history-humanitarian-system-western-origins-and-foundations>.

⁹ “The Nobel Peace Prize 1901,” NobelPrize.org, accessed July 28, 2020, <https://www.nobelprize.org/prizes/peace/1901/dunant/biographical/>.

and gave birth to International Humanitarian Law.¹⁰ Subsequently, the end of the Second World War marked a turning point in the history of the IHS, as the United Nations (UN) was born, along with specialized humanitarian agencies like the United Nations High Commissioner for Refugees (UNHCR).¹¹ UNHCR was created to help the millions of European refugees who were displaced during World War II, and it is still today one of the most influential actors in the IHS. Broadly, the IHS is comprised of states, non-governmental organizations (NGOs), international agencies, faith-based organizations, and the Red Cross/Red Crescent movement, who are all connected with one another at varying levels of collaboration, complementarity, and competition.¹² Figure 1 below displays the IHS in terms of the three largest categories of actors (UN agencies, NGOs, and the Red Cross/Red Crescent Movement), their number of personnel and expenditures.

¹⁰ “The Nobel Peace Prize 1901”.

¹¹ Randolph C. Kent, *The Anatomy of Disaster Relief* (London: Pinter, 1987).

¹² Davey et al., “A History of the Humanitarian System” 2013.

	UN agencies*	NGOs (estimates)	Red Cross/Red Crescent Movement**
Field personnel	79,000	331,000	159,700
569,700	• 68,000 nationals • 11,000 internationals	• 304,000 nationals • 27,000 internationals	• 14,000 ICRC/IFRC, nationals • 2,700 ICRC/IFRC, internationals • 143,000 RCS, nationals
Humanitarian expenditure (not cumulative)***	\$16 billion	\$16.8 billion	\$15.7 billion

Figure 1: The IHS in Numbers by Personnel, Spending and Agency Type, 2017.

Source: Knox, Clarke. “The State of the Humanitarian System 2018 - Full Report.” SOHS.ALNAP.org. ALNAP, December 5, 2018. <https://sohs.alnap.org/help-library/the-state-of-the-humanitarian-system-2018-full-report>.

Since the passing of UN General Assembly resolution 46/182 in December of 1991 and the 2005 Humanitarian Reform Agenda, the IHS is organized into and coordinates via a Cluster Approach, where clusters are groups of humanitarian organizations, both UN and non-UN, who have specific responsibilities in each of the main sectors of humanitarian action.¹³ The current clusters include camp coordination and camp management (led by the International Organization for Migration (IOM) and UNHCR), early recovery (led by the United Nations Development Program (UNDP)), education (led by the United Nations International Children’s Emergency Fund (UNICEF) and Save the Children), emergency telecommunications (led by World Food Program

¹³ “What Is the Cluster Approach?” Humanitarian Response (UN OCHA) accessed July 28, 2020, <https://www.humanitarianresponse.info/en/about-clusters/what-is-the-cluster-approach>.

(WFP)), food security (led by WFP and Food and Agricultural Organization (FAO)), health (led by the World Health Organization (WHO)), logistics (led by WFP), nutrition (led by UNICEF), protection (led by UNHCR), shelter (led by the International Federation of the Red Cross (IFRC) and UNHCR), and Water, Sanitation and Hygiene (WASH) (led by UNICEF).¹⁴

The IHS is further divided into categories based on the legal status of those receiving assistance: refugee or non-refugee. The Cluster Approach applies to humanitarian assistance given to non-refugee populations, such as internally displaced persons (IDPs), and the Refugee System (also referred to as the Refugee Coordination Model), applies to those with refugee-status.¹⁵ For the purposes of this paper, when the IHS is referred to, both the Cluster Approach and Refugee Systems are included. This is because interventions associated with both systems often interface in so called “mixed situations” where refugees, host communities, IDPs and populations affected by natural and/or man-made disasters are all being assisted alongside one another in the same geographic area.¹⁶ When necessary and possible, distinctions will be made to denote whether a program or evaluation applies to the Cluster Approach or the Refugee system. Figure 2 displays the organization of the IHS coordination mechanisms, including the associated Development System, while Figure 3 displays the Cluster Approach.

¹⁴ “What Is the Cluster Approach?” Humanitarian Response (UN OCHA).

¹⁵ “Emergency Handbook,” unhcr.org (UNHCR), accessed April 19, 2021, <https://emergency.unhcr.org/entry/41813/cluster-approach-iasc>.

¹⁶ “UNHCR Refugee Response Coordination Model,” unhcr.org (UNHCR), accessed April 19, 2021, <https://www.unhcr.org/53679e2c9.pdf>.

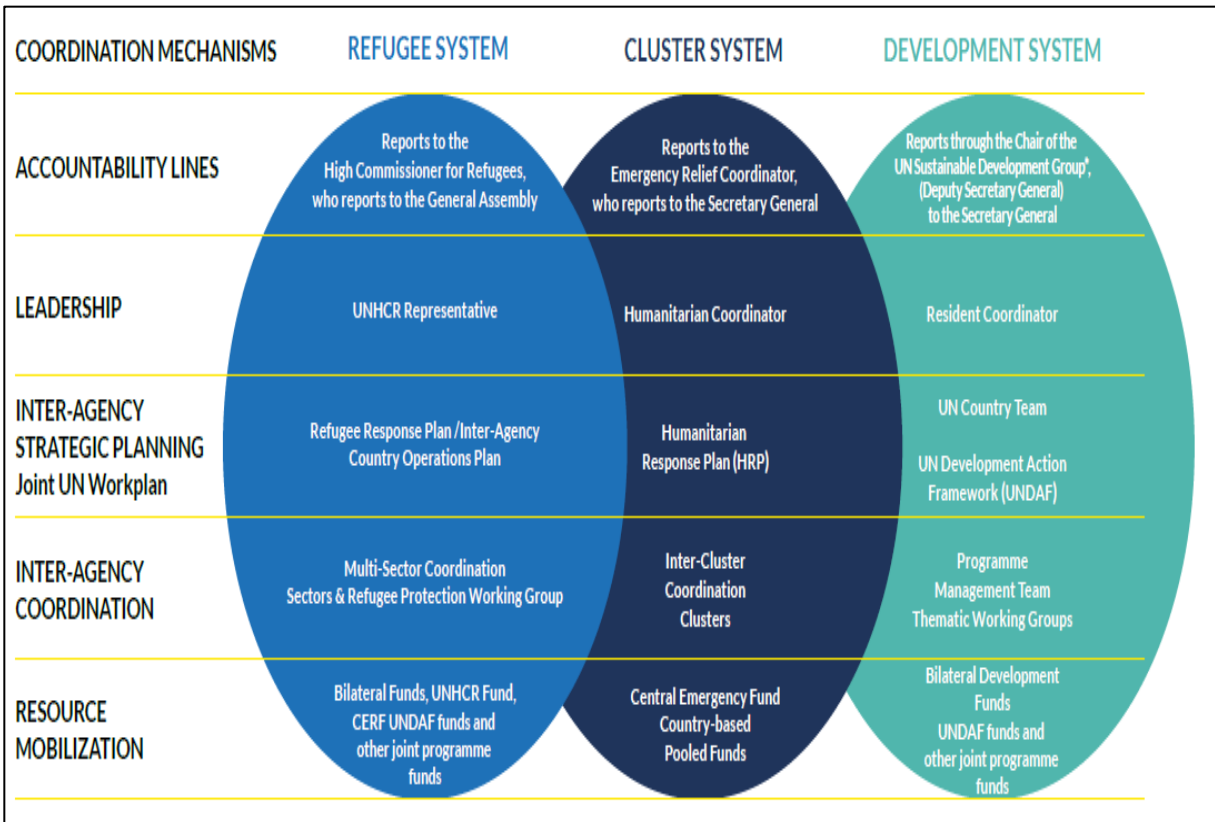


Figure 2: IHS Coordination Mechanisms

Source: “Refugee Coordination Model (RCM) - UNHCR.” [unhcr.org](https://data2.unhcr.org/en/documents/download/65407). UNHCR. Accessed April 19, 2021. <https://data2.unhcr.org/en/documents/download/65407>.

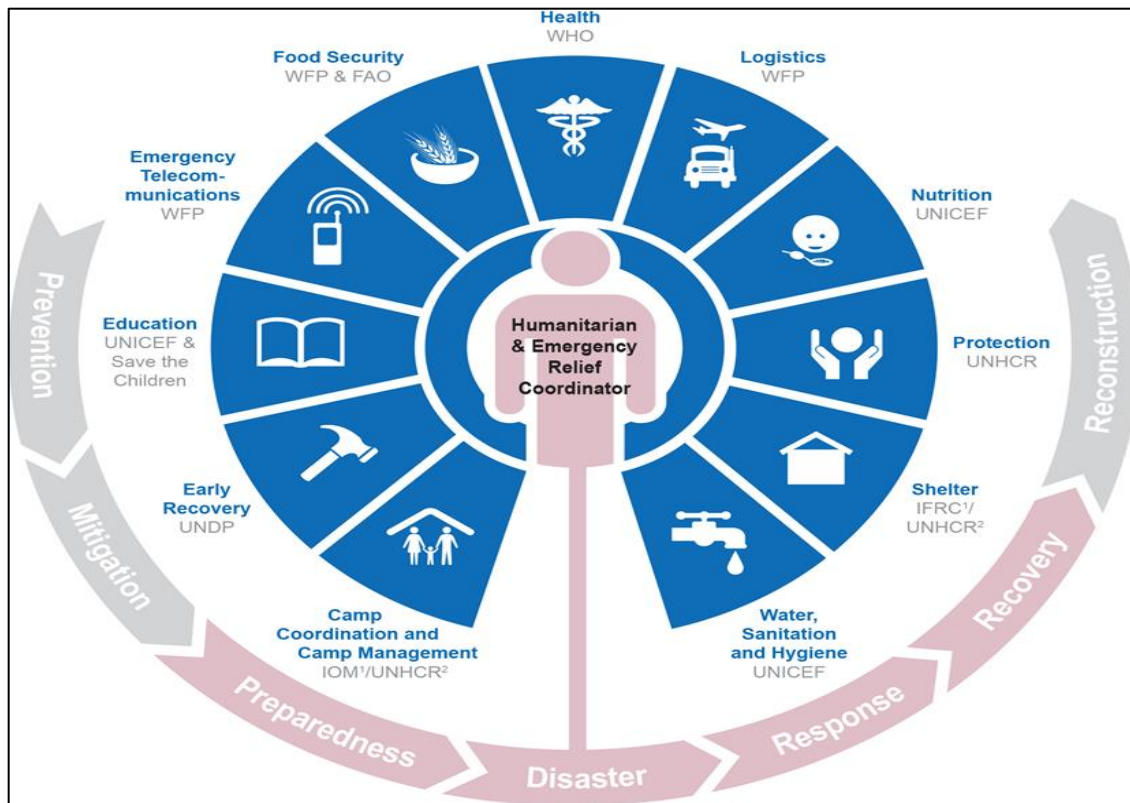


Figure 3: The Cluster Approach

Source: “Emergency Handbook.” unhcr.org. UNHCR. Accessed April 19, 2021.

<https://emergency.unhcr.org/entry/61190/cluster-approach-iasc>.

1.2 Humanitarian Protection

Humanitarian protection refers to a broad range of activities aimed at obtaining full respect for the rights of individuals (recipients of humanitarian aid) in accordance with the letter and spirit of the relevant bodies of law (International Human Rights, Humanitarian and Refugee law).¹⁷ This

¹⁷ “IASC Policy on Protection in Humanitarian Action, 2016,” interagencystandingcommittee.org (IASC, October 14, 2016), <https://interagencystandingcommittee.org/iasc-protection-priority-global-protection-cluster/iasc-policy-protection-humanitarian-action-2016>.

comprehensive definition comes from the Inter Agency Standing Committee (IASC), a high-level and authoritative humanitarian coordination forum comprised of 18 United Nations (UN) and non-UN organizations. This definition is intentionally broad to encompass all the different strategies and methods that different actors employ to achieve protection.¹⁸ Still, definitions, perceptions and operational realities of protection vary greatly among different stakeholders within the IHS, and this constitutes both a source of debate and confusion in the sector, as almost all humanitarian action influences protection outcomes. It becomes even more difficult to narrow the definition of protection considering recent protection mainstreaming policies within the IHS which assert that protection is central to the purpose of all humanitarian response (the “centrality” of protection). For example, the IASC states that “[A]ll humanitarian actors share responsibility for ensuring that activities ... are carried out with a protection lens and ... do not lead to or perpetuate discrimination, abuse, violence, neglect or exploitation.”¹⁹ This demonstrates the broader principle of “do no harm” that humanitarians are increasingly held accountable to after many unintended consequences. Placing protection at the center of humanitarian action is a good thing because in theory it reduces risks for beneficiaries, but at the same time, over-applying the label of protection to all aspects of humanitarian response creates a confusing situation where it is not clear who the protection experts are and what they are doing. Definitional confusion and debate aside, protection is fundamentally about reducing risk, and for the purposes of this paper, protection refers to interventions whose programmatic objective is to address and reduce risk for recipients of humanitarian aid.

¹⁸ “IASC Policy on Protection in Humanitarian Action, 2016”

¹⁹ “Cluster Working Group on Protection, Progress Report and Supporting Documents,” interagencystandingcommittee.org (IASC, December 9, 2005), <https://interagencystandingcommittee.org/working-group/content/cluster-working-group-protection-progress-report-and-supporting-documents>.

When it comes to protection, states bear the primary duty to protect and uphold the rights of civilians under their control, and humanitarians only intervene when primary duty bearers are unwilling or unable. Nonetheless, humanitarian protection actors cannot completely physically protect civilians from armed violence or other types of harm associated with crisis and disaster.²⁰ This was made evident in the 1996 evaluation of the international response to the Rwandan genocide, which showcased how humanitarian action is not a substitute for politico-military action.²¹ Another more obvious example of this is that protection actors cannot prevent natural disasters like hurricanes or earthquakes, but they can help build resilience and capacity among populations to control damage when these events happen. Humanitarian protection is consequently more about reducing risk to the maximum extent possible than it is about ensuring physical security. This is logical because most events that put populations at risk are outside humanitarians' control. Therefore, in practice, protection boils down to the prevention, mitigation and ending of actual and potential risks. This is facilitated by various types of protection intervention models such as providing remedy to victims of harm (remedial action), reducing risk exposure (prevention), increasing resilience (environment-building) and attempting to change harmful behaviors of primary duty bearers (persuasion).²² Real-world examples include clearing unexploded ordinance after war, issuing personal documentation to refugees, counselling survivors of sexual violence, and reunification of separated families. Protection can be further categorized as specialized protection actions (known as vertical or stand-alone), integrated protection activities

²⁰ Urban Reichhold and Andrea Binder, "Scoping Study: What Works in Protection and How Do We Know?," gppi.net (The Global Public Policy Institute, March 1, 2013), <https://www.gppi.net/2013/03/01/scoping-study-what-works-in-protection-and-how-do-we-know>.

²¹ Anne Mackintosh, "International Response to Conflict and Genocide: Lessons from the Rwanda Experience," ed. David Millwood, academic.oup.com (Oxford University Press, September 1, 1996), <https://academic.oup.com/jrs/article-abstract/9/3/334/1553872?redirectedFrom=fulltext>.

²² This is a combination of the protection interventions laid out in the GPPi Scoping Study and the SPHERE Handbook

(the integration of protection into sectoral and multi-sectoral interventions) and protection mainstreaming (the process of incorporating protection principles and promoting meaningful access, safety and dignity in humanitarian aid).^{23 24} Needless to say, protection manifests itself in many different ways, although they are collectively united by one primary objective: to reduce risk to populations of concern. Risk is best conceptualized by the following equation:²⁵

$$Risk = \frac{Threats \times Vulnerabilities}{Capacities}$$
 This equation illustrates how protection risks are directly

proportional to: 1) the level and nature of the threat; 2) the vulnerabilities of affected persons, while inversely correlated to: 1) capacities to cope with the threat.²⁶

Within the IHS, the humanitarian protection sector (HPS) is comprised of the Global Protection Cluster (GPC) and the Refugee Protection Working Group (RPWG), both of which are led by UNHCR who has a specific and internationally recognized protection mandate. The GPC primarily deals with non-refugee populations while the RPWG deals with refugees. The GPC is comprised of five National Protection Clusters (Africa, Asia and the Pacific, Americas, Europe and the Middle East), as well as a Strategic Advisory Group, a Task Team, Areas of Responsibility leaders, and an Operations Cell.²⁷ The GPC areas of responsibility subclusters are child protection, gender-based violence, housing land and property, and mine action. It is supported by specialized humanitarian protection organizations who also have mandates for protection such as the

²³ J Cosgrave, M Buchanan-Smith, and A Warner, “Evaluation of Humanitarian Action Guide,” ALNAP, October 10, 2016, <https://www.alnap.org/help-library/evaluation-of-humanitarian-action-guide>.

²⁴ “Protection Mainstreaming,” [globalprotectioncluster.org](https://www.globalprotectioncluster.org) (Global Protection Cluster), accessed April 27, 2021, <https://www.globalprotectioncluster.org/themes/protection-mainstreaming/>.

²⁵ “Humanitarian Protection: Improving Protection Outcomes to Reduce Risks for People in Humanitarian Crises,” ec.europa.eu (European Commission, September 6, 2020), https://ec.europa.eu/echo/what/humanitarian-aid/protection_en.

²⁶ “Humanitarian Protection: Improving Protection Outcomes”

²⁷ “WHO WE ARE,” [globalprotectioncluster.org](https://www.globalprotectioncluster.org) (Global Protection Cluster), accessed April 19, 2021, <https://www.globalprotectioncluster.org/about-us/who-we-are/>.

International Committee of the Red Cross (ICRC), UNICEF, the United Nations Population Fund (UNFPA), UNMAS, the International Rescue Committee (IRC), Care, Oxfam and more. The GPC coordinates the international protection response for Internally Displaced Persons (IDPs) and attempts to ensure that humanitarian assistance generally yields protection outcomes.²⁸ The RPWG operates under the Refugee Coordination Model and is responsible for the coordination and mainstreaming of refugee protection services throughout all operational sectors.²⁹ Like the Cluster System and the Refugee Coordination Model, the distinction between protection interventions carried out in the GPC and the RPWG can be ambiguous due to the high-level of interface between the two systems during “mixed situations”. Additionally, some actors conducting humanitarian protection work do not fall under these two formal coordination systems, such as UN Peacekeeping programs, human rights organizations, and smaller NGOs. The multitude of protection actors is extensive, ranging from state authorities to human rights groups. For the purposes of this paper, when the HPS is referred to, both the GPC, the RPWG and other relevant humanitarian protection actors are referred to. When necessary and possible, distinctions will be made to denote what system a program or evaluation applies to. Figure 4 gives an overview of the “Global Protection Architecture” (GPA) encompassing the HPS and the additional actors having a role in protection globally.³⁰ This research focuses on the HPS within the larger GPA.

²⁸ “The Protection of Human Rights in Humanitarian Crises - A Joint Background Paper by OHCHR and UNHCR,” ALNAP.org (IASC, May 8, 2013), <https://www.alnap.org/help-library/the-protection-of-human-rights-in-humanitarian-crises-a-joint-background-paper-by-ohchr>.

²⁹ “UNHCR Refugee Response Coordination Model,” unhcr.org (UNHCR), accessed April 19, 2021, <https://www.unhcr.org/53679e2c9.pdf>.

³⁰ Urban Reichhold and Andrea Binder, “Scoping Study: What Works in Protection and How Do We Know?,” gppi.net (The Global Public Policy Institute, March 1, 2013), <https://www.gppi.net/2013/03/01/scoping-study-what-works-in-protection-and-how-do-we-know>.

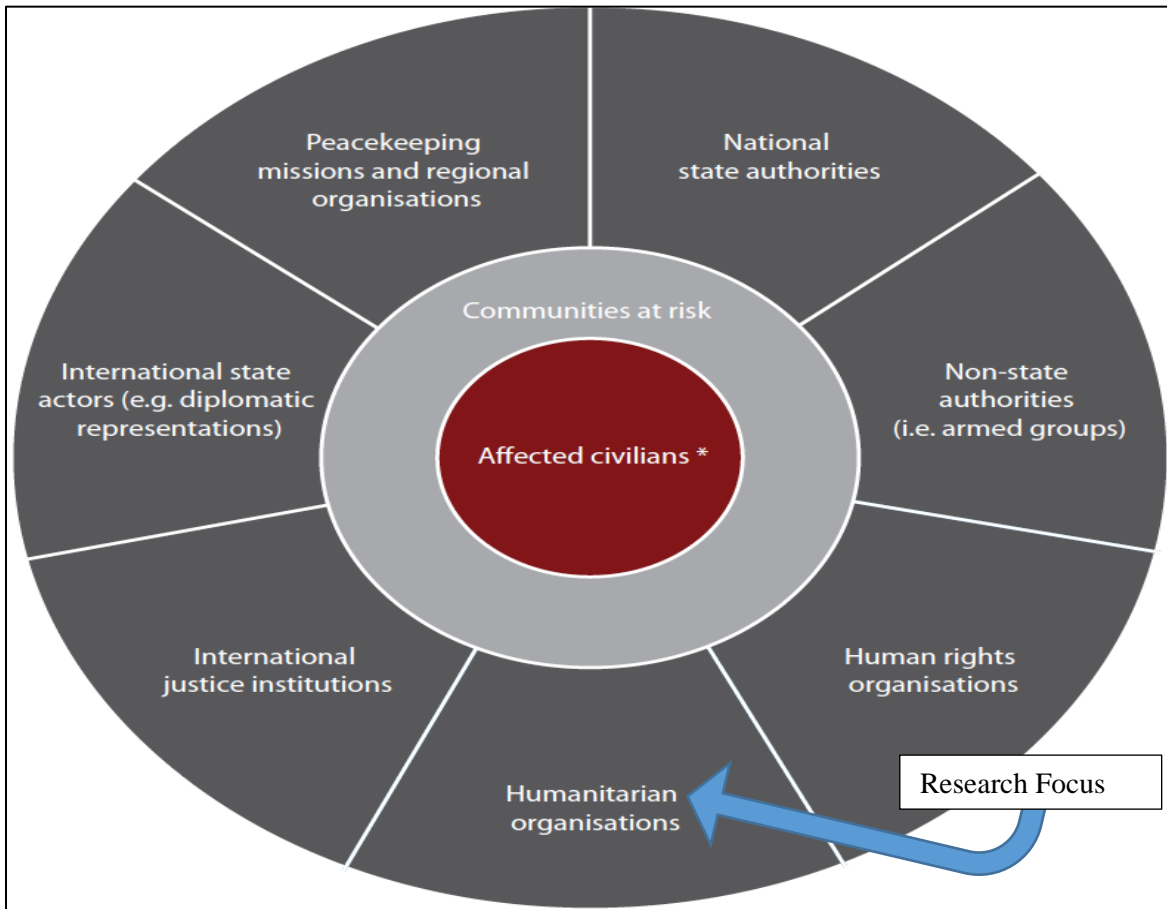


Figure 4: The Global Humanitarian Architecture

Source: Reichhold, Urban, and Andrea Binder. “Scoping Study: What Works in Protection and How Do We Know?” gppi.net. The Global Public Policy Institute, March 1, 2013.

<https://www.gppi.net/2013/03/01/scoping-study-what-works-in-protection-and-how-do-we-know>.

Like the rest of the IHS, the Humanitarian Protection Sector (HPS) is under increasing pressure from donors and aid recipients for greater staff professionalization, accountability, and evidence-based practice. Historically, protection has not been given the degree of attention or funding that more tangible humanitarian activities such as material assistance receive.³¹ Increasingly, however, the humanitarian community is placing greater emphasis on both the

³¹ J Cosgrave, M Buchanan-Smith, and A Warner, “Evaluation of Humanitarian Action Guide,” ALNAP, October 10, 2016, <https://www.alnap.org/help-library/evaluation-of-humanitarian-action-guide>.

importance and shared responsibility of protection by all actors involved in humanitarian assistance, as noted by the 2012 World Vision Minimum Standards for Protection Mainstreaming, the 2013 IASC Policy on the Centrality of Protection, the 2016 IASC Policy on System Wide Approaches to Protection, and the Global Protection Cluster’s Protection Mainstreaming Toolkit.³² This growing body of policies, regulations and norms surrounding protection shows how it is increasingly perceived to be at the core of all humanitarian action, as a central outcome that all humanitarian organizations must aim for. Additionally, the protection sector has made strides towards professionalizing the work standards of staff, as seen in the 2009, 2013 and 2018 ICRC Professional Standards for Protection Work, the Sphere Handbook Protection chapter, and the Core Humanitarian Standards on Quality and Accountability’s Protection Mainstreaming Imperative.³³

1.3 Evaluation of Humanitarian Protection

Evaluation, according to Owens, 2007, is the production of knowledge based on systematic inquiry to assist decision-making about a program.³⁴ While research can be viewed as the process of proving something, evaluation is the process of improving something. Evaluation, when applied to humanitarian protection programming, can facilitate informed judgements on the value of

³² “Evaluation of Humanitarian Action Guide,” 2016.

³³ Christoplos, Dillon, and Bonino, “ALNAP Guide: Evaluation of Protection in Humanitarian Action,” ALNAP, October 25, 2018, <https://www.alnap.org/help-library/alnap-guide-evaluation-of-protection-in-humanitarian-action>.

³⁴ John M. Owen and C. Alkin, *Program Evaluation: Forms and Approaches* (New York: The Guilford Press, 2007).

interventions and their results.³⁵ Evaluation can answer questions of effectiveness (is the intervention achieving its objectives?), coherence (how well does the intervention fit?), sustainability (will the benefits last?), efficiency (how well are resources used?), relevance (is the intervention doing the right things?), and most importantly, impact (is the intervention achieving its objectives).³⁶ Figure 5 below displays the most widely-used humanitarian and development evaluation criteria from the Organization for Economic Cooperation and Development (OECD).



Figure 5: OECD Evaluation Criteria

Source: “Evaluation Criteria.” oecd.org. OECD. Accessed April 19, 2021.

<https://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm#:~:text=The%20OECD%20DAC%20Network%20on,two%20principles%20for%20their%20use.>

³⁵ Cosgrave, Buchanan-Smith, and Warner, “Evaluation of Humanitarian Action Guide,” ALNAP, October 10, 2016, <https://www.alnap.org/help-library/evaluation-of-humanitarian-action-guide>.

³⁶ “Evaluation Criteria,” oecd.org (OECD), accessed April 19, 2021, <https://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm#:~:text=The%20OECD%20DAC%20Network%20on,two%20principles%20for%20their%20use.>

High-quality evaluations are evaluations that “help agencies and stakeholders identify successful programs to expand or pitfalls to avoid.”³⁷ They can arm humanitarians with the necessary tools to build a body of evidence and subsequently make evidence-based judgements upon the “degree to which a program was successful, and the nature of the success.”³⁸ This understanding of what works, what does not work, and why, is paramount to the protection sector fulfilling its mandate of preventing the harm of populations affected by crisis. Mistakes in this field constitute greater risk that can lead to suffering and death. Still, despite the importance and relevance of evaluation to improve, justify and measure the impact of programs, the field of humanitarian protection evaluation is lacking and there is a startling gap of evidence and data for what works and what does not.³⁹ This means that the billions of dollars spent on humanitarian protection programming every year is not backed by hard evidence, and that it is potentially squandered. According to Reichhold and Binder, 2013, this is because “there is a perception among both program 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.”⁴⁰ This is in contrast to existing high-quality evaluation practices and associated evidence in the humanitarian sectors of health, nutrition and peacebuilding.⁴¹ Generally, evaluation guidance focused on protection in humanitarian action

³⁷ Kerry Bruce, “What Is Quality in an Evaluation?” socialimpact.com (Social Impact, May 24, 2017), <https://socialimpact.com/what-is-quality-in-an-evaluation/>.

³⁸ Cosgrave, Buchanan-Smith, and Warner, “Evaluation of Humanitarian Action Guide,” ALNAP, October 10, 2016, <https://www.alnap.org/help-library/evaluation-of-humanitarian-action-guide>.

³⁹ “Evaluation of Humanitarian Action Guide,” 2016.

⁴⁰ Urban Reichhold and Andrea Binder, “Scoping Study: What Works in Protection and How Do We Know?,” gppi.net (The Global Public Policy Institute, March 1, 2013), <https://www.gppi.net/2013/03/01/scoping-study-what-works-in-protection-and-how-do-we-know>.

⁴¹ Jyotsna Puri et al., “What Methods May Be Used In Impact Evaluations of Humanitarian Assistance” iza.org, January 2015, <https://www.iza.org/publications/dp/8755/what-methods-may-be-used-in-impact-evaluations-of-humanitarian-assistance>.

is limited, fragmented and confined to specific programming manuals.⁴² This is especially problematic, given that the way in which a protection program is designed and implemented determines whether it will put people at greater risk or protect them from harm.⁴³ There is thus a need to assess and increase the quality of humanitarian protection evaluations so that the HPS can understand more fully what works in protection and what does not.

⁴² Christoplos, Dillon, and Bonino, “ALNAP Guide: Evaluation of Protection in Humanitarian Action,” ALNAP, October 25, 2018, <https://www.alnap.org/help-library/alnap-guide-evaluation-of-protection-in-humanitarian-action>.

⁴³ “ALNAP Guide: Evaluation of Protection in Humanitarian Action,” 2018.

2.0 Research Approach and Methodology

2.1 Purpose and Focus

The purpose of this research was to review current evaluative practice within the humanitarian protection sector (HPS) of the IHS to determine what criteria are most important when assessing the quality of protection-specific evaluation reports. Once the criteria were determined, a selection of ten different evaluation reports were assessed for quality to see how they performed in relation to the criteria. This research focused specifically on a narrow aspect of the larger IHS and the GPA: evaluation of protection-specific humanitarian interventions. It acknowledges the influence that human rights actors, development actors, peacekeeping/peacebuilding actors, as well as other political and military actors have on protection, but they are not the primary focus of this research. Understanding the quality level of the humanitarian protection evaluation reports selected for this research is part of the broader research focus of this thesis, which seeks to discover how “fit” the IHS is to evaluate humanitarian protection in general. Fit here means capable of evaluating humanitarian protection in such a way that existing evaluations are of high quality and decision makers can be confident in the findings.

2.2 Research Questions

This research centered around the following question: What criteria should be used when assessing the quality of humanitarian protection evaluations? Sub-questions which helped to answer the main research question included:

- What level of quality do recent humanitarian protection evaluations have?
- What theories of change apply to humanitarian protection?
- Are attributions of protection outcomes being evaluated?
- What humanitarian organizations have the highest quality evaluations?

2.3 Process

The research process for this thesis was organized around the following four steps:

2.3.1 Information Collection

A strategic review of evaluation reports, academic literature, and grey literature⁴⁴ was conducted to identify and analyze resources that were relevant to the research questions.⁴⁵ Inclusion criteria for the selection of evaluation reports was defined as English language reports published between January 1st, 2013 and April 1st, 2021, and which had a central focus on humanitarian protection. This temporal parameter was chosen to review the most recent (last 8 years) evaluation literature and gain a contemporary and up to date perspective. Exclusion criteria was defined as non-protection interventions, non-evaluative studies, or non-humanitarian reports (for example, human rights-specific and peacekeeping interventions were excluded). The two main databases used for the identification of evaluation reports were the Active Learning Network and Accountability Project (ALNAP) Humanitarian Evaluation, Learning and Performance Library

⁴⁴ Grey literature includes handbooks, protection standards, evaluation guides, professional guidelines, project descriptions, protection policies of donors and operational actors, etc.

⁴⁵ Strategic here refers to reviewing a selection of the most relevant and up to date humanitarian protection literature.

(HELP) and the United Nations Evaluation Group (UNEG) Database of Evaluation Reports. In addition, the document libraries of protection specific organizations (UNHCR, UNICEF, Oxfam, UNFPA, IFRC), the International Initiative for Impact Evaluation library and Relief Web were searched. The selection of evaluation reports was purposive and is illustrative of a diverse range of stakeholders in the HPS as it includes twenty-two humanitarian organizations conducting protection work across twenty-one countries and six regions (the Middle East, Central America, South America, Europe, Sub-Saharan Africa and Oceania).⁴⁶ A total of ten evaluation reports were reviewed, all of which used external evaluators to conduct the evaluations and write the reports, with support from team members of the organization being evaluated. The evaluation reports (A-J) are summarized below in Table 1, the evaluation report matrix:

⁴⁶ “IASC Protection Priority: Global Protection Cluster,” interagencystandingcommittee.org (IASC), accessed April 19, 2021, <https://interagencystandingcommittee.org/iasc-protection-priority-global-protection-cluster#:~:text=Participants%20in%20the%20GPC%20Task,on%20the%20Human%20Rights%20of>.

Table 1: Evaluation Matrix

Evaluation	Title	Year	Location	Organizations	Crisis Type	Protection
				Evaluated		Intervention
A	Joint Evaluation: CSMC in Informal Tented Settlements	2016	Lebanon	Concern Worldwide, Norwegian Refugee Council, Anonymous Organization	War and displacement: Syrian Conflict	Collective site management and coordination of informal tented settlements, collective shelters, and collective centers
B	Performance Evaluation in LAC Urban DRR Programming: The Neighborhood Approach	2018	Colombia, Guatemala, Haiti, Honduras, Jamaica, Peru	USAID, Save the Children, PREDES, COOPI, Global Communities, Project Concern International, World Concern, Habitat for Humanity, GOAL	Natural Disaster: earthquakes, landslides and cyclones	Urban disaster risk reduction
C	Evaluation of Family Centres as Community Level Service Delivery Mechanisms Reaching Vulnerable Children in Gaza	2018	Palestine	UNICEF, Ma'an, Tamer	War and protracted displacement: Israel-Palestine Conflict	Child Protection via family centers
D	Evaluation of The Effects of Cash Based Interventions on Protection Outcomes In Greece	2018	Greece	UNHCR, Action Against Hunger	Displacement: refugees and asylum seekers in Greece	Cash Based Intervention
E	Working with Men to Prevent Intimate Partner Violence in A Conflict-Affected Setting	2014	Cote d'Ivoire	International Rescue Committee	War: "The Crisis" armed conflict within Cote d'Ivoire	Men's Intimate Partner Violence Prevention Discussion Group

F	Government of Uganda / UNFPA 8TH Country Program 2016 – 2020: Evaluation Report	2020	Uganda	UNFPA	Displacement: Refugees and IDPs in Uganda	Addressing gender-based violence and violence against children
G	Awareness Raising and Problems Of Protection: Evaluating the Impact of The Within and Without the State Pilot	2019	Democratic Republic of the Congo	Oxfam	Protracted conflict, political instability and failed state	Improvement of state-citizen relationships through dialogue
H	Impact Evaluation of the EU-IOM Joint Initiative Program for Migrant Protection and Reintegration	2020	Djibouti, Ethiopia, Eritrea, Kenya, Somalia, Uganda, South Sudan, Sudan	IOM	Displacement due to regional instability	Reintegration assistance
I	Final Evaluation: Building Resilient Communities in Fiji Project	2019	Fiji	ICRC	Natural Disaster: Tropical Cyclones	Building resilient communities to increase capacity and resilience in disaster preparedness
J	Evaluation of Child Friendly Spaces Iraq Field Study Report	2014	Iraq	Save the Children, UNICEF, World Vision	Displacement: Syrian Conflict	Child protection via child friendly spaces

2.3.2 Creating a Protection-Specific Evaluation Quality Assessment Framework

The protection-specific evaluation quality assessment framework (PSEQAF) created for this report was synthesized using the following sources: 1) quality standards set by the United Nations Evaluation Group (UNEG), the largest and most authoritative evaluation network globally

with over 50 members, 2) the OECD Better Criteria for Better Evaluation, and 3) a review of humanitarian evaluation guidebooks, manuals and grey literature.

The UNEG Quality Checklist for Evaluation Reports serves as a guideline for UNEG members in preparation and assessment of evaluation reports and includes eight critical indicators for what the UNEG deems a high-quality evaluation report.⁴⁷ These eight critical indicators are: 1) report structure, 2) object of evaluation, 3) evaluation purpose, objective(s) and scope, 4) evaluation methodology, 5) findings, 6) conclusions, 7) recommendations, and 8) gender and human rights. Of these eight UNEG indicators, report structure, evaluation methodology, conclusions, and recommendations were used to construct the PSEQAF. While these four indicators do not address protection-specific evaluation report concerns, they were used because they form the basis for what makes a high-quality evaluation report, that is a report that helps decision makers identify successful programs to expand and/or pitfalls to avoid based on evidence.⁴⁸

The OECD Better Criteria for Better Evaluation define six evaluation criteria (relevance, coherence, effectiveness, efficiency, impact, and sustainability) designed to improve the quality and usefulness of evaluation.⁴⁹ These criteria are a cornerstone of evaluation practice in development and are widely used, but do not offer specific guidance on assessing an evaluation report's quality. Rather, the OECD posit that using their six criteria play a normative role that leads to high quality, useful evaluation. Of the six OECD evaluation criteria, impact, effectiveness,

⁴⁷ "UNEG Quality Checklist for Evaluation Reports," UNevaluation.org (UNEG, June 2010), <http://www.unevaluation.org/document/detail/607>.

⁴⁸ Kerry Bruce, "What Is Quality in an Evaluation?" socialimpact.com (Social Impact, May 24, 2017), <https://socialimpact.com/what-is-quality-in-an-evaluation/>.

⁴⁹ "Evaluation Criteria," oecd.org (OECD), accessed April 19, 2021, <https://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm#:~:text=The%20OECD%20DAC%20Network%20on,two%20principles%20for%20their%20use.>

and relevance were selected for application in the PSEQAF. Impact was selected because protection evaluators must attempt to understand what difference a protection intervention makes in terms of reduced risk to produce a useful evaluation that helps decision makers understand what works and what does not. Effectiveness, which is closely related to impact, as it assesses whether an intervention met its objectives, which almost always include making a difference (impact), was incorporated into the PSEQAF as it is essential for assessing whether a program’s theory of change produced the intended results and whether its assumptions were correct. Relevance was incorporated into the PSEQAF because it is a necessary assessment to establish whether the intervention is using the appropriate protection strategy regarding the context, or whether the intervention is “doing the right things.”⁵⁰

Conducting a review of humanitarian protection guidebooks, manuals and grey literature resulted in identifying protection-specific evaluation concerns that were not addressed by the UNEG standards or the OECD criteria. ALNAP published the Evaluation of Humanitarian Action (EHA) Guide in 2016, as well as the Evaluation of Protection in Humanitarian Action (EPHA) Guide in 2018. These guides offer useful guidance about the specific challenges of evaluation in a humanitarian context and strategies for overcoming these challenges. The ALNAP EHA establishes that theories of change (ToC) are important aspects of both program and evaluation design. The concept of a theory of change represents the underlying logic that explains why a program’s inputs produce its outputs and objectives (the results chain). A well-defined theory of change can help humanitarian protection actors as well as evaluators understand the mechanisms

⁵⁰ “Evaluation Criteria,” oecd.org (OECD), accessed April 19, 2021, <https://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm#:~:text=The%20OECD%20DAC%20Network%20on,two%20principles%20for%20their%20use.>

that reduce risk for individuals, groups and communities targeted by the intervention. Theories of change also add transparency and rigor to humanitarian action because they reveal program details, assumptions, expected results, as well as the conceptual frameworks that causal linkages are based on. Moreover, considering the complex environment that surrounds an evaluation of humanitarian protection, theories of change are a solution to simplify and break up complex results chains into a series of causal linkages that can be tested directly.⁵¹ Needless to say, theories of change are a critical aspect of high-quality humanitarian programming and evaluation, and for this reason the Theory of Change Effectiveness criteria was included as an aspect of the PSEQAC.

Furthermore, the EPHA Guide discusses the importance of assessing the impact of protection work in terms of reduced risk, and whether any reduced risk is attributable to the protection intervention. Attribution of an outcome means that a specific programmatic activity causes an impact-level effect.⁵² This is also referred to in the literature as cause and effect or causality, and it is the gold standard of evidence-based decision making. Attribution of outcome is assessed via impact evaluations, a special type of evaluative inquiry that seek to provide causal statements regarding the effects of program interventions. In the humanitarian context, however, proving attribution or impact is traditionally seen as impossible, and is referred to by ALNAP as “the attribution challenge.”⁵³ This is because of the complex, chaotic and constantly evolving nature of humanitarian situations, as well as issues of spillover and complementarity that cloud causal linkages. Still, since impact evaluation is the gold standard of evidenced-based decision

⁵¹ Cosgrave, Buchanan-Smith, and Warner, “Evaluation of Humanitarian Action Guide,” ALNAP, October 10, 2016, <https://www.alnap.org/help-library/evaluation-of-humanitarian-action-guide>.

⁵² Christoplos, Dillon, and Bonino, “ALNAP Guide: Evaluation of Protection in Humanitarian Action,” ALNAP, October 25, 2018, <https://www.alnap.org/help-library/alnap-guide-evaluation-of-protection-in-humanitarian-action>.

⁵³ Christoplos, Dillon, and Bonino, “ALNAP Guide: Evaluation of Protection in Humanitarian Action,” ALNAP, October 25, 2018, <https://www.alnap.org/help-library/alnap-guide-evaluation-of-protection-in-humanitarian-action>.

making, and because millions of dollars and lives are at stake, the humanitarian context excuse is no longer valid. This development is supported by recent movements within both the development and humanitarian sectors to conduct rigorous impact evaluations regardless of context difficulty. For example, the Abdul Latif Jameel Poverty Action Lab (JPAL), one of the most highly respected and prolific publishers of poverty reduction impact evaluations in the world is launching in 2021 a Humanitarian Initiative. This initiative will involve the design and implementation of a multi-million-dollar portfolio of rigorous randomized evaluations of humanitarian interventions, with a focus on protection.⁵⁴ Furthermore, the International Initiative for Impact Evaluation (3ie) has teamed up with UNOCHA to fund high-quality impact evaluations and systematic reviews to improve evidence on humanitarian programs.⁵⁵ This demonstrates that the tide is turning for attribution of outcomes in the IHS generally as well as the HPS, and that there are serious efforts being made to conduct impact evaluations of humanitarian action. It is no longer seen as an impossibility and this is a sign of progress. For those reasons, the Attribution of Protection Outcome / Impact Assessment Criteria was used in the PSEQAC.

Weights were assigned to each of the protection-specific evaluation criteria regarding their relative importance. Higher weights in the PSEQAC correspond to greater importance when assessing protection-specific evaluation reports for quality. The two largest weights in the PSEQAC were assigned to Theory of Change Effectiveness and Attribution of Protection Outcome / Impact Assessment, as these two criteria received the most attention and importance in the literature. Protection Relevance, Protection Indicators, Complementarity, and Accountability were each assigned 10% as they are less important than the 20% weighted criteria but more

⁵⁴ “Humanitarian Initiative Policy Associate Job Posting,” [parisschoolofeconomics.eu](https://www.parisschoolofeconomics.eu/IMG/pdf/humanit-initiat-march-2020-policy-associate.pdf) (JPAL Europe, 2020), <https://www.parisschoolofeconomics.eu/IMG/pdf/humanit-initiat-march-2020-policy-associate.pdf>.

⁵⁵ “Our Work: Humanitarian,” [3ieimpact.org](https://www.3ieimpact.org/our-work/humanitarian) (3ie, 2020), <https://www.3ieimpact.org/our-work/humanitarian>.

important than the 5% criteria, which mainly address non-protection-specific structural basics of the reports.

The resulting framework is to the author's knowledge the first framework created to specifically appraise the quality of humanitarian protection evaluation reports. Table 2 below outlines the framework.

Table 2: Protection-Specific Evaluation Quality Assessment Framework

Protection-Specific Evaluation Criteria	Associated assessment questions	Weight
Report Structure	Is the report well structured, logical, clear, and complete? Are sections sequential?	5%
Evaluation Methodology	Does the report present a transparent description of the evaluation methodology and why it was used to address the protection evaluation criteria? Are limitations addressed?	5%
Theory of Change Effectiveness	Does the report critically assess the protection theory of change, expected risk reduction results chain and assumptions underlying the object’s intervention? Does the report determine which type of protection program model fits the intervention? (ex: stand-alone, integrated)	20%
Attribution of Protection Outcome / Impact Assessment	Do findings assess the impact in terms of reduced risk and whether it is attributable to the intervention?	20%
Protection Relevance	Does the report assess the appropriateness of the protection strategy for the context?	10%
Protection Indicators	Does the report assess whether the program’s protection indicators were well defined, SMART and based on consultations with those at risk?	10%
Complementarity	Does the report address the possibility of complementarity regarding the intervention’s outcomes?	10%
Accountability	Does the report assess the accountability of the protection actors to the local population and other key stakeholders?	10%
Unintended Consequences	Does the report assess whether the protection object adheres to the “do no harm”? principle?	5%
Conclusion and Recommendations	Are conclusions reasonable, substantiated by evidence and provide valuable insights? Are recommendations useful?	5%
		Total: 100%

2.3.3 Quality Assessment of the Evaluation Reports

To assess the quality of the selected evaluation reports using the newly developed PSEQAC, each evaluation report was read and given a weighted score from 0-4 for each criterion

based on the Global Evaluation Report Oversight System (GEROS) tool scoring system, shown in Table 3 below:⁵⁶

Table 3: The GEROS Quality Rating Scale

Source: “GEROS Handbook Summary,” UNICEF.org (UNICEF Evaluation Office, June 2017), <https://www.unicef.org/evaluation/global-evaluation-reports-oversight-system-geros>.

Rating	Implication	GEROS Score
Highly Satisfactory	Exceeds UNEG standards for evaluation reports. Decision makers may use the evaluation with a high degree of confidence.	4
Satisfactory	Meets UNEG standards for evaluation reports. Decision makers may use the evaluation with confidence	3
Fair	Meets UNEG standards for evaluation reports in some regards, but not all. Decision makers may use the evaluation with caution. Substantive improvements in some areas are needed.	2
Unsatisfactory	Does not sufficiently meet the UNEG standards for evaluation reports. Decision makers cannot rely on the evaluation.	1
Missing	Important aspects of the evaluation that are required by the UNEG standards were found to be absent. The evaluation report is incomplete.	0

2.3.4 Calculating a Quality Meta-Score for the Selected Reports

Lastly, each report’s score on the PSEQAC was used to calculate a weighted average score,

using the equation: $W = \frac{\sum_{i=1}^n w_i X_i}{\sum_{i=1}^n w_i}$ where W is the weighted average score, n is the number of

criteria to be averaged (10), w_i is weights applied to each criteria, and X_i are the criteria scores to

⁵⁶ “Global Evaluation Reports Oversight System (GEROS),” unicef.org (UNICEF Evaluation Office, 2009), <https://www.unicef.org/evaluation/global-evaluation-reports-oversight-system-geros>.

be averaged.⁵⁷ Weights were assigned to the different criteria according to their relative importance when evaluating humanitarian protection. Higher weights correspond to higher importance in regard to evaluating humanitarian protection and vice versa. The quality level of evaluation reports the humanitarian protection sector is capable of producing was observed by calculating a meta-score which took the average of the ten different weighted evaluation report scores. Using the GEROS scale, an average score of 3 or higher for the ten reports would mean that the IHS is producing high quality humanitarian protection evaluation reports, as reports meet standards and decision makers may use the evaluations with confidence. If the overall average was less than 3, then the selected reports would not be considered high quality, because reports meet standards only in some regards, so decision makers cannot rely on the evaluations. This would mean that major substantive improvements are needed in the IHS to effectively evaluate humanitarian protection. Another calculation factored into the assessment is the percentage of the ten reports with scores of 3 or higher, which translates into the percentage of selected reports with satisfactory versus fair/unsatisfactory quality. The assumption in this analysis is that as the quantity of high-quality protection evaluations increases, the overall ability of the IHS to evaluate protection increases. Logically, this means that the relationship between high quality evaluations and the overall evaluative ability of the IHS are positively correlated.

⁵⁷ “Weighted Mean,” mathisfun.com, accessed April 19, 2021, <https://www.mathisfun.com/data/weighted-mean.html>.

2.4 Limitations

While the criteria from the UNEG, OECD, and ALNAP, as well as the GEROS scorings system are useful in applying a rigorous and systemic method of quality appraisal to the evaluation reports, ultimately the judgements of a report's scores on the criterion are subjective. Due to time constraints and the nature of independent thesis research, there was no way to fully counter this limitation, such as the "four-eyes principle" which involves having at least two people involved in quality assessment.⁵⁸ To minimize this subjectivity bias, all evaluation reports were analyzed on two separate occasions to double-check scoring consistency.

Purposive sampling of evaluation reports was used for this research. As such, the reports reviewed represent only a portion of the universe of humanitarian protection evaluations that exist. A systematic review for this material was impossible due to the way evaluation reports are cached in 100s of different organizational databases and the time allotted. There are thousands of reports online but limited time to search, retrieve and review them. The author was only able to select the reports that were perceived as relevant and characteristic of the protection sector, so the selection of reports is by no means representative. The selection of reports simply attempted to cover a diverse range of relevant protection stakeholders on a global scale.

Lastly, in the broadest sense of the term, protection can be perceived as an aspect of almost all humanitarian action, because the international legal frameworks underpinning it, such as the Universal Declaration of Human Rights, touch all aspects of humanitarian programming. Therefore, it is difficult to separate protection evaluations and from other types of general

⁵⁸ Urban Reichhold and Andrea Binder, "Scoping Study: What Works in Protection and How Do We Know?," gppi.net (The Global Public Policy Institute, March 1, 2013), <https://www.gppi.net/2013/03/01/scoping-study-what-works-in-protection-and-how-do-we-know>.

humanitarian evaluation. This made it difficult to find purely protection-specific evaluation reports during the report selection process, so this analysis could be skewed by the selection of reports that might not be universally perceived as protection-specific. This limitation was minimized by screening evaluation reports in such a way that only those which evaluated programs with protection-related outcomes were included.

3.0 Findings

3.1 Quality Level of the Selected Evaluation Reports

Based on the calculation of the meta-score from ten evaluation reports, the selected reports are of only “fair” quality according to PSEQAC and the GEROS scoring tool. The meta-score of the ten reports was 2.79 on a scale from 0-4, indicating that reports meet standards only in some regards, that decision makers cannot rely on the evaluations and that improvements are needed to effectively evaluate humanitarian protection among the organizations in the selection. Five out of the ten evaluation reports earned scores of 3 or higher, contributing to an overall satisfactory evaluation rate of only 50%. Of the ten different protection specific evaluation criteria, Evaluation Methodology had the highest average score at 3.4, while Unintended Consequences had the lowest average score at 1.7. Five of the ten criteria average scores were above the satisfactory threshold of 3. Table 4 on the next page displays the evaluation reports and their corresponding scores on the different criteria, while Figure 6 displays a bar chart with regard to the satisfactory threshold of 3.

Table 4: Evaluation Report Scores

Protection Specific Evaluation Criteria	Weight	A	B	C	D	E	F	G	H	I	J	
Report Structure	5%	3	4	4	4	3	4	2	2	3	2	average: 3.1
Theory of Change / Intervention Logic	20%	1	3	4	4	2	3	4	3	3	2	average 2.9
Evaluation Methodology	5%	3	3	4	3	4	3	4	4	3	3	average: 3.4
Attribution of Protection Outcome / Impact assessment	20%	2	4	3	4	4	3	3	4	2	4	average: 3.3
Complementarity	10%	2	3	1	3	1	3	2	3	1	1	average: 2
Protection Context	10%	3	4	4	4	2	3	2	3	3	2	average: 3
Accountability	10%	4	1	3	2	1	4	1	3	2	1	average: 2.2
Protection Indicators	10%	2	2	3	3	2	3	2	4	3	2	average: 2.6
Unintended Consequences	5%	2	1	2	1	1	2	3	3	1	1	average: 1.7
Conclusions and recommendations	5%	3	4	4	3	4	4	4	2	3	2	average: 3.3
Weighted Scores	100%	2.25	3	3.2	3.35	2.4	3.15	2.75	3.35	2.4	2.2	meta-score: 2.79 (fair)



Figure 6: Quality Scores of Selected Reports

3.2 Theories of Change Found in Selected Reports

Ideally, most humanitarian protection programs would have an explicit and well-defined theory of change before they implementation commenced. Unfortunately, this is not always the case, as theories of change are not yet widespread practice in humanitarian programming generally or in protection programming more specifically, though there are some signs that this is changing.⁵⁹ This means that often, evaluators have to develop a theory of change based on what is implicit in the program design and implementation. This can lead to methodological problems

⁵⁹ Cosgrave, Buchanan-Smith, and Warner, “Evaluation of Humanitarian Action Guide,” ALNAP, October 10, 2016, <https://www.alnap.org/help-library/evaluation-of-humanitarian-action-guide>.

within the evaluation, like evaluators using the same data to both generate and test the theory, as well as confusion in interpreting results over whether invalid causal linkages are due to a flawed theory of change or a flawed evaluation model.⁶⁰

The ten selected evaluation reports scored mixed results in this category, ranging from 1 to 4, with an average of 2.9. While most (seven) evaluations at least acknowledged or discussed the theory of change of the intervention, only three assessed it critically and meaningfully. Overall, not enough attention was given to assessment of the theory of change in the evaluation reports selected. Figure 11 below displays the different theories of change represented in the selection:

Table 5: Protection Theories of Change Identified in Selection

Evaluation Report	Associated Theory of Change
A	Program lacked a formal or explicit theory of change
B	Disaster risk reduction programming prevents new risk, reduces existing risk, and manages residual risk. This contributes to strengthening resilience, ultimately reducing risk even more in the long run.
C	Program lacked a formal or explicit theory of change.
D	Cash based interventions allows persons of concern to meet basic needs with dignity and choice. This contributes to reduction of risky negative coping strategies. External influences include support from local community and access to local economy.
E	Gender transformative prevention strategy. Program targets gender inequitable normative beliefs and behaviors that condone or encourage violence against women, with the aim of confronting gender norms related to negative manifestations of masculinity.
F	Increasing the availability and use of sexual and reproductive health services that are gender responsive and meet human rights standards will decrease maternal and newborn deaths
G	When citizens and local authorities are trained in laws and universal rights, as well as in the processes of negotiations and petitions, they will have the knowledge, the competencies and the confidence to identify protection and governance problems. If they have a neutral space to discuss these problems, a regular dialogue can be established that permits the two parties to identify and implement actions to resolve these problems.

⁶⁰ “Evaluation of Humanitarian Action Guide,” 2016.

H	Sustainable integration via reintegration into the economic, social and cultural processes of the country of origin, accompanied by feelings of safety and security in the environment of return. This is achieved through individual support, community-based assistance and structural interventions.
I	Increasing the capacity and resilience of communities in disaster preparedness through community-based disaster management programming leads to reduced vulnerability and risk during disasters.
J	Child friendly spaces provide a protective and restorative environment for children in emergency settings via the protection of children from risk, the promotion of children’s psychosocial wellbeing and supporting caregivers and communities in strengthening systems of child protection.

Another finding relevant to theories of change in humanitarian protection was that even though much of the guidance material on humanitarian protection classifies protection interventions based on specific models (stand-alone, integrated, mainstreamed, remedial, etc.), nine out of ten evaluation reports did not attempt to make this distinction (only the UNHCR evaluation did). Therefore, either evaluators in the IHS do not understand the different types of protective actions or they are not concerned with distinguishing between protection models. Either way, it detracts from the usefulness and generalizability of the evaluation results if the specific type of protection model is not categorized, because without a typology, the overall connectedness and coherence of protection evidence is limited. Overall, despite their necessity in generating logical, transparent, and evidence-based programming, the evaluation reports selected did not do a satisfactory job of critically assessing humanitarian protection theories of change.

3.3 Attribution of Protection Outcomes

From the reports selected there was satisfactory evidence of evaluations rigorously assessing the attribution of protection outcomes or impact of programs, demonstrated by the criteria's average score of 3.3. Only two reports received scores of 2, three received a score of 3 and five received the highest score of 4. This suggests progress being made in the wider IHS to take impact evaluation seriously. Like the theory of change assessment, the attribution of outcome evaluation criteria was given a heavy weight (20%), due to its critical contribution in generating high-quality evidence. Two reports out of the ten (Impact Evaluation of the EU-IOM Joint Initiative Program for Migrant Protection and Reintegration as well as the IRC Cluster Randomized Controlled Trial in Côte d'Ivoire) were impact-specific, meaning their evaluation methodology did not just include an impact assessment, but was designed specifically around impact measurement. The other evaluations who had scores of 3 or higher on this criterion included formal impact measurements with baseline and endline measurements of protection indicators, or at least acknowledged that they were unable to measure impact given constraints that they faced. Evaluations with scores of 2 either did not make a meaningful attempt to evaluate the impact or conflated impact with effectiveness, showing that there is some confusion in the selected reports between evaluation of effectiveness versus evaluation of impact.

Overall, the evidence of satisfactory impact assessment in this report selection is one of the more positive and hopeful findings of this research, especially considering the long-held beliefs that humanitarian action generally cannot be evaluated for impact. The high-quality evidence for what works and what does not in humanitarian protection is growing, and this is a promising development.

4.0 Conclusion

It is not hugely surprising that the selected evaluation reports are not of satisfactory quality using the PSEQAC and the GEROS scoring tool, as protection is traditionally viewed as unevaluatable and less tangible than other aspects of humanitarian action. Still, the results of this study demonstrate two things: 1) the movement towards evidenced based decision making via rigorous evaluation in the IHS is well underway despite extensive challenges; and 2) the IHS has a substantial evidence gap to fill before it has a solid understanding of what works and what does not in humanitarian protection. The implications of these findings are both hopeful and jarring. Overall, the majority of the organizations represented in the selection of reports cannot understand the impact of their work. This is jarring because US\$ billions of dollars are potentially being squandered every year on their humanitarian protection programming.⁶¹ That is an incredible sum of money to take from donor governments and their taxpayers to implement programs whose impact is not yet understood, especially now that the COVID-19 pandemic has created a global economic recession. If and when donors start tying impact evidence to funding for protection organizations like UNHCR, they may not have enough evidence to prove their worth.

It should be noted that some organizations take evaluation more seriously than others in the IHS. From this selection, UNICEF, UNHCR and the IOM had the highest quality evaluations according the PSEQAC and the GEROS scoring tool. UNICEF especially is an example of best practice when it comes to how an organization can foster a culture of continuous learning and

⁶¹ 1.45 billion is the field budget for UNHCR's 2020-21 protection related objectives. Source: "UNHCR's 2020-2021 Financial Requirements," [reporting.unhcr.org](https://reporting.unhcr.org/sites/default/files/ga2020/pdf/Chapter_Financial.pdf) (UNHCR), accessed April 19, 2021, https://reporting.unhcr.org/sites/default/files/ga2020/pdf/Chapter_Financial.pdf.

improvement via evaluation. This is demonstrated by their yearly Best of Research and Evaluation publication, which showcases high quality research and evaluations from UNICEF offices around the world.⁶² Furthermore UNICEF created the incredibly useful GEROS tool which allows for systematic quality appraisal of evaluation reports, contributing to overall IHS evaluation connectedness and coherence. Other humanitarian organizations should look to UNICEF as an example of best practice when it comes to evaluation of humanitarian action.

Thousands of evaluations were found to exist in the humanitarian sector. Between the ALNAP HELP library and the UNEG Database of Evaluation reports there are 5,908 evaluations available. Only a small portion, however, relate directly and specifically to protection programming evaluations since 2013 (270 in the UNEG database and 31 in ALNAP). There is considerable difficulty in finding protection-specific evaluation reports, as programs and evaluations include both assistance and protection interventions as the evaluation object. This lack of distinction between protection interventions and material assistance clouds organizations' ability to conduct high quality evaluations, but the forthcoming JPAL Humanitarian Initiative with a focus on protection as well as the Journal of International Humanitarian Action's special collection on protection are signs that the IHS is aware of this and the needle is moving towards progress. Still, it is impossible to know what protection interventions work unless organizations takes meaningful steps to distinguish between humanitarian assistance from protection. This is one of the central findings of the research, that the IHS is only fairly fit for evaluation of protection because protection actors and evaluators do not draw clear conceptual and theoretical distinctions between material humanitarian assistance and protection. This may be due to lack of rigor or

⁶² "Best of UNICEF Research and Evaluation 2020," unicef.org (UNICEF Office of Research - Innocenti, 2020), <https://www.unicef-irc.org/boure2020>.

simply because the two are inseparable. Still, protection outcomes must be separated from assistance outcomes if the IHS is ever to make evidence-based decisions about what programs work and which do not.

In conclusion, this study has developed a framework for assessing the quality of protection-specific humanitarian evaluation reports and used that framework to assess the quality of ten evaluation reports in an attempt to understand the quality of current evaluative practice in the humanitarian protection of the IHS. The ten selected reports were found to be of only fair quality, which demonstrates the need for further investment of time and resources into humanitarian protection evaluation. This investment is imperative if we are ever to make decisions based on evidence in the humanitarian protection sector. In a world where humanitarian needs are increasing and the resources available to address them are shrinking, evaluation and evidence-based decision making hold an unprecedented importance.

5.0 Recommendations

The novel protection-specific evaluation quality assessment criteria, or PSEQAC developed for this research should be considered for use by organizations and independent evaluators who evaluate protection interventions. It is a useful starting point to start asking the right questions when assessing the quality of protection evaluations and seeking out evidence for what works and what does not in humanitarian protection. This framework addresses protection-specific evaluation concerns and because of that can lead to uncovering more useful and relevant evaluation findings than non-specific evaluation quality assessment criteria.

Since the ability of the selected humanitarian protection organizations to evaluate protection is average, a discussion of what can be done and by whom to fix this problem is warranted. As aforementioned, the movement to conduct rigorous evaluation of humanitarian intervention has begun and is slowly starting to get traction. The author believes that a few system-wide approaches could be taken to increase the adoption of high-quality evaluative practice within the IHS.

First, efforts must be made at the organizational level to create cultures of continuous learning and improvement through evaluation. Evaluation must change from being a secondary activity of organizations to being a permanent and central activity. This can be accomplished through the adoption of permanent evaluation offices like the UNICEF Evaluation Office and the IRC Research Technical Unit. These offices must be embraced, and their work must be mainstreamed at every level of the organization. Evaluation reports should be disseminated and discussed by relevant teams in simple, straightforward ways that allow the findings to be translated into actionable results. High-quality evaluations should be incentivized and rewarded to encourage

good practice. Weaving high-quality evaluative inquiry into the culture of humanitarian organizations is a bottom-up approach that can serve as the starting point for long-term evidence building in the IHS.

Secondly, humanitarian donors, be they governmental or private, must start requiring high-quality evaluation as a pre-requisite for funding. This top-down enforcement of evaluation is likely the most effective way at getting actors within the IHS to stop using their built-in excuse that humanitarian action is too difficult to evaluate due its context. The vertical resource dependency in the IHS means that humanitarian actors can be pressured to conform with the demands of donors, because humanitarians depend on donated resources for their very existence.⁶³ It should be relatively straightforward to convince governments that their taxpayers' money should only go to organizations who are prioritizing evaluation, because these organizations are basing their operations on facts and evidence, rather than assumptions, anecdotal experience, and intuition. Something like the Food and Drug Administration in the United States could theoretically be applied to humanitarian intervention to filter programs through a rigorous and standardized approval process. Humanitarian organizations would have to submit evidence that demonstrates the effectiveness of their programming to secure funding. The vetting process would include a quality analysis of their evaluations, and only organizations with high quality evidence above an agreed-upon standard (like 3 or higher on the GEROS scale) would be funded. This probably already occurs on an informal basis when humanitarian agencies submit proposals or grants, but I argue that this process should be formalized to enforce a level of rigor equivalent to that of the Food and Drug Administration. This is logical because in both the IHS and the United States food

⁶³ Taylor B. Seybolt, "Harmonizing the Humanitarian Aid Network: Adaptive Change in a Complex System," *International Studies Quarterly* 53, no. 4 (December 2, 2009): pp. 1027-1050, <https://doi.org/10.1111/j.1468-2478.2009.00567.x>.

and drug markets, the products and services offered have the goal of sustaining and saving lives. The fact that no formal process to selectively distribute humanitarian funding based on high-quality evaluative practice must change, and governmental and private donors are best positioned to develop this process.

Lastly, different sectors within the IHS should develop and standardize different evaluation quality appraisal frameworks for their interventions. Ideally this would mean each cluster having its own evaluation criteria, like the protection-specific framework developed in this research. This would establish evaluation coherence and generalizability among the 100s of different organizations in the IHS conducting diverse types of operations. Current evaluative standards from the UNEG and the OECD are helpful but are too general to capture the crucial differences between the many distinct types of interventions. The GPC, for example, could hold a conference to establish how best to conduct rigorous impact evaluations of protection, and then once a framework is agreed upon, ask all humanitarian protection stakeholders to sign a memorandum of understanding that incorporates the agreed upon framework into their evaluative practice. A common framework for conducting evaluations and assessing their quality would mean that the results would be more usable and dependable across different sectors. This has long been the practice in medicine, where frameworks like the Oxford Center for Evidence-Based Medicine and the Recommendation Taxonomy grading system used by the American Academy of Family Physicians are used to gauge levels of evidence and inform recommendations.⁶⁴ Different clusters and sectors within the IHS already coordinate on a high number of initiatives, and creating specific evaluative frameworks for each would be an achievable solution that could greatly increase the

⁶⁴ Opeyemi O. Daramola and John S. Rhee, "Rating Evidence in Medical Literature," *Journal of Ethics | American Medical Association* (American Medical Association, January 1, 2011), <https://journalofethics.ama-assn.org/article/rating-evidence-medical-literature/2011-01>.

quality of evaluations over time. Widespread practice in other fields of study indicate that the IHS is long overdue for this.

Appendix A Evaluation Report Selection

- Dunlop, Katherine, Mariagni Ellina, and Elizabeth Smith. "Evaluation of the Effects of Cash Based Interventions on Protection Outcomes in Greece - Final Report." unchr.org. UNHCR, December 2018.
<https://www.unhcr.org/research/evalreports/5c9217c87/evaluation-effects-cash-based-interventions-protection-outcomes-greece.html>.
- Eager, Rachel, Andrew Pinney, Michael Loevinsohn, Elisa Sandri, Leonora Evans-Gutierrez, Becka Kindler, and Richard Williams Chris Barnett. "IMPACT - Impact Evaluation of the EU-IOM Joint Initiative for Migrant Protection and Reintegration in the Horn of Africa Region." iom.org. EU and IOM, October 22, 2020.
https://ronairobi.iom.int/sites/default/files/document/publications/IOM_Methodological_Report_FINAL_20102020.pdf.
- Garcia, Ofelia. "Joint Evaluation: Collective Site Management and Coordination (CSMC) in Informal Tented Settlements (ITSs) - Lebanon." ALNAP.org, June 1, 2016.
<https://www.alnap.org/help-library/joint-evaluation-collective-site-management-and-coordination-csmc-in-informal-tented>.
- Hossain, Mazedra, Cathy Zimmerman, Ligia Kiss, Tanya Abramsky, Drissa Kone, Monika Bakayoko-Topolska, Jeannie Annan, Heidi Lehmann, and Charlotte Watts. "Working with Men to Prevent Intimate Partner Violence in a Conflict-Affected Setting: a Pilot Cluster Randomized Controlled Trial in Rural Côte D'Ivoire." bmcpublichealth.com. BMC Public Health, April 10, 2014. <https://www.ncbi.nlm.nih.gov/pubmed/24716478>.
- Kembo, Joshua, Paul Bukuluki, and John Mark Mwesigwa. "GOVERNMENT OF UGANDA / UNFPA 8 COUNTRY PROGRAMME 2016 2020." unfpa.org. UNFPA, January 2020.
https://www.unfpa.org/sites/default/files/CPE_Uganda.pdf.
- Lilley, Sarah, Abduljabar Atrooshi, Janna Metzler, and Alastair Ager. "Evaluation of Child Friendly Spaces." savethechildren.net. Save the Children, October 2014.
https://resourcecentre.savethechildren.net/sites/default/files/documents/iraq_sc_cfs_evaluation_field_study_report.pdf.
- Low, Iris, Leaine Robinson, and David Hesaie. "Final Evaluation: Building Resilient Communities in Fiji Project." alnap.org. IFRC, November 5, 2019.
<https://www.alnap.org/help-library/final-evaluation-building-resilient-communities-in-fiji-project>.
- Ngambwa, Janvier. "Awareness Raising and Problems of Protection: Evaluating the Impact of Active Citizenship Work in the DRC." oxfamamerica.org. Oxfam, February 2019.
<https://www.oxfamamerica.org/explore/research-publications/evaluating-impact-active-citizenship-work-pilot-equateur-drc/>.

Sarmiento, Juan Pablo. "Performance Evaluation in LAC Urban DRR Programming: The Neighborhood Approach ." usaid.gov. USAID, April 2018. https://pdf.usaid.gov/pdf_docs/PA00T34V.pdf.

Wilson, Theresa. "Evaluation of Family Centres as Community Level Service Delivery Mechanisms Reaching Vulnerable Children in Gaza for the Period June 2015 to October 2017 ." unicef.org. UNICEF, May 2018. <https://www.unicef-irc.org/publications/pdf/9.State%20of%20Palestine.pdf>.

Appendix B Figure References

- Figure 1: Knox, Clarke. “The State of the Humanitarian System 2018 - Full Report.” SOHS.ALNAP.org. ALNAP, December 5, 2018. <https://sohs.alnap.org/help-library/the-state-of-the-humanitarian-system-2018-full-report>.
- Figure 2: “Refugee Coordination Model (RCM) - UNHCR.” unhcr.org. UNHCR. Accessed April 19, 2021. <https://data2.unhcr.org/en/documents/download/65407>.
- Figure 3: “Emergency Handbook.” unhcr.org. UNHCR. Accessed April 19, 2021. <https://emergency.unhcr.org/entry/61190/cluster-approach-iasc>.
- Figure 4: Reichhold, Urban, and Andrea Binder. “Scoping Study: What Works in Protection and How Do We Know?” gppi.net. The Global Public Policy Institute, March 1, 2013. <https://www.gppi.net/2013/03/01/scoping-study-what-works-in-protection-and-how-do-we-know>.
- Figure 5: “Evaluation Criteria.” oecd.org. OECD. Accessed April 19, 2021. <https://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm#:~:text=The%20OECD%20DAC%20Network%20on,two%20principles%20for%20their%20use>.
- Figure 6: “Global Evaluation Reports Oversight System (GEROS).” unicef.org. UNICEF Evaluation Office, 2009. <https://www.unicef.org/evaluation/global-evaluation-reports-oversight-system-geros>.

Bibliography

- Barrett, Christopher B. "Food Aid's Intended and Unintended Consequences." SSRN Electronic Journal, March 2006. <https://doi.org/10.2139/ssrn.1142286>.
- "Best of UNICEF Research and Evaluation 2020." unicef.org. UNICEF Office of Research - Innocenti, 2020. <https://www.unicef-irc.org/boure2020>.
- Bruce, Kerry. "What Is Quality in an Evaluation?" socialimpact.com. Social Impact, May 24, 2017. <https://socialimpact.com/what-is-quality-in-an-evaluation/>.
- "CASP Checklists - CASP - Critical Appraisal Skills Programme." CASP, April 27, 2020. <https://casp-uk.net/casp-tools-checklists/>.
- Chaudri, Chimran, Kristina Cordes, and Nathan Miller. "Humanitarian Health Programming and Monitoring in Inaccessible Conflict Settings: a Literature Review." Global Health Cluster Remote Programming and Monitoring Team, April 8, 2019. <https://jhumanitarianaction.springeropen.com/track/pdf/10.1186/s41018-019-0055-x>.
- Christoplos, Dillon, and Bonino. "ALNAP Guide: Evaluation of Protection in Humanitarian Action." ALNAP, October 25, 2018. <https://www.alnap.org/help-library/alnap-guide-evaluation-of-protection-in-humanitarian-action>.
- "Cluster Working Group on Protection, Progress Report and Supporting Documents." interagencystandingcommittee.org. IASC, December 9, 2005. <https://interagencystandingcommittee.org/working-group/content/cluster-working-group-protection-progress-report-and-supporting-documents>.
- Cosgrave, Buchanan-Smith, and Warner. "Evaluation of Humanitarian Action Guide." ALNAP, October 10, 2016. <https://www.alnap.org/help-library/evaluation-of-humanitarian-action-guide>.
- Daramola, Opeyemi O., and John S. Rhee. "Rating Evidence in Medical Literature." Journal of Ethics | American Medical Association. American Medical Association, January 1, 2011. <https://journalofethics.ama-assn.org/article/rating-evidence-medical-literature/2011-01>.
- Davey, Eleanor, John Borton, and Matthew Foley. "A History of the Humanitarian System: Western Origins and Foundations." Humanitarian Policy Group. Overseas Development Institute, June 1, 2013. <https://www.odi.org/publications/7535-history-humanitarian-system-western-origins-and-foundations>.
- "Emergency Handbook." unhcr.org. UNHCR. Accessed April 19, 2021. <https://emergency.unhcr.org/entry/41813/cluster-approach-iasc>.

- “Evaluation Criteria.” oecd.org. OECD. Accessed April 19, 2021.
<https://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm#:~:text=The%20OECD%20DAC%20Network%20on,two%20principles%20for%20their%20use.>
- “GEROS Handbook Summary.” UNICEF.org. UNICEF Evaluation Office, June 2017.
[https://www.unicef.org/evaluation/global-evaluation-reports-oversight-system-geros.](https://www.unicef.org/evaluation/global-evaluation-reports-oversight-system-geros)
- “GHD Principles and Good Practice.” Good Humanitarian Donorship. Accessed July 28, 2020.
[https://www.ghdinitiative.org/ghd/gns/principles-good-practice-of-ghd/principles-good-practice-ghd.html.](https://www.ghdinitiative.org/ghd/gns/principles-good-practice-of-ghd/principles-good-practice-ghd.html)
- “Global Humanitarian Overview 2020.” unocha.org. OCHA, December 3, 2019.
[https://www.unocha.org/global-humanitarian-overview-2020.](https://www.unocha.org/global-humanitarian-overview-2020)
- “Humanitarian Initiative Policy Associate Job Posting.” parisschoolofeconomics.eu. JPAL Europe, 2020. [https://www.parisschoolofeconomics.eu/IMG/pdf/humanit-initiat-march-2020-policy-associate.pdf.](https://www.parisschoolofeconomics.eu/IMG/pdf/humanit-initiat-march-2020-policy-associate.pdf)
- “Humanitarian Protection: Improving Protection Outcomes to Reduce Risks for People in Humanitarian Crises.” ec.europa.eu. European Commission, September 6, 2020.
[https://ec.europa.eu/echo/what/humanitarian-aid/protection_en.](https://ec.europa.eu/echo/what/humanitarian-aid/protection_en)
- “IASC Policy on Protection in Humanitarian Action, 2016.” interagencystandingcommittee.org. IASC, October 14, 2016. [https://interagencystandingcommittee.org/iasc-protection-priority-global-protection-cluster/iasc-policy-protection-humanitarian-action-2016.](https://interagencystandingcommittee.org/iasc-protection-priority-global-protection-cluster/iasc-policy-protection-humanitarian-action-2016)
- “IASC Protection Priority: Global Protection Cluster.” interagencystandingcommittee.org. IASC. Accessed April 19, 2021. <https://interagencystandingcommittee.org/iasc-protection-priority-global-protection-cluster#:~:text=Participants%20in%20the%20GPC%20Task,on%20the%20Human%20Rights%20of.>
- Jobe, Kathleen. “Disaster Relief in Post-Earthquake Haiti: Unintended Consequences of Humanitarian Volunteerism.” *Travel Medicine and Infectious Disease* 9, no. 1 (January 2011): 1–5. [https://doi.org/10.1016/j.tmaid.2010.10.006.](https://doi.org/10.1016/j.tmaid.2010.10.006)
- Kent, Randolph C. *The Anatomy of Disaster Relief*. London: Pinter, 1987.
- Mackintosh, Anne. “International Response to Conflict and Genocide: Lessons from the Rwanda Experience .” Edited by David Millwood. academic.oup.com. Oxford University Press, September 1, 1996. [https://academic.oup.com/jrs/article-abstract/9/3/334/1553872?redirectedFrom=fulltext.](https://academic.oup.com/jrs/article-abstract/9/3/334/1553872?redirectedFrom=fulltext)
- Martínez, José Ciro, and Brent Eng. “The Unintended Consequences of Emergency Food Aid: Neutrality, Sovereignty and Politics in the Syrian Civil War, 2012-15.” *International Affairs* 92, no. 1 (2016): 153–73. [https://doi.org/10.1111/1468-2346.12502.](https://doi.org/10.1111/1468-2346.12502)

- “The Nobel Peace Prize 1901.” NobelPrize.org. Accessed July 28, 2020. <https://www.nobelprize.org/prizes/peace/1901/dunant/biographical/>.
- “Our Work: Humanitarian.” 3ieimpact.org. 3ie, 2020. <https://www.3ieimpact.org/our-work/humanitarian>.
- Owen, John M., and C. Alkin. *Program Evaluation: Forms and Approaches*. New York: The Guilford Press, 2007.
- “Protection Mainstreaming.” globalprotectioncluster.org. Global Protection Cluster. Accessed April 27, 2021. <https://www.globalprotectioncluster.org/themes/protection-mainstreaming/>.
- “The Protection of Human Rights in Humanitarian Crises - A Joint Background Paper by OHCHR and UNHCR.” ALNAP.org. IASC, May 8, 2013. <https://www.alnap.org/help-library/the-protection-of-human-rights-in-humanitarian-crisis-a-joint-background-paper-by-ohchr>.
- “Protection: What Is It Anyway?” IASC. Inter-Agency Standing Committee, January 4, 2016. <https://interagencystandingcommittee.org/protection/content/protection-what-it-anyway>.
- Puri, Jyotsna, Anastasia Aladysheva, Vegard Iversen, Yashodhan Ghorpade, and Tilman Brück. “What Methods May Be Used in Impact Evaluations of Humanitarian Assistance.” iza.org, January 2015. <https://www.iza.org/publications/dp/8755/what-methods-may-be-used-in-impact-evaluations-of-humanitarian-assistance>.
- Reichhold, Urban, and Andrea Binder. “Scoping Study: What Works in Protection and How Do We Know?” gppi.net. The Global Public Policy Institute, March 1, 2013. <https://www.gppi.net/2013/03/01/scoping-study-what-works-in-protection-and-how-do-we-know>.
- Seybolt, Taylor. “The Myth of Neutrality.” *Peace Review* 8, no. 4 (December 1996). <https://doi.org/10.1080/10402659608426005>.
- Seybolt, Taylor B. “Harmonizing the Humanitarian Aid Network: Adaptive Change in a Complex System.” *International Studies Quarterly* 53, no. 4 (December 2, 2009): 1027–50. <https://doi.org/10.1111/j.1468-2478.2009.00567.x>.
- Snyder, Hannah. “Literature Review as a Research Methodology: An Overview and Guidelines.” *Journal of Business Research*. Elsevier, August 1, 2019. <https://www.sciencedirect.com/science/article/pii/S0148296319304564>.
- “The Sphere Handbook: Standards for Quality Humanitarian Response.” spherestandards.org. SPHERE, 2018. <https://spherestandards.org/handbook/>.
- Taplin, Dana, and Helén Clark. “Theory of Change Basics.” www.actknowledge.org, March 2012. http://www.theoryofchange.org/wp-content/uploads/toco_library/pdf/ToCBasics.pdf.

“Theory-Based Approaches to Evaluation: Concepts and Practices.” Canada.ca. President of the Treasury Board, October 25, 2012. <https://www.canada.ca/en/treasury-board-secretariat/services/audit-evaluation/centre-excellence-evaluation/theory-based-approaches-evaluation-concepts-practices.html>.

“UNEG Quality Checklist for Evaluation Reports.” UNevaluation.org. UNEG, June 2010. <http://www.unevaluation.org/document/detail/607>.

“UNHCR - Global Trends 2019: Forced Displacement in 2019.” UNHCR Global Trends 2019, 2020. <https://www.unhcr.org/globaltrends2019/>.

“UNHCR Refugee Response Coordination Model.” unhcr.org. UNHCR. Accessed April 19, 2021. <https://www.unhcr.org/53679e2c9.pdf>.

“UNHCR’s 2020-2021 Financial Requirements.” reporting.unhcr.org. UNHCR. Accessed April 19, 2021. https://reporting.unhcr.org/sites/default/files/ga2020/pdf/Chapter_Financial.pdf.

United Nations High Commissioner for Refugees. “Refugee Statistics.” Accessed April 19, 2021. <https://www.unhcr.org/refugee-statistics/>.

“Weighted Mean.” mathisfun.com. Accessed April 19, 2021. <https://www.mathisfun.com/data/weighted-mean.html>.

“What Is the Cluster Approach?” Humanitarian Response. UN OCHA. Accessed July 28, 2020. <https://www.humanitarianresponse.info/en/about-clusters/what-is-the-cluster-approach>.

“WHO WE ARE.” globalprotectioncluster.org. Global Protection Cluster. Accessed April 19, 2021. <https://www.globalprotectioncluster.org/about-us/who-we-are/>.