



**MONITORING &
EVALUATION**

ALNAP

Review of the OECD DAC criteria for evaluating humanitarian action

A MAPPING OF LITERATURE, GUIDANCE AND PRACTICE

ALNAP is a global network of NGOs, UN agencies, members of the Red Cross/Crescent Movement, donors, academics and consultants dedicated to learning how to improve response to humanitarian crises.

www.alnap.org

About the authors

Jo-Hannah Lavey is a Principal Consultant at Alinea International with 20 years' experience in humanitarian action across four continents. She has a deep passion for improving outcomes for those affected by crisis.

Christina Northey is a Senior Consultant at Alinea International with over 25 years' experience in humanitarian program design, implementation, monitoring and evaluation in Asia and the Pacific, Middle East and Africa.

David Poulton is an Associate Consultant at Alinea International with experience in research, design, and monitoring and evaluation of humanitarian and development programs throughout Asia and the Pacific.

Claire Bowyer is an Associate Consultant at Alinea International with over 5 years' experience in the delivery, monitoring, and evaluation of humanitarian and development projects across Asia, the Middle East, and the Pacific.

Acknowledgements

This research paper is part of a broader project to update ALNAP's existing guidance on using the OECD DAC Evaluation criteria for evaluating humanitarian action.

The project is managed by Susanna Morrison-Métois, a Senior Research Fellow for Evaluation and Learning at ALNAP.

A global advisory group provided oversight, advice and guidance. The advisory group is chaired by John Mitchell (ALNAP). Advisory group members include:

Helene Juillard, Evaluation Consultant
Henri Van den idert, UNHCR
Kassem El Saddik, Evaluation Consultant
Katy Bobin, independent
Margie Buchanan-Smith, Evaluation Consultant
Megan Kennedy-Chouane, OECD DAC
Michael Carbon, WFP
Miki Tsukamoto, International Federation of the Red Cross
Neelofar Shahzad, independent
Olivia Roberts/ Michele Tarsilla, UNICEF
Sylvestre Musengimana, Evaluation Consultant

Suggested citation

ALNAP (2023) *Review of the OECD DAC criteria for evaluating humanitarian action*. London: ALNAP/ODI

ISBN: 9781913526351

© ALNAP/ODI 2023. This work is licensed under a Creative Commons Attribution-non Commercial Licence (CC BY-NC 4.0).

Communications management by Maria Gili and Molly Maple
Typesetting by Inkwell Design and Communications Studio, www.inkwellcommunications.org
Design by Soapbox, www.soapbox.co.uk

CONTENTS

ACRONYMS 3

EXECUTIVE SUMMARY 5

INTRODUCTION 10

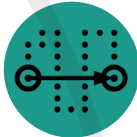
BACKGROUND AND OVERVIEW 15



EFFECTIVENESS 26



RELEVANCE/APPROPRIATENESS 35



EFFICIENCY 43



IMPACT 49



COVERAGE 57



COHERENCE 64



CONNECTEDNESS AND SUSTAINABILITY 71



ADDITIONAL CRITERIA AND CROSS-CUTTING THEMES 80

CONCLUSION 95

BIBLIOGRAPHY 96

ANNEX 1: LIST OF EVALUATIONS 105

Tables

Table 1: Coding framework	12
Table 2: IAHE application of criteria in evaluations since 2018	23
Table 3: Evaluation criteria by guide or standard	24
Table 4: Organisational guidance on effectiveness	30
Table 5: Organisations evaluations on effectiveness	31
Table 6: Organisational guidance on relevance/appropriateness	38
Table 7: Organisational evaluations on relevance/appropriateness	39
Table 8: Organisational guidance on efficiency	45
Table 9: Organisational evaluations on efficiency	46
Table 10: Organisational guidance on impact	53
Table 11: Organisational evaluations on impact	54
Table 12: Organisational guidance on coverage	59
Table 13: Organisational evaluations on coverage	61
Table 14: Organisational guidance on coherence	67
Table 15: Organisational evaluations on coherence	68
Table 16: Organisational guidance on connectedness/sustainability	75
Table 17: Organisational evaluations on connectedness and sustainability	76
Table 18: Definitions of additional criteria in sector-wide guidance	83
Table 19: Application of additional criteria and cross-cutting themes in 40 evaluations	86

ACRONYMS

AHP	Australian Humanitarian Partnership
ALNAP	Active Learning Network for Accountability and Performance in Humanitarian Action
BHA	Bureau for Humanitarian Assistance
CaLP	Cash Learning Partnership
CHS	Core Humanitarian Standard
CRS	Catholic Relief Services
DEC	Disasters Emergency Committee
DFAT	Department of Foreign Affairs and Trade
DFID	UK Department for International Development
DRC	Democratic Republic of the Congo
EHA	Evaluation of Humanitarian Action
FAO	Food and Agriculture Organization
FCDO	Foreign, Commonwealth and Development Office
GLAM	Global Learning for Adaptive Management
HAG	Humanitarian Advisory Group
HDP	Humanitarian–development–peace
HPG	Humanitarian Policy Group
HR	Human resources
HRP	Humanitarian Response Plan
IAHE	Inter-Agency Humanitarian Evaluation
IASC	Inter-Agency Standing Committee
ICRC	International Committee of the Red Cross
IDP	Internally displaced persons
IFRC	International Federation of Red Cross and Red Crescent Societies
IHL	International humanitarian law
IOM	International Organization for Migration
M&E	Monitoring & evaluation
MSF	Médecins Sans Frontières
NGO	Non-governmental organisation
NRC	Norwegian Refugee Council

OCHA	Office for the Coordination of Humanitarian Affairs
OECD	Organisation for Economic Co-operation and Development
OECD DAC	Organisation for Economic Co-operation and Development's Development Assistance Committee
OFDA	USAID's Office of Foreign Disaster Assistance
PNG	Papua New Guinea
SEAT	Social Equity Assessment Tool
SOHS	State of the Humanitarian System
UN	United Nations
UNEG	United Nations Evaluation Group
UNFPA	United Nations Population Fund
UNHCR	United Nations High Commission for Refugees
UNICEF	United Nations Children's Fund
USAID	United States Agency for International Development
WASH	Water, sanitation and hygiene
WFP	World Food Programme

EXECUTIVE SUMMARY

In 2006, ALNAP published [Evaluating humanitarian action using the OECD DAC criteria, an ALNAP guide for humanitarian agencies](#) to strengthen the evaluation of humanitarian action. The primary purpose of this paper is to provide evidence, using existing literature for ALNAP to review and update its 2006 guide. It also serves as a review on the use of the evaluation criteria in humanitarian settings, contributing to the broader body of evaluative knowledge.

This paper, a mapping of literature, guidance and practice, provides a high-level overview of each of the OECD DAC criteria for evaluating humanitarian action, how they have been interpreted in sector-wide and organisational guidance, key issues identified in the literature and application in contemporary evaluations. It does not identify contemporary, undocumented debates on the criteria; these will be addressed through forthcoming consultations.

The OECD DAC evaluation criteria are the pre-eminent criteria for evaluating development and humanitarian assistance (Kennedy-Chouane 2020, Picciotto 2013, cited in Patton 2020). Adapted in 2019, the six criteria are: **effectiveness, relevance, efficiency, impact, sustainability and coherence**. ALNAP's 2006 guide interprets the criteria for application in humanitarian action as: **effectiveness, relevance/appropriateness, efficiency, impact, coverage, coherence and connectedness**.

Box 1: Select OECD DAC and ALNAP documents informing criteria for evaluating humanitarian action

- 1991: OECD DAC Principles for Evaluation of Development Assistance
- 1999: OECD DAC Guidance for Evaluating Humanitarian Action in Complex Emergencies
- 2002: OECD DAC Glossary of Key Terms
- 2006: ALNAP Evaluating Humanitarian Action Using the OECD DAC Criteria: An ALNAP guide for humanitarian agencies
- 2016: ALNAP Evaluation of Humanitarian Action Guide
- 2019: OECD DAC Better Criteria for Better Evaluation: Revised Evaluation Criteria Definitions and Principles for Use
- 2021: OECD DAC Applying Evaluation Criteria Thoughtfully

The criteria

More evaluations of humanitarian action are being conducted than ever before, and the OECD DAC criteria remain one of the most used frameworks (Darcy and Dillon 2020). Of the seven criteria in ALNAP's 2006 guide, a sample of 120 evaluations¹ found the most commonly applied criteria to be (from most commonly applied to least commonly applied):

1. Effectiveness
2. Appropriateness/Relevance
3. Efficiency
4. Impact
5. Coverage
6. Coherence
7. Connectedness

¹ Identified for the 2018 State of the Humanitarian System Report, in Darcy and Dillon 2020

Effectiveness is the most applied of all the OECD DAC evaluation criteria. Effectiveness is generally defined in terms of achieving objectives. Challenges were identified in determining what constitutes an objective, and how they are best measured in humanitarian contexts. Sector-wide evaluation guidance differently emphasises aspects of effectiveness, such as coordination, timeliness, inclusion and quality.

Relevance/Appropriateness is the second most-used OECD DAC criterion in humanitarian evaluations (Darcy and Dillon 2020). The literature highlights two main issues related to its application. The first: that the power differential in humanitarian aid means that the evaluators' or evaluation managers' interpretation of what is relevant and appropriate may be very different from what is considered relevant and appropriate by the affected population. The second, and related issue: that the criterion is not systematically applied across evaluations, meaning two evaluations of the same programme could potentially come to different conclusions about its relevance and appropriateness.

Efficiency is the third most-applied OECD DAC criterion. Cost-efficiency and timeliness are common components of the criterion. Operational efficiencies, such as consortia and partnerships, were often included in evaluations. While important for decision-making, learning and accountability, the literature indicates that evaluation of efficiency often suffers from weak data sources and variable methodologies, and fails to account for social and environmental costs.

Impact is the fourth most-applied OECD DAC criterion, considered in almost 50% of humanitarian evaluations. It includes positive and negative, intended and unintended, and direct and indirect impacts. While valued as an important criterion to evaluate, evaluators face challenges in establishing cause and effect and attribution, particularly in single-organisation evaluations.

Coverage is an additional criterion in ALNAP's 2006 guide, and one of the least used of the criteria. It is, however, highly relevant to current issues identified in the performance of humanitarian action, and a priority criterion for some. There is some variance in how it is defined and applied in terms of how explicitly the evaluation considered geographic coverage, socioeconomic coverage and proportionality to need.

Coherence is one of the least-applied criteria in humanitarian evaluations. While it has long been a humanitarian evaluation criterion, it has only recently been added to the OECD DAC evaluation criteria for development (OECD DAC 2019). The literature focuses on the sub-division of coherence into internal coherence – within an organisation's own policies and standards – and external coherence with other actors and standards, with varying levels of focus on coherence with humanitarian principles. The criterion was represented mainly in UN, International Federation of Red Cross and Red Crescent Societies (IFRC) or donor-led evaluations, supporting literature findings on the complexity of assessing coherence for a single organisation or programme, particularly for NGOs.

Connectedness replaces the OECD DAC criterion of 'sustainability' in ALNAP's 2006 guide. It is among the least-used criteria in evaluating humanitarian action. Perhaps surprisingly, sustainability has been applied slightly more often than connectedness. It is 'unusual among the OECD DAC performance criteria in that there is disagreement as to whether connectedness should be used as a measure of humanitarian performance at all' (ALNAP 2018a: 239). This disagreement is anchored in the long history and continued debate regarding the relationship between humanitarian action and development (ALNAP 2018a). It is defined variably across evaluations. The most commonly applied definition is a version of 'lasting benefits', similar to the OECD DAC 2019 criteria of sustainability. This

was followed by links with development, resilience, risk reduction and peacebuilding. **Additional criteria and cross-cutting themes.** There is a plethora of additional and cross-cutting themes applied in humanitarian evaluations. ALNAP's original 2006 guidance on the criteria included: local context, human resources, protection, participation of primary stakeholders; coping strategies and resilience, gender equality, HIV/AIDS and the environment. ALNAP's State of the Humanitarian System (SOHS) publications have included sufficiency, complementarity, accountability and participation (ALNAP 2018a). While not explicitly including cross-cutting themes, the 2021 OECD DAC guidance focuses on inclusion and encourages application of a gender lens to evaluations. Gender, inclusion, protection, accountability and participation are common evaluative cross-cutting themes in individual organisational guidance. This is reflected in an analysis of a sample of humanitarian evaluations, which found gender, equity and inclusion, followed by accountability, participation and communication with communities to be the most commonly applied cross-cutting themes, followed by protection. These, together with adaptive management, are explored briefly.

Common issues identified in applying OECD DAC criteria to humanitarian action

This paper sought to identify documented issues evaluators face in interpreting or applying the criteria to humanitarian action. These issues have given rise to questions that could usefully be explored in forthcoming consultations designed to inform ALNAP's update to its guidance on the criteria.

A foundational question for future humanitarian guidance is how closely it should align to the OECD DAC guidance (and the adaptation to the criteria in 2019). Other common issues identified across the criteria include:

- The importance of **positionality** and whose perspective is used in defining the evaluative questions and who conducts the evaluation. The views of the person who defines what is effective and how the performance of an intervention is measured will likely have a significant impact on findings. This is related to calls for the **decolonisation of evaluation**. Chilisa and Mertens (2021) find that evaluation is dominated by Western culture and approaches, reinforcing biased power relations (2021: 242). Ofir (2017) applies this directly to the OECD DAC criteria, calling out insufficient recognition of the importance of culture and cultural differences (Ofir 2017).
- There are widespread challenges in providing guidance that supports the humanitarian sector's ability to compare findings across contexts and over time, and the substantive resources and skill set that are required to conduct evaluations well in accordance with such guidance and common definitions. This is reflected in a desire for more guidance to **improve standardisation** (Darcy and Dillon 2020), **while maintaining flexibility** in application (DEval 2018).
- The **utility and application of the criteria** are variable and depend largely on the type of programme and organisation and the intent of the evaluation.

Going forward

The OECD DAC criteria have proven highly popular in the evaluation of humanitarian action. Their use has contributed to the sector's ability to improve the quality of evaluation, and to compare findings across evaluations. This mapping of literature, guidance and evaluations has aimed to provide a foundation for consultations on the contemporary debate on the use of humanitarian evaluation criteria and how ALNAP should update its guidance. This debate is taking place amid increasing calls to question existing evaluation criteria to reflect complexity, systems thinking, transformational change and decolonisation, among other concerns.

Moreover, there is an opportunity to improve how we understand and evaluate 'good' humanitarian action – foundational to improving humanitarian action itself, including how to better ensure that affected communities and their views are central in humanitarian accountability and performance criteria. The interpretation and application of the OECD DAC criteria to evaluate humanitarian action is a small, but important, contribution.

INTRODUCTION

INTRODUCTION

In 2006, ALNAP published [Evaluating humanitarian action using the OECD DAC criteria, an ALNAP guide for humanitarian agencies](#) to strengthen the evaluation of humanitarian action. ALNAP is now in the process of reviewing its 2006 guide.

The humanitarian sector has since gained substantial experience applying the criteria to humanitarian action. The primary purpose of this paper is to provide evidence, using existing literature for ALNAP to review and update its 2006 guide. It also serves as a review on the use of the evaluation criteria in humanitarian settings, contributing to the broader body of evaluative knowledge.

This review of ALNAP's 2006 guide to evaluating humanitarian action using the OECD DAC (Organisation for Economic Co-operation and Development's Development Assistance Committee) criteria is occurring in the context of a world emerging from the COVID-19 pandemic, exposed to increasing impacts of climate change intersecting with insecurity, and ever-increasing numbers of forcibly displaced people.

The latest State of the Humanitarian System (SOHS) report (ALNAP 2022) reported that the number of people in need of humanitarian assistance reached 493.2 million at the peak of the pandemic, with UN appeals barely aiming to assist 60% of them. The number of conflicts more than doubled in the decade to 2020, and record numbers of people were displaced by 2021. The world needs to assist and protect those in need in the most effective way possible.

The global systems needed to respond effectively are, however, simultaneously being challenged. Established power structures are struggling to respond to calls for decolonisation, and for power to shift towards more local actors. Direct funding to local actors declined in 2021 to around 1.5% of all international humanitarian funding (ALNAP 2022). Multilateralism is under strain and governments are rejecting humanitarian norms.

The humanitarian sector has evolved substantially since 2006. It is far larger and more diverse. It has undertaken a series of reform efforts, to become more accountable, more effective and better protect those impacted by crises. It is, however, yet to deliver the transformative change some have envisioned to meet the global contextual challenges.

Evaluation criteria frame the priorities for what is considered good practice when evaluating humanitarian action. The criteria therefore have an important role to play in supporting the transformative changes that are required to assist those most in need within the current global context.

About this paper

This paper begins with a brief overview of the OECD DAC criteria, ALNAP's guidance and other approaches. It then proceeds with separate sections for each of the criteria identified in the 2006 ALNAP guide, comparing them with the adapted OECD DAC criteria definitions, literature, guidance and application in evaluations. The final section is dedicated to cross-cutting themes. This paper does not identify contemporary, undocumented debates on the criteria; these will be addressed through forthcoming consultations.

Box 2: Shortcut terminology

Throughout this paper we will use terms as follows:

- ALNAP's 2006 guide refers to the guidance for applying the seven criteria, in Evaluating humanitarian action using the OECD DAC criteria (ALNAP 2006). These criteria are: effectiveness, appropriateness /relevance, efficiency, impact, coverage, coherence and connectedness.
- The OECD DAC criteria refers to the six criteria established by the OECD DAC in 2019 (unless otherwise specified). These criteria are: effectiveness, relevance, efficiency, impact, sustainability and coherence.

Methodology

The paper employed three main research methods, namely document reviews of the following:

1. **Humanitarian evaluation guidance and standards**, intended for sector-wide application and those published by individual agencies.
2. **Academic and grey literature**, including peer-reviewed journal articles, humanitarian and organisational publications and blogs by evaluation experts.
3. **Humanitarian evaluations**, focusing on evaluations of responses to crises undertaken in the last five years and across geographies.

To supplement these main research methods, the following was also carried out:

- A formal feedback process through an advisory group established by ALNAP to support the process for updating its guidance.
- Key informant interviews.

Humanitarian evaluation guidance and standards

The paper reviewed 43 guidance documents and standards. These include ALNAP and OECD DAC guidance, which are used as a point of comparison to other interpretations of the criteria, including:

- Sector-wide guidance and standards such as the Inter-Agency Humanitarian Evaluations Process Guidelines and the Core Humanitarian Standard. These are explored further in Chapter 1: Overview and Background.
- Evaluation guidance published by individual agencies, including the United States Agency for International Development (USAID), UN agencies, NGOs and the IFRC

The guidance documents were identified through a web search and consultations with the advisory group. All guidance documents considered relevant to the update of the ALNAP guide have been included. Each are analysed by criterion using the definition in the ALNAP 2006 guide as the basis for comparison, and against key issues arising throughout the literature. Guidance is excluded from analysis tables in this paper where there was insufficient detail in the guidance on the topic.

Academic and grey literature

The research team compiled a comprehensive list of grey and academic literature related to the application of the OECD DAC criteria to identify the key issues in defining, interpreting and applying the criteria. The literature was identified through:

- Lexical searches of 21 humanitarian, development, academic and evaluation databases and websites, using keywords together with each of the criteria.
- Search strings on the most relevant publications.
- Recommendations from evaluation experts.

The list focused on literature published since 2006, the year of the original ALNAP guide. Earlier literature is included where it is a foundational document or fills an important gap.

After identifying an initial 80 publications, the research team undertook a rigorous prioritisation process, assessing the literature across four criteria of relevance and one of quality. A final list of 53 publications were analysed in NVivo against a coding framework based on the research questions and parent codes, summarised in [Table 1](#). Sub-codes were also applied to differentiate between humanitarian and development interpretation and application.

Table 1: Coding framework

Research question	Parent code
<i>Overarching research question: What are the key issues and recommendations arising from applying the OECD DAC criteria to the evaluation of humanitarian assistance since the original ALNAP guide was produced in 2006?</i>	
<i>A. What has been learned through the OECD DAC review (2019) that can be applied to the review of the 2006 ALNAP guide?</i>	Relevance, effectiveness, impact, coherence, efficiency, sustainability, inclusion, gender
<i>B. How have the humanitarian criteria, including additional criteria, been applied and interpreted since the ALNAP guidance on applying the criteria to humanitarian action was produced in 2006?</i>	Relevance, effectiveness, impact, coherence, efficiency, sustainability, connectedness, coverage
<i>C. What are the options for addressing related issues in updated humanitarian guidance e.g., cross-cutting issues, application to nexus programming, localisation of humanitarian assistance?</i>	19 possible cross-cutting themes or additional criteria, plus 'other' ¹²

Humanitarian evaluations

Fifty-nine evaluations have been reviewed to inform this paper (see [Annex 1](#) for a list of the evaluations). A sample of 40 evaluations were initially selected and analysed to provide an indication of how the criteria are being applied in practice. The ALNAP HELP database was the primary source for identifying evaluations, supplemented with advice from evaluation experts, and agency and donor evaluation databases and websites where there was a need to diversify the sample. The sample prioritised evaluations conducted since 2017 for a more contemporary analysis, and evaluations that focus on crisis response given the relatively small sample size and focus on humanitarian evaluation. Joint evaluations were of particular interest, where multiple stakeholders had agreed on the evaluation criteria.

¹² The list of 19 cross-cutting themes can be found in [Table 18: Application of additional criteria and cross-cutting themes to 40 evaluations](#). Each evaluation criterion is assessed in at least 10 evaluations framed by the issues identified in

the literature review. These evaluations were chosen for sufficient consideration of the criterion, as well as a diversity of agency, geography and type of crisis. An additional 19 evaluations were added to the initial sample of 40 where the initial evaluations were found to be of insufficient depth for a particular criterion. For example, an evaluation may not have provided a definition or related questions for the criterion, making it too difficult to analyse in accordance with the key issues identified in the literature and evaluation guidance. Evaluations were also added where there was value in reviewing targeted evaluations that were not included in the initial sample.

Limitations

This paper aims to provide a high-level overview of key issues and their application across seven criteria, as well as additional criteria and cross-cutting themes. It is limited in the depth of analysis for each criterion due to the number of criteria and themes, the specificity (and occasionally absence) of literature for each criterion, and the time required for targeted analysis against key issues identified separately for each criterion across both organisational guidance and evaluations.

Structure

The paper begins by providing background to the development of evaluation criteria in the context of humanitarian action and how the criteria have been interpreted by various organisations. It is followed by a chapter for each of the seven OECD DAC criteria as defined in ALNAP's 2006 guide: relevance/appropriateness, coherence, effectiveness, efficiency, impact, connectedness (and sustainability) and coverage. A final chapter reviews additional and cross-cutting criteria.

Each of the chapters focusing on a criterion follows the same structure:

- **Definitions:** definitions from the ALNAP 2006 guide and the OECD DAC 2019 adaptation, providing a common reference point for the review of each criterion.
- **Sector-wide guidance and standards:** definitions and interpretations of the criteria by the more prominent guidance and standards in the humanitarian sector. The review team has selected the original ALNAP 2006 guide, subsequent ALNAP publications, the OECD DAC's criteria and guidance, the Core Humanitarian Standard and the Inter-Agency Standing Committee's Inter-Agency Humanitarian Evaluation process guidelines.
- **Key issues:** key issues discussed in the academic and grey literature are identified and applied to individual organisational guidance and humanitarian evaluations to understand how the criterion is being applied in practice.
- **Summary and questions to explore:** a summary of findings from the literature, guidance and evaluations, with a list of questions arising for exploration.



BACKGROUND AND OVERVIEW

BACKGROUND AND OVERVIEW

The OECD DAC evaluation criteria

Original OECD DAC evaluation criteria and adaptation for complex emergencies

The OECD DAC evaluation criteria have origins dating back to 1991 in the publication: [Principles for evaluation of development assistance](#) (OECD 1991). The five original criteria were: **effectiveness, relevance, efficiency, impact and sustainability**. **Coherence** was added as a sixth criteria in 2019. They have been developed, adopted and reaffirmed through consultation over the years.

While the criteria are intended to be appropriate for humanitarian assistance, the OECD DAC found that, in the context of complex emergencies, there are necessary differences in the ways in which assistance is provided and evaluations are conducted. The OECD DAC 1999 [Guidance for evaluating humanitarian assistance in complex emergencies](#) made the following adaptations to the original five criteria to account for the different context and approach in these circumstances:

- added two additional criteria: **coverage and coherence**
- included **appropriateness** with relevance as complementary criteria
- replaced sustainability with **connectedness**.

It also noted the importance of coordination and protection as critical to the effectiveness of humanitarian action. The adaptations made in the OECD DAC 1999 *Guidance for evaluating humanitarian assistance in complex emergencies* formed the basis of ALNAP's 2006 guidance.

At the time of drafting its guidance for evaluation in complex emergencies, the OECD DAC defined complex emergencies as: situations where armed conflict and political instability are the principal causes of humanitarian needs, as distinct from situations where natural hazards are the principal cause of such needs (OECD 1999).

Adapted OECD DAC criteria 2019

The original OECD DAC evaluation criteria underwent an adaptation process from 2017 to 2019, resulting in the publication of the OECD DAC evaluation criteria in 2019 and supported by guidance published in 2021. The adapted criteria are: **effectiveness, relevance, efficiency, impact, sustainability and coherence**. **Coherence** was added to 'better capture synergies, linkages, partnership dynamics and complexity'. The original five criteria were adapted to improve clarity, with new definitions and better integration of equity, gender equality and the imperative to 'leave no one behind' – particularly across the relevance and effectiveness criteria (OECD 2019).

In the 2019 adaptation, the OECD DAC included two important principles (see [box](#) below) to guide the use of the criteria. These are intended to be supplemented by guidelines and standards and applied as appropriate to the broader context of an evaluation (OECD 2019).

Box 3: Principles to guide the use of the OECD DAC evaluation criteria

Principle 1

The criteria should be applied thoughtfully to support high-quality, useful evaluation. They should be contextualised – understood in the context of the individual evaluation, the intervention being evaluated and the stakeholders involved. The evaluation questions (what you are trying to find out) and what you intend to do with the answers, should inform how the criteria are specifically interpreted and analysed.

Principle 2

Use of the criteria depends on the purpose of the evaluation. The criteria should not be applied mechanically. Instead, they should be covered according to the needs of the relevant stakeholders and the context of the evaluation. More or less time and more or fewer resources may be devoted to the evaluative analysis for each criterion depending on the evaluation purpose. Data availability, resource constraints, and timing and methodological considerations may also influence how (and whether) a particular criterion is covered (OECD 2019: 6).

The six evaluation criteria are intended to ‘be a complete set that fully reflects all important concepts to be covered in evaluations ... across the sustainable development and humanitarian fields’ (OECD 2021). It is useful to emphasise that the 2021 OECD DAC guidance is explicitly intended to apply to humanitarian action.

The adaptation of the original OECD DAC criteria has been characterised as conservative. While the approach has been applauded by some, it has been critiqued by others as ‘business as usual’ (Patton 2020: 59).

Box 4: Timeline of foundational documents informing criteria for evaluating humanitarian action

- 1991:** OECD DAC Principles for Evaluation of Development Assistance
- 1998:** Relief and Rehabilitation Network Good Practice Review No. 7 ‘Evaluating Humanitarian Assistance Programmes in Complex Emergencies’ (Hallam) (prepared in tandem with 1999 OECD DAC Guidance)
- 1999:** OECD DAC Guidance for Evaluating Humanitarian Action in Complex Emergencies
- 2002:** OECD DAC Glossary of Key Terms
- 2006:** ALNAP Evaluating Humanitarian Action Using the OECD DAC Criteria: An ALNAP guide for humanitarian agencies
- 2010:** OECD DAC Quality Standards for Development Evaluation
- 2014:** Core Humanitarian Standard on Quality and Accountability
- 2016:** ALNAP Evaluation of Humanitarian Action Guide
- 2018:** Core Humanitarian Standard Updated Guidance Notes and Indicators
- 2018:** Inter-Agency Standing Committee Inter-Agency Humanitarian Evaluation process guidelines
- 2019:** OECD DAC Better Criteria for Better Evaluation: Revised Evaluation Criteria Definitions and Principles for Use
- 2021:** OECD DAC Applying Evaluation Criteria Thoughtfully

Strengths of the OECD DAC evaluation criteria

The OECD DAC evaluation criteria are the pre-eminent criteria for evaluating development and humanitarian assistance (Kennedy-Chouane 2020, Picciotto 2013, cited in Patton 2020). They are widely applied – even more so than originally expected (Lundgren 2017). This has important advantages. It makes meta-evaluation easier, helping to capture common weakness in humanitarian action, and creates common terminology making it easier for evaluators to work together (ALNAP 2016).

Perhaps the most obvious indication of the OECD DAC evaluation criteria's strength is their durability across decades. Lundgren (2017) observes that the criteria are relatively easy to understand and use, and cover the key issues that are important to consider when assessing the performance of an intervention. This observation is supported by those engaged in evaluations of development interventions, who note the following strengths of the OECD DAC evaluation criteria (OECD 2018):

- Universal acceptance and use, creating a common language and understanding
- standardisation, with scope for comparability across evaluations and contexts
- adaptable and flexible, able to be tailored for different evaluations and contexts
- comprehensive, covering key areas required for accountability and learning
- simple, clear and useful
- neutral and universal across cultures and political contexts, institutions and policy areas
- concise and feasible; the limited number of criteria is more realistic to implement.

Issues common to OECD DAC evaluation criteria

As can be expected, such popular and widely applied criteria have been subject to critique. Critique of individual criteria are included within the relevant chapters of this report. Some of the broader criticisms include:

- An inability to effectively evaluate **transformational change** (Patton 2020; Ofir 2017). Given the critical need for 'fundamental systems transformation' to address 'existential threats to the future of humanity', such as climate change and the COVID-19 pandemic, Patton (2020: 53) is a strong advocate for different criteria, finding that 'the greatest danger for evaluators in times of turbulence is not the turbulence—it is to act with yesterday's criteria'. Related critiques have found that the criteria are limited in their application to **system-wide performance** or organisational performance across a range of humanitarian contexts (Darcy and Dillon 2020), and in their ability to reflect complexity and synergy, or evaluations of strategy, policy or corporate approaches (OECD 2018).
- Insufficient focus on **gender, equity or human rights** concerns (OECD 2018). This is likely related to a broader lack of high-level commitment to inclusion in humanitarian action, as well as a lack of tools to evaluate progress on inclusion (Lough et al. 2022). The criteria were similarly critiqued in advance of the update in 2019 for insufficient reference to **important norms** (Ofir 2017), such as the 2030 Agenda for Sustainable Development (DEval 2018; OECD 2018).

This paper has also identified common issues that recurred across multiple criteria and that would be useful to consider in future humanitarian guidance. A foundational question is **how closely the new guidance should align to the OECD DAC guidance and the adaptation to the criteria in 2019**. Other common issues include:

- The importance of **positionality** and whose perspective is used in defining the evaluative questions and who conducts the evaluation. This is reflected in increasing calls for the **decolonisation of evaluation**. Chilisa and Mertens (2021) find that evaluation is dominated by Western culture and approaches; that evaluation has ‘become a colonial prejudice that reinforces uneven and biased power relations’ (2021: 242). They call for the use of indigenous approaches to evaluation to challenge the culture, ethics and values of an evaluation’s terms of reference, conduct and findings. Ofir (2017) applies this directly to the OECD DAC criteria, calling out insufficient recognition of the **importance of culture and cultural differences** (Ofir 2017). The review of the OECD DAC criteria identified related concerns regarding the uncontextualised implementation of the criteria (OECD 2018).

There is substantive work to draw on to progress the decolonisation of evaluation. Chilisa and Mertens (2021), for example, note that Indigenous groups around the world have developed approaches appropriate to their culture, such as the Made in Africa approach to evaluation.

Literature across the criteria indicate that evaluation is impacted by positionality; whose perspective is used in defining the evaluative questions and who conducts the evaluation. For example, who defines what is effective, and how it is measured, may have a significant impact on findings, even if affected populations’ views are incorporated as an addition to or in place of international perspectives.

- The **desire for more guidance to improve standardisation** (Darcy and Dillon 2020), **while maintaining flexibility** in application (DEval 2018). There are challenges in providing guidance that supports evaluations that produce findings that can be compared across contexts and over time.
- Finally, the **utility and application of the criteria** are variable and depend largely on the type of programme and organisation and the intent of the evaluation. Therefore, it may be helpful to provide guidance on which criteria are most useful in which context or how to adapt the criteria as needed for the type of evaluation and specific context.

The ALNAP guidance

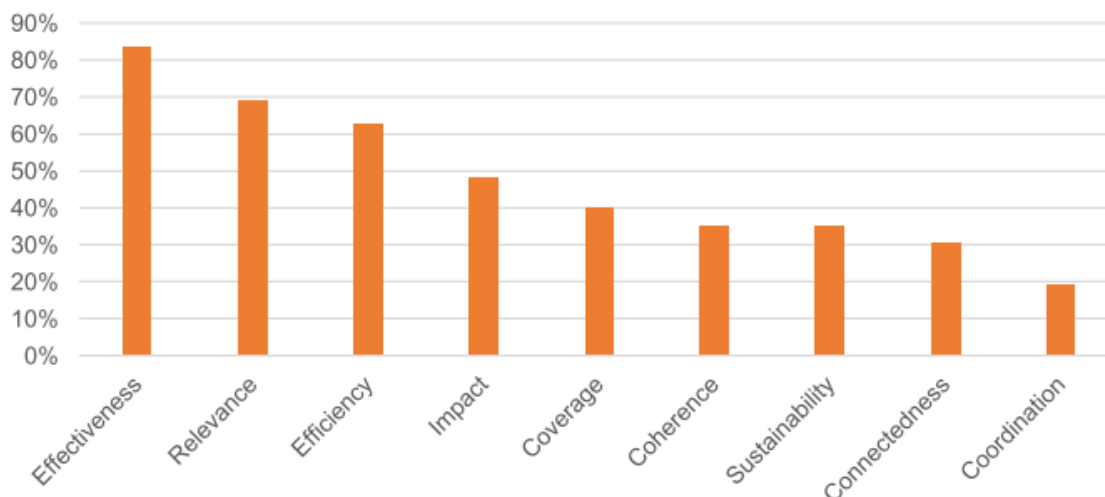
ALNAP’s 2006 guide

ALNAP’s 2006 guide was developed to assist with the quality of humanitarian evaluations and to help evaluation practitioners and managers interpret the OECD DAC criteria in a humanitarian context. It applies the criteria established by the OECD DAC for evaluating humanitarian action in complex emergencies (OECD 1999) to all humanitarian action: **effectiveness, appropriateness/relevance, efficiency, impact, coverage, coherence and connectedness**.

The 2006 ALNAP guide outlines definitions for those criteria, key messages, issues to consider in the application of criteria, and examples of good practice. The 2006 guide is one of ALNAP’s most popular publications, perhaps because more evaluations of humanitarian action are being conducted than ever before and the OECD DAC criteria remains one of the most-used frameworks (Darcy and Dillon 2020).

Of the seven criteria in ALNAP's 2006 guide, a sample of 120 humanitarian evaluations found the most commonly applied criteria to be effectiveness, appropriateness/relevance and efficiency, followed by impact, coverage and coherence respectively. Though connectedness was the least used criteria, it was still applied in 30% of the evaluations reviewed.

Figure 1: Application of the OECD DAC criteria across a sample of 120 evaluations



Source: Darcy and Dillon 2020

Additional ALNAP guidance

ALNAP Evaluation of humanitarian action guide 2016

In 2016, ALNAP published its [Evaluation of Humanitarian Action \(EHA\) guide](#), in response to demand for more extensive guidance on how to evaluate humanitarian interventions. The EHA guide explains how to plan, design, implement and utilise humanitarian evaluation. It recommends developing evaluation questions first, and then checking them against an appropriate evaluation framework, such as the OECD DAC criteria.

The 2016 EHA guide touches briefly on the criteria and definitions, adapting them slightly from the 2006 guide and including sample evaluation questions. The key adaptations from 2006 are the addition of 'coordination' as a separate criterion and replacing 'relevance/appropriateness' with 'appropriateness' only. These changes were presented to assist evaluators and did not reflect a formal review of the criteria.

ALNAP State of the Humanitarian System (SOHS) reports

The SOHS is an independent study by ALNAP published every two to four years to assess the overall performance and progress of the humanitarian system. It includes a list of criteria and definitions against which the assessment is structured. These are adapted from the OECD DAC criteria in consultation with the SOHS Methods Group.

For the purpose of the 2018 and 2022 SOHS reports, ALNAP included three criteria usually assessed under other OECD DAC criteria: **accountability and participation**, **sufficiency** and **complementarity** (ALNAP 2018a).³ Accountability and participation were

³ The 2018 SOHS report is referenced in this paper as the 2022 report was not yet available at the time of writing. The 2022 report applies the same criteria.

added to emphasise the central perspectives of the affected population; not just as tools to achieve relevance (the criterion under which it is otherwise usually evaluated), but as their own critical component of performance. Complementarity was added to distinguish the international humanitarian system's relationship with development and peacebuilding activities from its relationship with local and national actors. ALNAP therefore focused the criterion of connectedness on links with development and peacebuilding, and the complementarity criterion on links with local and national actors.

The SOHS defines its additional three criteria as follows:

- **Accountability and participation:** the degree to which actors within the international humanitarian system can be held to account by crisis-affected people and the degree to which crisis-affected people are able to influence decisions related to assistance and protection.
- **Sufficiency:** the degree to which the resources available to the international humanitarian system are sufficient to cover humanitarian needs.
- **Complementarity:** the degree to which the international humanitarian system recognises and supports the capacities of national and local actors, in particular governments and civil society organisations.

ALNAP guidance for Real-Time Evaluations

ALNAP has published two papers on applying the criteria to Real-Time Evaluations, or Real-Time Learning: [Real-Time Evaluations of humanitarian action: An ALNAP guide pilot version](#) by Cosgrave, Ramalingam and Beck in 2009, and [From Real-Time Evaluation to Real-Time Learning: Exploring new approaches from the COVID-19 response](#) by Buchanan-Smith and Morrison-Métois in 2021. Both papers discuss the OECD DAC criteria, while also referring readers to other frameworks and standards. While Real-Time Evaluations take a different approach from more 'traditional' forms of evaluation, they are a useful point of reference given their applicability and use in complex crises.

The 2009 *Real-Time Evaluations of humanitarian action* guide includes two additional criteria in its discussion of the OECD DAC criteria. The first, coordination (included in ALNAP's 2006 guide under effectiveness), refers to an organisation's own internal management and coordination processes as well as its contribution to host government and international coordination mechanisms. The second additional criterion, depending on the circumstance of the evaluation, is proportionality, or the extent to which the response is both proportional to the needs of the affected people and comparable with other humanitarian responses (Cosgrave et al. 2009).

ALNAP's 2021 paper on Real-Time Learning (Buchanan-Smith and Morrison-Métois 2021) found that the OECD DAC criteria has not been the primary framework for most Real-Time Evaluations and Real-Time Learning exercises it reviewed (undertaken during the COVID-19 response from 2020 to 2021). While some UN agencies applied the OECD DAC evaluation criteria, international NGOs tended to make more use of the Core Humanitarian Standard (see [below](#)).

Additional frameworks and standards

The advantage of using the OECD DAC criteria and the ALNAP interpretation of the criteria for humanitarian action is that it supports comparability across evaluations and provides an opportunity for meta-analyses and syntheses (Cosgrave et al. 2009). There are, however, other frameworks and standards that could be applied. Many agencies have developed their own guidance for evaluations, and have defined the criteria slightly differently, applied other standards or developed their own. This section briefly discusses other sector-wide accountability standards. ALNAP's *Evaluating humanitarian action guide* (2016) provided a brief overview of alternative normative frameworks, standards, guides and manuals to consider when designing evaluations. As this paper is focused on the OECD DAC criteria, we have included other sector-wide standards that are often used in addition to or, in some cases, in place of the common criteria. This paper does not fully explore alternative evaluation criteria, a subject which would merit its own research.

It is helpful to understand that any set of criteria is a reflection of the concerns of the time, intended to help address recurring areas of weakness in the assistance provided (Cosgrave et al. 2009). For example, the criteria recommended by the OECD DAC for complex emergencies in 1999 (see above) reflected the key issues for humanitarian action in the 1990s (ALNAP 2016). Perceptions of the key issues and biggest challenges evolve over time. This can be seen in the contemporary literature, guidance and evaluations of humanitarian action, where themes such as inclusion, localisation and accountability to affected populations are more prominent than they were in 2006.

Sphere

The Sphere Handbook (2018) comprises the Humanitarian Charter, Protection Principles, Core Humanitarian Standard (CHS) and minimum standards for four technical areas of humanitarian action: water supply, sanitation and hygiene promotion; food security and nutrition; shelter and settlement; and health. It is one of the most widely recognised and applied standards for humanitarian action (Sphere Association 2018).

The Sphere Project produced guidance in 2015 for applying the Sphere standards in monitoring and evaluation (Mountfield 2015). It notes the different lens provided by the OECD DAC criteria as compared to the 'participatory approaches' implied by the *Sphere Handbook* (2018). The Charter, Protection Principles and CHS are applicable across the criteria, while the technical chapters most usefully relate to relevance, effectiveness and impact.

This paper will focus on the CHS as the more commonly applied standard to humanitarian evaluations.

Core Humanitarian Standard

The Core Humanitarian Standard on Quality and Accountability (CHS) is a set of nine commitments to improve the quality and effectiveness of humanitarian assistance (CHS 2014 and 2018).⁴ It places communities and people affected by crisis at the centre of humanitarian action, aiming to improve the accountability of humanitarian action to those it is intended to assist.

⁴ The Core Humanitarian Standard is undergoing a revision at the time of writing this report.

The CHS appears to be increasingly important to evaluation. In a review of approaches to Real-Time Evaluation and Learning, Buchanan-Smith and Morrison-Métois (2021) found that international NGOs have tended towards application of the CHS as a framework over application of the OECD DAC criteria. The Disasters Emergency Committee (DEC), for example, used the CHS as a key reference point for Real-Time Learning, while World Vision drew on both the CHS and the OECD DAC criteria to develop its own blended criteria. Key informants reported that the CHS was more oriented towards affected people, clearer and easier to apply (Buchanan-Smith and Morrison-Métois 2021).

Given the reported utility of the CHS, this paper maps the CHS commitments against ALNAP's 2006 criteria where they are directly relevant.

Inter-Agency Standing Committee (IASC)

An Inter-Agency Humanitarian Evaluation (IAHE) is an independent evaluation conducted by member organisations of the IASC of a humanitarian response (OCHA 2018). The IASC is made up of nine UN agencies and nine standing invitees including the Red Cross Red Crescent Movement, NGO consortia, the World Bank and human rights agencies. IAHEs are usually initiated through inclusion in the IAHE Steering Group's Workplan and launched by the Emergency Relief Coordinator (IASC n.d.).

As a multi-agency humanitarian evaluation process, the IAHE evaluation guidelines, published by the Office for the Coordination of Humanitarian Affairs (OCHA) in 2018, are a useful point of reference for this review. The IAHE guidelines establish a six-question framework for evaluating a specific crisis response, which includes some OECD DAC criteria and additional criteria: relevance, effectiveness, sustainability, partnerships, localisation and coordination. It frames thematic evaluations around just three themes: relevance, effectiveness and the delivery of collective outcomes (OCHA 2018). This paper maps the IAHE criteria for specific crisis responses against the ALNAP 2006 criteria.

As there are variations in the application of the OECD DAC criteria, which are demonstrated throughout this paper, there are also variations in application of the IAHE guidelines. The greatest divergence from the IAHE guidelines can be seen in the terms of reference for the evaluation of the COVID-19 response (IASC 2021), which applies the OECD DAC criteria outlined in ALNAP's 2006 guide. The four IAHE evaluations conducted since the guidelines were published in 2018 have all included appropriateness/relevance, effectiveness and coordination (IASC 2022, 2020, 2020a, 2019). The IAHE thematic evaluation of gender was the only evaluation to add coherence (IASC 2020a). The others also included partnerships, with the IAHE evaluation of Mozambique including connectedness, coverage and localisation, and the evaluation of IAHE evaluation in Ethiopia (2018) including localisation and sustainability.

Table 2: IAHE application of criteria in evaluations since 2018

IAHE Criteria	Yemen 2022	Gender equality 2020	Mozambique 2020	Ethiopia 2018
Effectiveness	Effectiveness	Effectiveness	Effectiveness	Effectiveness
Relevance	Appropriateness	Relevance	Appropriateness	Relevance
Coordination	Coordination (with Partnerships)	Coordination	Coordination	Coordination
Partnerships	Partnerships (with Coordination)		Partnerships	Partnerships (with Localisation)
Localisation			Localisation	Localisation (with Partnerships)
Sustainability				Sustainability
Other criteria		Coherence	Coverage Connectedness	

Source: Compiled by author. Data drawn from the IAHE evaluation reports.

Sector-wide guidance and standards: A summary

This paper reviews each criterion using the ALNAP 2006 guide as the primary reference point. Each criterion is compared against sector-wide guidance and standards. [Table 3](#) lists the criteria used by each of the sector-wide guidance documents reviewed in this paper. It demonstrates some evolution through ALNAP guidance, key points of difference with the OECD DAC criteria and IAHE guidelines, and substantive difference with the CHS. Criteria are placed on the same line where they are the same or have some similarities with other criteria.

Table 3: Evaluation criteria by guide or standard

ALNAP OECD DAC guidance 2006	ALNAP EHA guide 2016	ALNAP SOHS 2018	OECD DAC 2019 (Development)	IASC's IAHE Process Guidelines	Core Humanitarian Standard (in brief)
Effectiveness	Effectiveness	Effectiveness	Effectiveness	Effectiveness	Effective and timely
Relevance/ Appropriateness	Appropriateness	Relevance and appropriateness	Relevance	Relevance	Appropriate and relevant
Efficiency	Efficiency	Efficiency	Efficiency		Resources managed and used responsibly
Impact	Impact	Impact	Impact		
Coherence	Coherence	Coherence	Coherence		Coordinated and complementary
Coverage	Coverage	Coverage			
Connectedness	Connectedness	Connectedness	Sustainability	Sustainability	Strengthens local capacities and avoids negative effects
	Coordination			Coordination	
		Complementarity		Localisation	
				Partnerships	
		Sufficiency			
		Accountability and participation			Based on communication, participation and feedback; complaints are welcome and addressed
					Humanitarians continually learn and improve
					Staff supported and treated fairly

A large, stylized teal number '2' is centered on a solid teal background. The number is composed of two overlapping shapes: a semi-circle at the top and a trapezoidal shape at the bottom. The word 'EFFECTIVENESS' is written in white, bold, uppercase letters across the middle of the number.

EFFECTIVENESS



EFFECTIVENESS

MOST-USED OECD DAC CRITERION

Definitions

ALNAP Guide 2006		OECD DAC Criteria 2019	
<p>Effectiveness: Measures the extent to which an activity achieves its purpose, or whether this can be expected to happen on the basis of the outputs. Implicit within the criterion of effectiveness is timeliness.</p>		<p>Effectiveness: Is the intervention achieving its objectives? The extent to which the intervention achieved, or is expected to achieve, its objectives, and its results, including any differential results across groups.</p>	
ALNAP EHA Guide 2016	ALNAP SOHS 2018	IASC's IAHE 2018	CHS 2018
<p>Effectiveness:</p> <p>The extent to which an activity achieves its purpose, or whether this can be expected to happen on the basis of the outputs.</p>	<p>Effectiveness:</p> <p>The degree to which humanitarian operations meet their stated objectives, in a timely manner and at an acceptable level of quality.</p>	<p>Effectiveness:</p> <p>To what extent were the results (in terms of assistance delivery as articulated in the Humanitarian Response Plan) achieved and to what extent were they effective in meeting the needs of the most vulnerable.</p>	<p>Commitment 2:</p> <p>Communities and people affected by crisis have access to the humanitarian assistance they need at the right time.</p> <p>Quality criterion:</p> <p>Humanitarian response is effective and timely.</p>

Sector-wide guidance and standards

ALNAP

The 2006 ALNAP guidance focuses the evaluation of effectiveness on an analysis of the extent to which objectives (rather than activities), stated in organisational planning documents, are met. It defines objectives as 'intermediate between outputs and outcomes' (51). Evaluations need to examine inclusion, using disaggregated data to understand who uses and benefits from the resources provided. Timeliness is a key element: whether interventions adequately supported the affected population at different phases of the crisis.

The 2006 guidance further states that evaluation should include an analysis of how objectives were formulated and who participated; in particular, whether and how primary stakeholders participated in the intervention's design. The perspectives of primary stakeholders, as compared to the perspectives of other humanitarian actors such as organisation staff, should be a key element in determining whether interventions have met their objectives.

The 2006 ALNAP guidance explicitly **includes the criterion of coordination under the effectiveness criterion**. This reflects the OECD DAC 1999 *Guidance for evaluating humanitarian assistance in complex emergencies*, which encouraged the evaluation of coordination separately and gave the option for it to be considered under effectiveness. ALNAP in 2006 noted three key issues to consider concerning coordination:

1) the multiplicity of actors, 2) the role of the host government and other national and local institutions and 3) non-traditional partners.

ALNAP's 2016 definition provides space for evaluators to determine whether an activity can be expected to achieve its purpose, on the basis of the outputs. Its example evaluation question reinforces the importance of timeliness: 'to what extent did the programme achieve its objectives, including the timely delivery of relief assistance'. It includes coordination as a separate criterion, while noting that it is often incorporated within the effectiveness criterion. It defines coordination as the extent to which the interventions of different actors are harmonised with each other, promote synergy and avoid gaps, duplication and resource conflicts.

The 2018 SOHS also emphasises timeliness in its definitions and introduced the benchmark of 'an acceptable level of quality'. It noted that many evaluations reviewed for the report did not make a clear statement about whether or not objectives had been achieved. The SOHS looked specifically at both life-saving objectives and objectives related to protection. In terms of factors affecting effectiveness, the SOHS looked at funding, leadership, coordination, preparedness, organisational factors and staffing.

OECD DAC criteria

As part of the definition of effectiveness outlined at the beginning of this chapter, the OECD DAC Guidance (2021: 52) includes the following note:

Analysis of effectiveness involves taking account of the relative importance of the objectives or results. The term effectiveness is also used as an aggregate measure of the extent to which an intervention has achieved or is expected to achieve relevant and sustainable impacts, efficiently and coherently.

The 2019 OECD DAC criteria ask: *Is the intervention achieving its objectives?* It distinguishes effectiveness from impact – with effectiveness focusing more closely on attributable results than impact, which looks at higher-order effects. The phrasing is intended to allow flexibility for evaluations to focus on the objectives and/or results that are of most interest. The 2021 guidance encourages evaluators to identify unintended effects, both positive and negative.

The addition of 'results' and 'differential results' from the previous OECD DAC definition of effectiveness is intended to encourage an in-depth consideration of equity between different groups – through an understanding of the distribution of results across different groups, and the intended and unintended effects on different groups. This supports the policy priority to 'leave no one behind', and is to be considered regardless of whether equity is a specific objective of the intervention.

The 2021 guidance identifies four key elements for analysis: achievement of objectives, the varying importance of objectives and results, differential results across groups and understanding the factors that influence outcomes.

Inter-Agency Standing Committee's 2018 IAHE Guidance

IAHE guidelines (OCHA 2018) introduced to the effectiveness criterion a focus on the needs of the most vulnerable groups affected by crisis.

Core Humanitarian Standard

The effectiveness of humanitarian interventions is addressed primarily under the Core Humanitarian Standard Commitment 2, explicit in its quality criterion that humanitarian response is effective and timely. In its indicators, the CHS reflects the IASC's focus on vulnerable groups and the broader attention to timeliness, and the 2018 SOHS's focus on protection. It specifies the importance of assessing whether the response met its timing, quality and quantity objectives.

Key issues arising

Effectiveness is the most applied of all the OECD DAC evaluation criteria, appearing in more than 80% of the 549 studies reviewed for the 2018 SOHS report (Darcy and Dillon 2020). The literature on evaluating effectiveness is, however, less extensive (Ibid.). The multisectoral guidance outlined above found differing emphasis on coordination, timeliness, quality, protection and vulnerable groups/equity/inclusion. The literature focuses on challenges in evaluating effectiveness: determining objectives, measuring the effects of interventions, and attribution (ALNAP 2018a; Darcy and Dillon 2020). In reviews of the criteria in advance of the OECD DAC's update in 2019, DEval, the German Institute for Development Evaluation (2018), noted challenges in implementing effectiveness, including that the levels of the objectives examined often differ in practice. The OECD DAC consultation processes found key conceptual challenges, including: insufficient consideration of external effects, limited requirement to assess quality and the (previous) inability to make changes to intended results in response to changing contexts (OECD 2018).

Challenges in determining and measuring objectives

Effectiveness is generally defined in terms of achieving objectives. What constitutes an objective is less clear. For example, the ALNAP 2006 guide defines objectives as *intermediate between outputs and outcomes* (51), while the US Bureau for Humanitarian Assistance (BHA) guidance (2022) asks as a sub-question for effectiveness: 'to what extent do the activity's interventions appear to have achieved their intended *outputs and outcomes*?' While this allows flexibility in terms of the level assessed as an 'objective', it also facilitates variable application. The 2018 SOHS report found some programmes expressed objectives as outputs 'rather than outcomes', while others set objectives that were aspirational and unlikely to be achieved.

Heider (2017) further criticises this approach as too lenient, in that it allows implementers to define the terms against which their intervention will be assessed. Darcy and Dillon (2020) argue that effectiveness must have as its reference point an effect, or objectives, situated 'in the real world'; relating to a change in people's situation, and not be based on the delivery of outputs and assumptions about the resulting effects. In other words, objectives must be defined in terms of change that is external to the project or action. This could include the contribution of humanitarian action to the prevention or mitigation of feared outcomes, such as threats to life, health, livelihoods or security.

Patton (2020), in the context of evaluating transformation, recommends changing the evaluation focus from whether transformation occurred, to transformational engagement. That is, assessing the trajectory towards transformation; the contribution of the intervention that helps to achieve systemic sustainable change with large-scale impact.

In practice, evaluators are challenged in their ability to find the evidence needed to determine whether objectives have been met, especially objectives outside the immediate results chain, and attribute change to a single actor or even combined humanitarian action (OECD 2007; Heider 2017; ALNAP 2018a; Darcy and Dillon 2020).

As mentioned, the updated OECD DAC guidelines (2021) for effectiveness incorporate the need to unpack the **distributional effects** of interventions, as well as **unintended effects**. Heider (2017) further emphasises this point:

Distributional effects of interventions, whether explicitly part of the intended outcomes or not, need to be assessed if we are serious about goals like 'no-one left behind'. ... If the analysis of the intended and unintended effects is differentiated by different stakeholder groups, we can get a better understanding of the actual effects or impacts of interventions.

The ALNAP 2006 guide emphasises the importance of understanding the perspectives of primary stakeholders in determining whether interventions have met their objectives. In their paper focused on the criterion of relevance/appropriateness, Field (2016) notes that, from the perspective of affected populations, perceptions of what is effective can differ within groups and across time. For example, one group prioritised physiological and material needs, while another prioritised a return to social normalcy and non-life-saving items.

Elements of effectiveness: Coordination, timeliness, inclusion, quality

The literature is less explicit on the extent to which the evaluation of effectiveness in humanitarian action should emphasise important elements, in particular: **coordination, timeliness, inclusion and quality**. These elements were identified through the multisectoral guidance: ALNAP in 2006 and 2018 includes coordination; the OECD DAC in 2021 makes prominent inclusion through encouraging analysis of differential results across groups – this is echoed in CHS and IAHE definitions emphasising communities and vulnerable groups; and quality is specified in the ALNAP (2018a) definition.

Organisational guidance

The analysis in [Table 4](#) includes a review of the definitions, explanatory text and suggested questions available in the guidance. In some organisational guidance, discussion of evaluation criteria is limited to a definition. The following analysis does not suggest that these agencies do not consider these elements to be important to effectiveness, just that it is not explicit in their guidance.

Reflecting the literature, there is wide variance in the terminology used to explain the level of objective or result expected to be assessed during evaluations of the effectiveness of humanitarian action. The most common terminology used is 'objectives': applied across six of the 10 individual agency guidance documents reviewed; half of these use 'objectives' in combination with other terminology. Results, immediate results, purpose, outputs and outcomes are terms also used. The World Food Programme (WFP), echoing the OECD DAC's definition of 'results', explicitly includes both intended and unintended and positive and negative results.

The inclusion of the four elements – inclusion, coordination, timeliness and quality – is more common in guidance that provides an explanation beyond the definition of effectiveness. Guidance by Save the Children, ACT Alliance and USAID BHA, for example, limits discussion of the criteria to definitions.

Inclusion is incorporated in 40% of the organisational guidance documents reviewed, and as a separate criterion in UNICEF's guidance for equity-focused evaluations. This perhaps is reflective of the contemporary nature of some guidance. All guidance reviewed that was published from 2019 onwards incorporates inclusion in some manner; for example, the WFP guidance published in 2021 applies the OECD DAC (2019: 9) definition of effectiveness, which includes an assessment of 'differential analysis across groups'. And for others, inclusion may be the focus of the evaluation guidance, as is the case with UNICEF's guidance for equity-focused evaluations.

Coordination and timeliness both appear in three of 10 of the organisational guidance documents reviewed, though only one of these (Médecins Sans Frontières) included coordination as a component of effectiveness. The others included coordination as a separate criterion (UNICEF and UNFPA). Timeliness was made explicit by IFRC, Catholic Relief Services (CRS) and Médecins Sans Frontières (MSF). Quality was not explicitly incorporated into guidance for assessing effectiveness except by CRS, which mentioned quality in relation to partnerships. This may be because the quality of the results assessed are assumed to be incorporated into the definitions of the results themselves.

Table 4: Organisational guidance on effectiveness

Organisation	Objectives	Inclusion (differentiated results)	Coordination	Timeliness
ALNAP (2006)	Objectives, defined as intermediate between outputs and outcomes	✓		✓
OECD (2021)	Objectives and results (results = output, outcome or impact – intended or unintended, positive or negative)		X	
IFRC (2011)	Intended, immediate results	X	✓	✓
Save the Children (n.d.)	Objectives	X	X	X
ACT Alliance (2012)	Objectives	X	X	X
Catholic Relief Services (2012)	Outputs	X	X	✓
WFP (2021)	Objectives and results (intended or unintended, positive or negative), outputs and outcomes	✓	X	X
Médecins Sans Frontières (2017)	Objectives and purpose, based on activities, outputs and outcomes	X	✓	✓
UNICEF (2014)	Objectives	Additional criteria: equity, gender equality	Additional criterion	X
USAID BHA (2022)	Outputs and outcomes	✓ (Individuals and households)	X	X
UNICEF – Equity Evaluations (2011)	Equity-focused results, in relation to equity-focused objectives	✓	X	X

UNFPA (2019)	Country programme outputs, and the extent to which these contribute to outcomes (comparing goals, outcomes, outputs and results)	✓	Additional criterion	X
--------------	--	---	----------------------	---

Application in evaluations

The same analysis was applied across 10 evaluations of humanitarian responses (as shown in [Table 5](#)) and provides similar trends regarding terminology. The term 'objective' was used in 50% of the evaluations, with the terms results, outcomes and outputs remaining common.

In terms of the additional elements to effectiveness identified – inclusion, coordination, timeliness and quality – these were assessed more often than was perhaps to be expected given the guidance. Each element was considered in almost all the evaluations, though not always under the effectiveness criterion. Inclusion was perhaps the most common element, considered in eight of the 10 evaluations under effectiveness and for the remaining two evaluations considered separately. Coordination and timeliness were evaluated about half the time under effectiveness, and half the time as separate criteria. Timeliness was evaluated as part of the efficiency criterion in 30% of the evaluations reviewed. Quality was more difficult to assess; while it was often implicit in the evaluation findings, it was not often explicitly mentioned using the word 'quality'.

Table 5: Organisations evaluations on effectiveness

Organisation	Title	Objectives	Inclusion (dis-aggregated results)	Coordination	Timeliness	Quality
UNHCR	Evaluation of UNHCR's Level-3 Emergency Response to Cyclone Idai (2021)	HRP objectives, protection objectives	✓	✓	✓	Not explicitly
IAHE	Evaluation of the Drought Response in Ethiopia 2015–2018 (2019)	Intended results/goals. Also looked at outcomes and outputs	As fairness and impartiality	As separate criterion	As separate criterion	✓
AHP	Response to the 2018 Papua New Guinea Highlands Earthquake (2019)	Outcomes and outputs	✓	X	Under efficiency	✓
DFID	Evaluation of DFID's Humanitarian Response to Typhoon Haiyan (Yolanda) (2015)	DFID priority outcomes	✓	✓	✓	✓
IFRC	Evaluation of IFRC West Africa Ebola Viral Disease Appeal Response Sierra Leone and Liberia (2018)	Outcomes and outputs	✓	✓	Under efficiency	Not explicitly
DEC	East Africa Response Review: South Sudan (2017)	Output objective. Also mentions results and outcomes	Under other criteria	✓	As own criteria (overall approach and timeliness)	✓

UNICEF	Review of the UNICEF L2 Response in Venezuela (2020)	Objectives and results, including differential results across groups.	✓	Separately (under leadership)	Under efficiency	✓
UNHCR	Independent evaluation of UNHCR's emergency response to the Rohingya refugees influx in Bangladesh 2017–2018 (2018)	Terminology not explicit. Included mortality and morbidity	Separately	Separately	✓	✓ (considered standards)
DFAT	Independent Review – Iraq Humanitarian and Stabilisation Package (2020)	Objectives, results, outcomes	✓	Separately	Separately	Not explicitly
European Union	Evaluation of the European Union's Humanitarian Response to the refugee crisis in Turkey (2019)	Objectives (findings specified strategic objective)	✓	Separately	✓	Yes, for some interventions

Summary and questions to explore

Effectiveness is the most applied of all the OECD DAC evaluation criteria. Effectiveness is generally defined in terms of achieving objectives. Challenges were identified in determining what constitutes an objective, and how they are best measured in humanitarian contexts. ALNAP (2006) defines objectives as 'intermediate between outputs and outcomes' (51). Darcy and Dillon (2020) argue that objectives must be defined in terms of a change that is external to the project or action. Organisational evaluation guidance varies widely in the terminology used to explain the level of objective expected to be assessed, and includes the terms: objectives, results, immediate results, purpose, outputs and outcomes.

- **Is more precise guidance needed on what constitutes an objective?**

A further consideration is whether the intervention's original objectives remain relevant in changing contexts and therefore whether their achievement is an indicator of effectiveness (with links to the exploration of adaptive management under cross-cutting themes later in this paper).

- **Should objectives be tested for relevance in the current context before they are used as an indicator of effectiveness (considering adaptive management)?**

Sector-wide guidance variously emphasises coordination, timeliness, quality, protection and vulnerable groups/equity/inclusion as key components of effectiveness. ALNAP (2006) includes coordination under the effectiveness criterion. The WFP's guidance, echoing the OECD DAC's definition of 'results', explicitly includes both intended and unintended and positive and negative results.

- **Should coordination, timeliness, quality, protection and vulnerable groups/equity/inclusion be considered elements of effectiveness? If**

not, how should they be considered in relation to the other evaluation criteria?

- **Is it necessary to specify and explicitly measure unintended results and to explore negative, as well as positive, results? What are the challenges to doing this effectively in practice and are there situations in which this is not feasible?**

In practice, evaluators are often challenged in their ability to find the evidence needed to determine whether objectives have been met and whether change can be attributed to a single actor or even to combined humanitarian action (OECD 2007; Heider 2017; ALNAP 2018a; Darcy and Dillon 2020).

- **How can guidance help address challenges in assessing attribution and measuring effectiveness?**

ALNAP (2006) recommends that evaluations analyse whether and how primary stakeholders participated in the intervention's design and the formulation of objectives and include the perspectives of primary stakeholders – as compared to other humanitarian actors such as organisation staff – in determining whether interventions have met their objectives.

- **How important is the perspective of primary stakeholders in designing and evaluating effectiveness? If important, how can guidance enable this?**



**RELEVANCE/
APPROPRIATENESS**



RELEVANCE/APPROPRIATENESS

SECOND MOST-USED OECD DAC CRITERION

Definitions

ALNAP Guide 2006		OECD DAC Criteria 2019	
<p>Relevance/Appropriateness: Relevance is concerned with assessing whether the project is in line with local needs and priorities (as well as donor policy). Appropriateness is the tailoring of humanitarian activities to local needs, increasing ownership, accountability and cost-effectiveness accordingly.</p>		<p>Relevance: Is the intervention doing the right things? The extent to which the intervention's objectives and design respond to beneficiaries', global, country and partner/institutions' needs, policies and priorities and continue to do so if circumstances change.</p>	
ALNAP EHA Guide 2016	ALNAP SOHS 2018	IASC's IAHE 2018	CHS 2018
<p>Appropriateness: The extent to which humanitarian activities are tailored to local needs, increasing ownership, accountability and cost-effectiveness accordingly.</p>	<p>Relevance and Appropriateness: The degree to which the assistance and protection that the international humanitarian system provides addresses the most important needs of recipients (as judged both by humanitarian professionals and by crisis-affected people themselves).</p>	<p>Relevance: To what extent have the objectives set out in the Humanitarian Response Plan (HRP) or similar plan been based on identified needs of the most vulnerable groups affected by the crisis?</p>	<p>Commitment 1: Communities and people affected by crisis receive assistance appropriate and relevant to their needs.</p>

Sector-wide guidance and standards

ALNAP

In 2006, ALNAP included appropriateness with relevance, reflecting the 1999 OECD DAC guidance for evaluating humanitarian assistance in complex emergencies. Relevance looks at the 'why' – the purpose of the programme – while appropriateness is more about the activities, or the 'how' (ALNAP 2020). ALNAP's 2006 guide refers to the need for contextual analysis, participatory needs assessment of the differentiated needs of the affected population (women, girls, boys and men, different social groups) and support for livelihoods and capacities. It emphasises cultural appropriateness, questions of cultural relativism and the institutional capacity of the humanitarian organisation for ensuring a relevant and appropriate response. It suggests that many cross-cutting themes, such as gender and participation, could be considered under this criterion.

More recently, ALNAP (2016: 268) has emphasised the importance of the perspectives of crisis-affected people to assessing relevance: *'it is only by engaging with the affected population and hearing their perspectives and views that we can know whether humanitarian programmes and projects have in fact been relevant to their needs.'*

This approach is further strengthened in the 2018 SOHS, where ALNAP added a criterion for **accountability and participation** following discussions with the SOHS Methods Group. The purpose of this was to provide a more focused picture of how the humanitarian system engages with risk reduction, resilience and development programming, while also allowing for a focused assessment of the relationships between international agencies and local actors.

OECD DAC criteria

As part of the definition of relevance outlined at the beginning of this chapter, the OECD DAC guidance (2021: 38) includes the following note:

'Respond to' means that the objectives and design of the intervention are sensitive to the economic, environmental, equity, social, political economy and capacity conditions in which it takes place. 'Partner/institution' includes government (national, regional, local), civil society organisations, private entities and international bodies involved in funding, implementing and/or overseeing the intervention. Relevance assessment involves looking at differences and trade-offs between different priorities or needs. It requires analysing any changes in the context to assess the extent to which the intervention can be (or has been) adapted to remain relevant.

The 2019 OECD DAC criteria ask: *is the intervention doing the right thing?* It does not separate out appropriateness. The guidance (OECD 2021) identifies four main dimensions of relevance: responding to needs, policies and priorities; being sensitive and responsive to context; quality of design; and responsiveness over time. It suggests that the most important element for analysing relevance is the assessment of the extent to which an intervention addresses affected populations' needs and priorities.

'Responsiveness over time' is an important inclusion in the updated criteria, recommending an adaptive management analysis to ensure the intervention remains relevant across a context that changes over time.

Inter-Agency Standing Committee's IAHE 2018

The IASC's IAHE process guidance includes only relevance, not appropriateness, as a criterion. Per the above definition, it focuses on the needs of the most vulnerable groups affected by the crisis. It does not consider links to broader programming, such as risk reduction, resilience, development and peacebuilding programming.

Core Humanitarian Standard

Appropriateness and relevance are front and centre in the Core Humanitarian Standard (2014). Its first commitment is that 'communities and people affected by crisis receive assistance appropriate and relevant to their needs'. The emphasis is on the preference of communities and people affected by crisis; to 'consider the response takes account of their specific needs, culture, and preferences' (CHS 2018). It requires accounting for the capacities of the affected population and the need to adapt programmes to changing needs, capacities and context.

Key issues arising

Relevance is the second most-used OECD DAC criterion in humanitarian evaluations (Darcy and Dillon 2020). While expert opinions diverge (see for example, Heider 2017), the literature highlights the importance of relevance/appropriateness. It does, however, find two main issues related to its application. First: that the power differential in humanitarian aid means that the evaluator's interpretation of what is relevant and appropriate may be very different from what is considered relevant and appropriate by the affected population. Evaluations may find that humanitarian action has been relevant and appropriate for the donor, while ignoring what would have made it relevant and appropriate for the affected population. The second, and related, issue is that the criterion is not systematically applied across evaluations, meaning two evaluations of the same programme could potentially come to different conclusions about its relevance and appropriateness.

Who determines what is relevant and appropriate?

In current practice, Abdelmagid et al. (2019) found that the perspective of humanitarian actors primarily determines the components of appropriate humanitarian action. Field (2016) argues for better awareness of the difference in values and priorities of international humanitarian actors as compared to national actors. The asymmetric power dynamic may come at the expense of accurately evaluating local and national perspectives of what is relevant and appropriate humanitarian action. Field (2016) gives an example where local and national actors prioritised equity over the approach taken to targeting.

This was reflected at ALNAP's 2019 Annual Meeting (ALNAP 2020), which focused on relevance, using a working definition of relevance as 'being in line with the priority needs of affected people' (8). Through discussions on issues of unequal power and privilege, it noted that the international humanitarian system is often 'driven by its own culture, values' as much as 'those of the people it seeks to serve' (ALNAP 2020: 9). Many at the meeting called for a power shift to those who are best placed to judge what is relevant (ALNAP 2020).

It is useful to understand that the imbalance of national perspectives in determining what is relevant is linked to a second set of critiques identified at ALNAP's 2019 Annual Meeting – that the international humanitarian system is supply driven and determined by largely Western agencies and their donors (ALNAP 2020). The power and privilege to determine relevance is embedded in the system.

Organisational guidance

An analysis of guidance for evaluating humanitarian action reflects the tensions identified in the literature on how relevance and appropriateness are determined. Generally, guidance documents (as shown in [Table 6](#)) define relevance as focusing on the extent to which interventions are suited to the needs of affected communities, in addition to the needs of donors, global priorities and the broader political context. However, there are differing points of emphasis. The definitions of relevance and appropriateness given in guidance documents produced by ALNAP, the OECD DAC and UN agencies place greater emphasis on evaluating the extent to which assistance aligns with global priorities and policies. For example, UNICEF guidance refers to alignment with UN mandates and the United Nations Development Assistance Framework. However, in three of the four NGO guidance documents reviewed, NGOs placed greater emphasis in their definitions of relevance on the needs of affected populations. Guidance documents produced by NGOs were also more likely than those produced by UN agencies to emphasise the importance of active participation of local and national stakeholders in evaluations.

Despite broad acknowledgment across the guidance documents of the importance of assessing relevance in relation to local needs, ‘what’ is considered relevant is generally defined by an international evaluator or evaluation manager or by the humanitarian actor. The guidance documents analysed recommend that evaluation questions around relevance should be pre-determined by evaluation teams, and these questions should be used to assess whether the intervention met local needs. However, none of the guidance documents analysed explicitly suggest that relevance should be determined by local or national stakeholders themselves. This indicates that, largely, it is not local or national stakeholders who decide what relevance means, but evaluators.

Table 6: Organisational guidance on relevance/appropriateness

Organisation	Appropriateness or relevance	Emphasis on relevance to affected populations, or donor needs and political context	Whose perspective? (local/national/international)
ALNAP (2006)	Appropriateness and relevance	Emphasis on affected populations	International
OECD (2019)	Relevance	Emphasis on affected populations	International
IAHE (2019)	Relevance	Emphasis on affected populations	International
IFRC (2011)	Appropriateness and relevance	Equal emphasis on both	Undefined
Save the Children (n.d)	Relevance	Equal emphasis on both	Undefined. Includes participation of affected populations throughout
ACT Alliance (2012)	Relevance	Emphasis on donor needs and political context	Undefined
Catholic Relief Services (2012)	Relevance	Emphasis on affected populations	Undefined, however leans local based on evaluation questions
WFP (2021)	Appropriateness and relevance	Emphasis on donor needs and political context	International
Médecins Sans Frontières (2017)	Relevance	Emphasis on donor needs and political context	Undefined
	Appropriateness	Emphasis on affected populations	Undefined
UNICEF (2014)	Relevance	Emphasis on donor needs and political context	International
USAID Bureau for Humanitarian Assistance (2022)	Relevance	Emphasis on affected populations and activity goal	Undefined
UNFPA (2109)	Relevance	Emphasis on both	Undefined
IOM (2021)	Relevance	Emphasis on donor needs and political context (though includes affected populations)	Undefined
	Appropriateness	Emphasis on affected populations	Undefined

Application in evaluations

A reading of 10 evaluations conducted on humanitarian responses (as shown in [Table 7](#)) finds similar trends in whether relevance is defined from international, local or national perspectives. Nine of the evaluations analysed were led by international evaluation teams, reflecting a more technocratic, international perspective.

In terms of emphasis on relevance to affected populations as compared to donor needs and the political context, five of the six NGO evaluations reviewed focused on the needs of affected populations in their assessment of relevance. In contrast, three of the four UN agency evaluations reviewed focused on donor policies and global and sector priorities in their assessment of relevance.

Across all the evaluations reviewed there was little transparency around the involvement of local communities in defining relevance in the context of the humanitarian assistance received. Interestingly, evaluations conducted by local evaluation teams framed relevance as relating solely to community needs, without any reference to donor needs or the broader context. This suggests that relevance may be determined differently based on whether the evaluation is conducted by local, national or international evaluation teams.

That the make-up of the evaluation team, and their world view, might impact the interpretation of relevance reflects the established concept of positionality, which holds that the ‘positioning of the researcher in relation to the social and political context of the study ... affects every phase of the research process ...’ (Coghlan and Brydon-Miller 2014: 628). Of particular relevance to this concept is the researcher’s positionality as an insider or outsider to the community.

Table 7: Organisational evaluations on relevance/appropriateness

Organisation	Title	Emphasis on relevance to affected populations or donor needs and political context	Whose perspective? (local/national/international)
IAHE	Evaluation of the Drought Response in Ethiopia 2015–2018 (2019)	Greater emphasis on planning documents, broader sector priorities	International
Norwegian Refugee Council	Global Cash Evaluation (2019)	Emphasis on needs of affected populations	International
IAHE	Humanitarian Evaluation of the Response to Cyclone Idai in Mozambique (2019)	Emphasis on donor policies and documents	International
UN OCHA	Evaluation of Country-Based Pooled Funds Global Synthesis Report November (2019)	Emphasis on global priorities and commitments	International
WFP and The Cash Learning Partnership (CaLP)	Evaluation of the 2017 Somalia Humanitarian Cash-Based Response (2018)	Emphasis on affected populations	International
CARE International in Pakistan	Humanitarian Project in South Waziristan Tribal District (2019)	Emphasis on affected populations	International
ACTED	Final Evaluation Report: Disaster Response in Yemen (2019)	Emphasis on affected populations	International
Norwegian Church Aid	Sustainable WASH Assistance to the 2010 Flood-Affected Communities in Naushahro Feroze, Sindh (2017)	Equal emphasis on both	National
IFRC	Midterm Review of Malaysia COVID-19 Response (2021)	Emphasis on donor needs	International
IFRC	Timor-Leste Drought Operation (2017)	Emphasis on affected populations	International

Lack of systematic application

Abdelmagid et al. (2019) reviewed eight methods for assessing the appropriateness of humanitarian action and found that 'existing methods vary considerably in their definitions of "appropriateness", provide insufficient guidance on measurement, are vulnerable to interpretive bias and frequently report findings in an ambiguous manner'. While all approaches are framed around the needs of the affected population or context of the crisis, Abdelmagid et al. (2019) note that some focused on response objectives, others on humanitarian organisations and yet others on the actual assistance delivered. Approaches may also include the response design, the services received, the choice of interventions in relation to needs, the modality of the interventions and the extent to which the response caters to the needs of specific groups.

To address this issue, Abdelmagid et al. (2019: 11) proposed a new definition and conceptual framework for appropriateness. The authors define 'appropriate humanitarian assistance' as 'a combination of (i) an intervention/package of services that addresses objective needs and threats to the health or welfare of crisis-affected populations; (ii) a modality of delivery that reflects the context, enhances user acceptability and promotes sustainability where possible; and (iii) a target beneficiary population that is clearly defined, sufficient in size and prioritised according to need.' They propose a specific set of questions reflecting these three components of the definition to assist with more systematic application of the criteria.

Guidance on applying the relevance and appropriateness criteria to humanitarian innovation is focused on assessing whether the innovation was more relevant and appropriate to the needs of end users or of the innovating team (Obrecht, Warner and Dillon 2017). It is less explicit about the macro component and broader context related to needs.

This breadth and flexibility in approach may be a contributor to Caroline Heider's criticism of relevance in the development context. In *Rethinking evaluation – is relevance still relevant?*, Heider (2017: 5) argues that, in practice, evaluators often defer to government and organisational policies, under which a multitude of activities can be deemed to be relevant. Therefore, she concludes, 'meeting the bar for relevance is not all that hard'. Heider further questions whether existing approaches are suitable to today's complex world and proposes network analyses to map out situational problems as the basis for more targeted evaluation questions.

Organisational guidance

The systematic application of relevance and appropriateness as evaluative criteria is challenged at its foundations by the differences in definition of the criteria across the sector. This is reflected in organisational guidance documents. Of the 10 guidance documents analysed, none used the exact definition of relevance and appropriateness provided in the OECD DAC criteria guidance, or in ALNAP's 2006 guide.

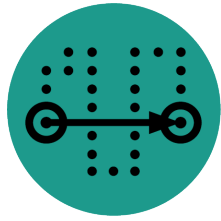
Summary and questions to explore

Relevance/appropriateness is the second most-used OECD DAC criterion in humanitarian evaluations (Darcy and Dillon 2020). The literature highlights two main issues related to its application:

1. Power differentials in humanitarian aid means the evaluator's interpretation of what is relevant and appropriate may be very different from what is considered relevant and appropriate by the affected population. Contemporary guidance does not specify who should define relevance and appropriateness in evaluations.
 - **How might the positionality of the evaluator and the author of the evaluation terms of reference and the composition of the evaluation team impact the evaluation questions, methodology, findings and recommendations?**
 - **Should guidance emphasise whose perspective should be prioritised when defining relevance/appropriateness or how differences in views can be reconciled? For example, is it more important to emphasise what is relevant to people affected by crisis, than what is relevant to donors and global policy objectives? How can principles related to Accountability to Affected Populations (AAP) help inform evaluative perspectives?**
 - **Should guidance put a greater emphasis on positionality and the importance of careful reflection on whose views and which perspectives will inform the evaluative questions in relation to relevance? Should guidance address this directly or specify who should define relevance and appropriateness in evaluations?**
2. Relevance/appropriateness is not systematically applied across evaluations, meaning two evaluations of the same programme could potentially come to different conclusions about its relevance and appropriateness. Of the guidance documents analysed, none used the exact definition of relevance and appropriateness provided in the OECD DAC criteria guidance, or in ALNAP's 2006 guide.
 - **How important is it that relevance and appropriateness are more systematically applied and comparable across evaluations?**

A large, stylized teal number '4' is centered on the page, serving as a background for the text. The number is composed of several overlapping rectangular shapes, creating a 3D effect.

EFFICIENCY



EFFICIENCY

THIRD MOST-USED OECD DAC CRITERION

Definitions

ALNAP Guide 2006		OECD DAC Criteria 2019	
<p>Efficiency: Efficiency measures the outputs – qualitative and quantitative – achieved as a result of inputs. This generally requires comparing alternative approaches to achieving an output, to see whether the most efficient approach has been used.</p>		<p>Efficiency: How well are resources being used? The extent to which the intervention delivers, or is likely to deliver, results in an economic and timely way.</p>	
ALNAP EHA Guide 2016	ALNAP SOHS 2018	IASC's IAHE 2018	CHS 2018
<p>Efficiency: The outputs – qualitative and quantitative – achieved as a result of inputs.</p>	<p>Efficiency: The degree to which humanitarian outputs are produced for the lowest possible amounts of inputs.</p>	<p>Efficiency: Not a criterion.</p>	<p>Commitment 9: Communities and people affected by crisis can expect that the organisations assisting them are managing resources effectively, efficiently and ethically.</p>

Sector-wide guidance and standards

ALNAP

The 2006 ALNAP guidance explains efficiency as the measure of how economically inputs (financial, human, technical and material resources) were converted to outputs. This usually focuses on financial data and factors in the urgency of the assessed needs of the affected population. It often involves comparison with alternatives and asserts the importance of consideration of needs in determining whether an intervention was implemented in the most efficient way. This is illustrated by an example of using food-aid rations that are more culturally appropriate and therefore more likely to suit the needs of the affected population.

More recently, ALNAP (2016) slightly expanded the understanding of efficiency by including focus on both the qualitative and quantitative outputs achieved as a result of inputs; however, there is limited mention of 'efficiency' as a criterion in the guidance. The 2018 SOHS defines 'efficiency' as the degree to which humanitarian outputs are produced for the lowest number of inputs. It also notes that efficiency is both an important and a limited criterion for judging the performance of humanitarian action. Important, because resources are far below needs and must be spread as far as possible. Limited, because the measurement focuses purely on outputs and not on broader achievements, such as lives saved. A further challenge is the nature of efficiency as a relative, rather than an absolute, measure, and comparing the relationship between inputs and outputs for similar activities can be difficult due to programme scale and context – even within the same organisation.

OECD DAC criteria

As part of the definition of efficiency outlined at the beginning of this chapter, the OECD DAC guidance (2021: 58) includes the following note:

'Economic' is the conversion of inputs (funds, expertise, natural resources, time, etc.) into outputs, outcomes and impacts, in the most cost-effective way possible, as compared to feasible alternatives in the context. 'Timely' delivery is within the intended timeframe, or a timeframe reasonably adjusted to the demands of the evolving context. This may include assessing operational efficiency (how well the intervention was managed).

The 2019 OECD DAC criteria consider the extent to which the intervention delivers, or is likely to deliver, results in an economic and timely way. The timeliness dimension was added in 2019 to ensure that it is captured as an important aspect of an intervention's value (OECD 2019). The OECD views the efficiency criterion as an opportunity to check whether an intervention's resources can be justified by its results.

Inter-Agency Standing Committee's IAHE 2018

The IASC's IAHE guidance includes only a reference to efficiency in the context of it belonging to the OECD DAC criteria. It is not a criterion for crisis-specific evaluations.

Core Humanitarian Standard

The CHS highlights the importance of efficiency across a couple of standards: Standard 9 examines the responsible management and utilisation of humanitarian resources and Standard 2 emphasises the timeliness of humanitarian response. The guidance discusses timeliness as not only being the delivery of resources without delay, but also the provision of assistance in the right way at the right time, using the example of different needs of an affected population in summer and winter. It also notes that a balance needs to be struck between economy, effectiveness and efficiency, as 'economical' doesn't always equal value for money; for example, in the instance of a programme that is understaffed or under-resourced.

Key issues arising

Efficiency is one of the most commonly used of the OECD DAC criteria in humanitarian evaluations (Darcy and Dillon 2020), ranked third behind effectiveness and relevance. It is thought that efficiency is frequently applied due to its potential usefulness in decision-making and for learning, accountability and operational improvements (Aiazzi et al. 2011). However, the literature also highlights the limitations of efficiency, in that it often suffers from weak data sources and limited/unreliable/variable methodologies (Multilateral Organisation Performance Assessment Network 2018) and fails to account for social and environmental costs (OECD 2018). Participants in consultations updating the OECD DAC criteria (OECD 2018) recommended specifying the level of the results at which chain efficiency is evaluated, and clarifying the tools that are most appropriate; for example, whether cost-benefit, cost-effectiveness or cost-efficiency analysis would be most useful. Heider (2017) notes that the practice of cost-benefit analysis has been on the decline at the World Bank as it was increasingly not feasible to apply to its projects.

Darcy and Dillon (2020) note that efficiency is often applied to the performance of single projects and programmes, as opposed to more systemic and response-wide approaches. The SOHS (ALNAP 2018a) highlights that many agencies have agreed in principle to move towards comparable cost structures through their Grand Bargain commitments, but there is a lack of clarity on the purpose of this. It also notes that there is little guidance available on how to calculate efficiency in humanitarian operations – and often this takes the form of one-off studies focused on programme interventions, for example, the use of cash rather than other modes of assistance.

Donors, such as the UK's Foreign, Commonwealth and Development Office (FCDO), have long sought to deliver value for money on behalf of their taxpayers. The FCDO has used a '5E' framework to track value for money: economy, efficiency, effectiveness, equity and overall cost-effectiveness (DFID 2020). This somewhat reflects the CHS approach outlined above – aiming to balance economy, effectiveness and efficiency – and is more comprehensive than ALNAP's definitions.

Organisational guidance

Generally, organisational guidance documents (as shown in [Table 8](#)) adopt the original OECD DAC definition of efficiency in terms of the ratio between the inputs and outputs of an intervention. Several guides emphasise the importance of assessing timeliness in the delivery of humanitarian assistance, as well as staffing and other operational efficiencies. Nearly all guidance notes the importance of finding suitable comparisons but provide limited practical examples as to how an evaluator may approach this. The skills and experience of evaluators in approaching efficiency is put under the spotlight in several guidance documents, with OECD (2021) stressing the need for efficiency analysis expertise as part of the evaluation team skill set.

Guidance from NGOs was more likely to focus on the cost of an intervention and whether the same results could have been achieved with fewer resources. Whereas donor and multi-agency guidance provided broader scope for the definition of efficiency, including staffing and operational efficiencies, opportunities for adaptation and the application of different lenses (such as gender equality, disability and social inclusion) to evaluating efficiency. This discrepancy might be explained by time, in that most of the NGO guidance was developed prior to the release of the updated 2019 OECD DAC criteria – which promotes a broader definition of efficiency. It may also be indicative of increasing scarcity and competition for funds for NGOs, which may have resulted in an overstatement of the cost dimension of this criterion.

Table 8: Organisational guidance on efficiency

Organisation	Cost-efficiency	Timeliness	Other efficiencies
ALNAP (2006)	Cost-efficiency	-	-
OECD (2021)	Cost-efficiency	Timeliness	Operational efficiency
IFRC (2011)	Cost-efficiency	Timeliness	-
Save the Children (n.d.)	Cost-efficiency	Timeliness	Donor efficiency
ACT Alliance (2012)	Cost-efficiency	-	-
Catholic Relief Services (2012)	Cost per participant, ratio of programming to administration costs	-	Staffing efficiency, coordination
WFP (2021)	Cost-efficiency	Timeliness	-
Médecins Sans Frontières (2017)	Cost-efficiency	-	-

UNICEF (2014)	Cost-efficiency	Timeliness	Staffing efficiency
USAID BHA (2022)	Cost-efficiency	Timeliness	-
UNFPA (2019)	Cost-efficiency	Timeliness	Implementation efficiency

Application in evaluations

A reading of 10 evaluations conducted on humanitarian responses (as shown in [Table 9](#)) finds a similar focus on cost-efficiency in the delivery of a humanitarian intervention. Timeliness of delivery was examined in four out the 10 evaluations. Operational efficiencies, particularly in taking partnership or consortia approaches, was a key theme emerging from the evaluation review.

Several evaluation reports emphasised the lack or poor quality of data available to assess cost-effectiveness. There was also limited evidence of comparable models/frameworks being established as part of the evaluation methodology.

Table 9: Organisational evaluations on efficiency

Organisation	Title	Emphasis on cost or timeliness	Other efficiencies
World Vision	Real-Time Evaluation of World Vision's Response to the Refugee Influx in Lunda Norte, Angola (2018)	Cost	Organisational
CARE International in Pakistan	Humanitarian Project in South Waziristan Tribal District End of Project Evaluation Report (2019)	Cost and timeliness	Operational (Partnerships)
AHP	Australian Humanitarian Partnership (AHP) Joint Evaluation on Rohingya Crisis Response (2021)	Cost and timeliness	Value for money (Consortia)
UNHCR	Evaluation of the UNHCR Regional Refugee Response to the Venezuela Situation (2020)	Cost and timeliness	-
European Union	Evaluation of the European Union's Humanitarian Interventions in India and Nepal, 2013-2017 (2018)	Cost	Operational (Partnerships, Consortia) and Management
IAHE	Evaluation of the Drought Response in Ethiopia 2015–2018 (2019)	Cost and timeliness	Operational (Partnership)
ACTED and Concern Worldwide	Evaluation Report Disaster Response in Yemen (2019)	Cost	Operational (Partnership, Beneficiary involvement)
Christian Aid, CAFOD Trocaire, Oxfam, Tear Fund	Missed Out: The Role of Local Actors in the Humanitarian Response in the South Sudan Conflict (2016)	Cost	Operational (Partnership)
European Union	Evaluation of the European Union's Humanitarian Response to the Refugee Crisis in Turkey (2019)	Cost	-
IFRC	Timor-Leste Drought Operation Evaluation Report (2017)	Cost	-

Summary and questions to explore

Efficiency is the third most-applied OECD DAC criterion. Cost-efficiency and timeliness are common components of the criterion. Operational efficiencies, such as consortia and partnerships, were often also included in the evaluations reviewed.

The OECD DAC views this criterion as an opportunity to check whether an intervention's resources can be justified by its results. It takes a broad view of resources – including human, environment, financial and time resources – and covers the complete intervention results chain, from outputs to impact.

The CHS notes that a balance needs to be struck between economy, effectiveness and efficiency, as a programme that is economical isn't always value for money – for example, a programme that is understaffed or under-resourced (CHS 2018). Donors, such as the UK's FCDO, have long sought to deliver value for money on behalf of their taxpayers. The FCDO has used a '5E' framework to track value for money – economy, efficiency, effectiveness, equity and overall cost-effectiveness (DFID 2020).

Although evaluation of efficiency is important for decision-making, learning and accountability, the literature indicates that it often suffers as a result of weak data sources and variable methodologies and fails to account for social and environmental costs.

- **Should the 2006 definition of efficiency be broadened, similar to that of the 2019 OECD DAC definition? Or potentially more aligned to value-for-money approaches?**
- **Are cost-efficiency, timeliness and operational efficiencies the three most important components of efficiency? Should effectiveness and equity also be considered, in line with approaches to value for money (such as in FCDO's '5E' framework)?**
- **How can guidance support evaluators when they have limited resources and lack the technical skill sets required for assessing efficiency?**

A large, stylized number '5' in a light teal color, centered on a darker teal background. The number is composed of several overlapping rectangular and circular shapes, creating a layered, geometric effect.

IMPACT



IMPACT

FOURTH MOST-USED OECD DAC CRITERION

Definitions

ALNAP Guide 2006		OECD DAC Criteria 2019	
<p>Impact: Impact looks at the wider effects of the project – social, economic, technical, environmental – on individuals, gender- and age-groups, communities and institutions. Impacts can be intended and unintended, positive and negative, macro (sector) and micro (household).</p>		<p>Impact: The extent to which the intervention has generated or is expected to generate significant positive or negative, intended or unintended, higher-level effects.</p>	
ALNAP EHA Guide 2016	ALNAP SOHS 2018	IASC's IAHE 2018	CHS 2018
<p>Impact: The wider effects of the project – social, economic, technical, and environmental – on individuals, gender- and age-groups, communities and institutions. Impacts can be intended and unintended, positive and negative, macro (sector) and micro (household). (This is not exactly the same thing as 'Impact' in the results chain.)</p>	<p>Impact: The degree to which humanitarian action produces (intentionally or unintentionally) positive longer-term outcomes for the people and societies receiving support.</p>	<p>Impact: <i>Not a criterion.</i></p>	<p>Impact: <i>Not a criterion.</i></p>

Sector-wide guidance and standards

The definition adopted by ALNAP in 2006 (outlined above) aligns with the OECD DAC's 1999 *Guidance for evaluating humanitarian assistance in complex emergencies*. ALNAP explains impact by comparing it to effectiveness: while effectiveness considers whether intermediate objectives have been achieved, impact examines the longer-term consequences of achieving or not achieving those objectives. Understanding impact often requires going beyond the intervention scope to consider the intervention's impact in wider socioeconomic and political contexts. Its longer-term focus has links with the evaluation of connectedness.

It notes that evaluating impact is not relevant to all interventions, particularly those that are short term. Assessing impact requires a particular skill set on the part of evaluators and may be one of the most challenging criteria to assess. Attribution is particularly difficult to establish when assessing impact.

ALNAP'S 2016 guidance reflects the 2006 definition for the impact criterion. It notes the growing interest in evaluations that focus solely on impact, referred to as 'impact evaluations' (ALNAP 2016: 356). It finds that, in practice, some organisations focus their assessment of impact on which outcomes can be attributed to an intervention, while others assess wider impact, including the ripple effects – intended and unintended, positive and negative.

The fundamental challenge in conducting impact evaluation is attribution – that is, isolating the impact that is due to a given intervention from the many other factors at play.

The 2018 SOHS report (ALNAP 2018a) adjusts the definition of impact, focusing on longer-term outcomes for people and societies rather than wider effects as in ALNAP's 2006 definition. It finds that the question of what longer-term impact humanitarian action has on societies is not often asked and 'remains unanswered' (271). It identifies impact as one of the most important and least understood of the criteria, with little hard data measuring the impact of humanitarian responses on wider populations or across time. There was, for example, insufficient data for the SOHS report to draw conclusions on impact.

OECD DAC criteria

As part of the definition of impact outlined at the beginning of this chapter, the OECD DAC guidance (2021: 64) includes the following note:

Impact addresses the ultimate significance and potentially transformative effects of the intervention. It seeks to identify the social, environmental and economic effects of the intervention that are longer term or broader in scope than those already captured under the effectiveness criterion. Beyond the immediate results, this criterion seeks to capture the indirect, secondary and potential consequences of the intervention. It does so by examining the holistic and enduring changes in systems or norms, and potential effects on people's wellbeing, human rights, gender equality, and the environment.

The 2019 OECD DAC criteria definition focuses on higher-level effects, as compared to the 'wider effects' or 'longer-term outcomes' considered in ALNAP's earlier guidance. It is intended to align more with the common usage of the word 'impact'.

The impact criterion encourages consideration of the critical 'so what?' question (OECD 2021: 11). This is where evaluators look at whether the intervention created change that really matters to people. It provides an opportunity to centre participants in evaluations and actively seek their opinions of the impact a humanitarian intervention has had on their lives. The 2021 guidance notes the importance of considering differential impact, particularly any significant, unintended negative distributional effects.

Inter-Agency Standing Committee's IAHE 2018

Although impact is not a criterion in the IASC guidance, its process guidelines establish a coordinated humanitarian action impact pathway as the point of reference for all IAHEs. The impact pathway outlines the inputs, outputs, outcomes and core responsibilities that are intended to lead to the ideal longer-term impact: that affected people live in enhanced safety and dignity with better prospects of thriving as agents of their own destinations. IAHEs are intended to be conducted within that context.

Core Humanitarian Standard

Impact is not treated as a stand-alone criterion in the CHS. It is most closely related to Commitment 3: Communities and people affected by crisis are not negatively affected and are more prepared, resilient and less at-risk as a result of humanitarian action and the associated quality criterion: humanitarian response strengthens local capacities and avoids negative effects. This commitment is also related to the OECD DAC criteria of connectedness and sustainability, which are explored in the next chapter.

Key issues arising

Impact is applied as a criterion in almost 50% of humanitarian evaluations, ranking fourth of the seven criteria in frequency of use (Darcy and Dillon 2020). ALNAP identified a shift towards impact evaluation in its 2009 paper on improving humanitarian impact assessment (Proudlock and Ramalingam with Sandison 2009). Some organisations are, however, excluding impact from their suggested evaluation criteria, including both UNFPA (2019) and the IASC (in its IAHE process guidelines, OCHA 2018).

Two key challenges arise with the impact criterion. These are 1) defining impact and 2) measuring impact. Further, impact evaluation is also its own type of evaluation. This is further explained below.

How to define and assess impact?

While most organisations adhere to ALNAP's definition of impact in their guidance (see [Table 10](#)), it appears to be variably interpreted when it comes to evaluation. One possible reason for this is the joint challenge of defining and assessing impact, including the breadth of the definition.

Challenges defining impact

ALNAP adjusted the definition of impact for the 2018 SOHS, focusing on longer-term outcomes for people and societies rather than wider effects as in ALNAP's 2006 definition. Proudlock and Ramalingam with Sandison (2009) found that focusing on long-term effects or lasting change may not always be useful for humanitarian interventions, as some impacts are measurable within short timeframes and not all humanitarian goals are intended to be sustainable.

The OECD DAC's (2019) definition of impact focuses on higher-level effects. This creates challenges when evaluating humanitarian interventions. Clements (2020) indicates that there is often no requirement to establish a logic of project implementation to the point at which people experience changes in wellbeing. This means many evaluations conclude with project outputs rather than impact. Where longer-term interventions have a programme logic establishing impact, it is often the case that there is no adequate baseline data on key indicators related to wellbeing – for example, livelihoods – against which to assess whether there has been a change. This can be compounded by an absence of high-quality monitoring data over time.

Proudlock and Ramalingam with Sandison (2009: 5) further emphasised the importance of understanding impact from different perspectives, finding that impact assessment 'inevitably involves value judgments about which kinds of changes are significant, and for whom'. They outline a diverse range of stakeholders and their different interests in impact, including affected populations, host governments, donors and implementing organisations, and find that this diversity means there is no consistent view of impact assessment.

Challenges assessing impact

Perhaps the most prominent challenge is that of establishing causation and attribution. Brusset, Cosgrove and Macdonald (2010) note that the dynamic and unpredictable humanitarian context and multiple international organisations mean it is complex to attribute impact to any one cause. UNFPA's (2019) decision to exclude impact from their list of suggested evaluation criteria and focus instead on the more immediate results of its assistance was based on this challenge in attributing impact (or even showing contribution to impact) to high-level societal conditions. Further, UNFPA found that generating lessons relevant to the next country programme is a difficult task, given that 'societal changes are far removed from UNFPA programming decisions' (2019: 293).

Impact is more easily assessed in multi-year programmes, which remain the exception rather than the norm for humanitarian interventions. Multi-sector interventions may also provide greater scope to understand social, political or economic changes. Darcy and Dillon (2020) found a marked absence of studies that looked at systemic or response-wide issues.

Impact evaluations

Uniquely among the OECD DAC criteria, impact evaluations are also a type of evaluation focused on results at impact level (see [Box 5](#)). Some organisations are increasingly emphasising impact evaluation methods. The WFP has, for example, developed a strategy dedicated to impact evaluation (2019–2026) and is working with the World Bank and USAID's Bureau for Humanitarian Assistance to conduct impact evaluations (WFP 2021); UNICEF has produced research on impact evaluation in settings of fragility and humanitarian emergency (Bakrania et al. 2021), and the World Bank, together with UNHCR and the United Kingdom's FCDO, have invested in impact evaluations in forced displacement contexts (Elice 2021).

There are good arguments for conducting impact evaluations. They can provide a more evidence-based understanding of how aid ultimately affects the lives and livelihoods of people affected by crisis (Proudlock and Ramalingam with Sandison 2009). Puri et al. (2017: 531) argue that the 'dearth of rigorous causal evidence of what works and what does not work in the humanitarian sector means there is a high dividend to be earned' from impact evaluation. At a practical level, the WFP (2021) uses impact evaluations to quantify the impact of different programming choices, usually by comparing groups of intended recipients of humanitarian assistance.

Impact evaluations measure changes in outcomes of interest that can be attributed to a specific intervention through a credible counterfactual using a range of quantitative methods, usually complemented by qualitative methods (WFP 2021). There are substantive methodological and contextual challenges in conducting impact evaluations of humanitarian interventions (Bakrania et al. 2021). Research has, however, found that it is possible to assess the impact of humanitarian interventions ethically and with statistical confidence (Puri et al. 2017; Bakrania et al. 2021).

Box 5: Impact evaluation

ALNAP's *Guide to evaluating humanitarian action* (2016) defines 15 types of evaluations, one of which is impact evaluation. It describes impact evaluation as focusing on a particular level of results:

An evaluation that focuses on the wider effects of the humanitarian programme, including intended and unintended impact, positive and negative impact, macros (sector) and micro (household, individual) impact. (ALNAP 2016: 81)

ALNAP differentiates impact evaluation from normative evaluation focusing on a different level of results, comparing 'what is being implemented with what was planned or with specific standards' (ALNAP 2016: 82). It further explains that:

The essence of impact evaluation is that it explores the cause and effect or 'causal inference'. It shifts the focus away from the effectiveness and efficiency of the intervention, to examine whether people are better off or safer as a result. Establishing this causal relationship is the challenge. (ALNAP 2016: 357)

Organisational guidance

The guidance for evaluating humanitarian action reflects the above challenges and primarily focuses on the original definitions of impact. Several guidance documents highlight the need to establish a cause-and-effect relationship and the challenge this presents, particularly in single-organisation evaluations. Guidance by IFRC and MSF also acknowledge potential methodological constraints such as a lack of baseline or monitoring data with which to develop tools to assess impact. The extent to which a humanitarian intervention is *attributable* or has *contributed* to broader impact, and the challenges in assessment, is explored in more than half of the organisational guidance documents highlighted in [Table 10](#).

Table 10: Organisational guidance on impact

Organisation	Primary definition	Reference to contribution or attribution challenges
ALNAP (2006)	Positive and negative, intended and unintended, directly and indirectly	Attribution
OECD (2019)	Positive and negative, intended and unintended, directly and indirectly	Attribution
IFRC (2011)	Positive and negative, intended and unintended, directly and indirectly	Attribution
Save the Children (n.d)	Positive and negative, intended and unintended, directly and indirectly	Contribution
ACT Alliance (2012)	Positive and negative, intended and unintended, directly and indirectly	-
Catholic Relief Services (2012)	Positive and negative, intended and unintended, directly and indirectly	-
WFP (2021)	Positive and negative, intended and unintended, directly and indirectly	Attribution
Médecins Sans Frontières (2017)	Positive and negative, intended and unintended, directly and indirectly	Attribution
UNICEF (2014)	Positive and negative, intended and unintended, directly and indirectly	Contribution
USAID BHA (2022)	Positive and negative, expected and unexpected	Attribution
IOM (2021)	Positive and negative, intended and unintended, higher-level effects; direct and indirect (through evaluation questions)	-
Humanitarian Cash Transfers (2018)	Positive and negative, 'wider effects' examples: household relations, local markets, community self-help systems, local debt and credit markets	-

Application in evaluations

A reading of 10 evaluations of humanitarian interventions with impact as a criterion (as shown in [Table 11](#)) highlights similar challenges to those identified by the guidance documents and wider literature on the definition and assessment of impact. Two evaluation reports explicitly stated that long-term impact was not measured, given that it was a humanitarian intervention or that multiple organisations being involved in a response would make it difficult to attribute impact. Several evaluations reported on unintended impacts, including some negative outcomes, as a result of the intervention.

It seems clear from the evaluations analysed for this paper that organisations (and evaluators) view impact as an important criterion to include but need clearer and perhaps more realistic guidance on defining impact in a humanitarian context. Impact evaluations, as compared to

evaluations assessing impact as a criterion, don't seem to face this challenge. They are able to take a far more in-depth approach to evaluating impact.

Table 11: Organisational evaluations on impact

Organisation	Title	Assessment of impact
CARE	Final Report for the Final Evaluation of the OFDA Response Program (2021)	Explicitly states there is no long-term impact due to the nature of the response work, however increased access to services is noted
DFID	Evaluation of the Humanitarian Innovation and Evidence Program (HIEP) (2018)	Impact from the perspective of creating enabling environments; systems change through the application of lessons; research
Plan International UK	External Evaluation. Plan International UK's DEC-funded Response to the Rohingya Refugee Crisis In Bangladesh (2019)	Impact from cumulative contributions of stakeholders; positive and negative changes brought about both directly and indirectly, either intended or unintended
Plan International UK	DEC-Funded Response to the Nepal Earthquakes (2018)	Impact of different types of recovery interventions such as shelter strengthening and livelihoods skills training and how these contribute to overall better quality of life
IFRC	Timor-Leste Drought Operation Evaluation Report (2017)	Impact of interventions on future disaster preparedness and the coping abilities of the affected population
ACTED and Concern Worldwide	Evaluation Report Disaster Response in Yemen (2019)	Reports on negative impacts; considers attribution
IFRC	Real-Time Evaluation Indonesia: Earthquakes and Tsunami (Lombok, Sulawesi) (2019)	Impact of a 'localised'/locally led humanitarian response
IFRC	Evaluation of the West Africa Ebola Viral Disease Appeal Response, Sierra Leone and Liberia (2018)	Notes the challenges of attribution, assessment of unintended impact and lack of baseline data (particularly to assess the recovery phase)
CARE	Humanitarian Project in South Waziristan Tribal District (SWTD): End of Project Evaluation Report (2019)	Assesses impact by considering project outcomes, unintended impacts and the influence context has in optimising impact (or creating challenges)
UNICEF	Review of the L2 Response in Venezuela (2020)	Applies the OECD DAC 2019 definition; assesses impact in line with available data; focuses on decrease in infant mortality

Summary and questions to explore

Impact is the fourth most-applied OECD DAC criterion, considered in almost 50% of evaluations (Darcy and Dillon 2020). ALNAP (2006) explains the differences and similarities between impact and effectiveness: while effectiveness considers whether intermediate objectives have been achieved, impact examines the longer-term consequences of achieving or not achieving those objectives. The longer-term focus of this criterion links it with the connectedness criterion in evaluating lasting benefits. The 2019 OECD DAC criteria definition focuses on higher-level effects (such as changes in norms or systems), compared to ALNAP's concentration on wider effects (2006) or longer-term outcomes (2018).

The 2021 OECD DAC guidance notes the importance of considering differential impact across groups of people (such as groups of people disaggregated by gender, age, ability or vulnerability), particularly any significant, unintended negative distributional effects.

Proudlock and Ramalingam with Sandison (2009) further emphasised the importance of understanding impact from different perspectives.

Some organisations, including the WFP, UNHCR, the World Bank, USAID's Bureau for Humanitarian Assistance and the United Kingdom's FCDO have invested in impact evaluations (focused solely on impact) using very particular methodologies. Other organisations, including UNFPA (2019) and the Inter-Agency Standing Committee (in its IAHE process guidelines 2018) have excluded impact from their evaluation criteria entirely.

- **Is it more appropriate to focus on wider effects, longer-term outcomes or higher-level effects when evaluating the impact of humanitarian action? Does the answer vary by organisation and programme?**
- **Is there enough focus on unintended impact in humanitarian evaluations?**
- **How can guidance address the challenges of evaluating impact in short time-horizons and of establishing cause and effect and attribution in humanitarian settings?**
- **Should guidance include views on the types of situations and settings in which evaluating impact is more or less appropriate?**

A large, semi-transparent teal number '6' is centered on the page, serving as a background for the text.

COVERAGE



COVERAGE

FIFTH MOST-USED OECD DAC CRITERION

Definitions

ALNAP Guide 2006		OECD DAC Criteria 2019	
Coverage: The need to reach major population groups facing life-threatening suffering wherever they are.		Coverage: Not a criterion.	
ALNAP EHA Guide 2016	ALNAP SOHS 2018	IASC's IAHE 2018	CHS 2018
Coverage: The extent to which major population groups facing life-threatening suffering were reached by humanitarian action.	Coverage: The degree to which action by the international humanitarian system reaches all people in need.	Coverage: Not a criterion. Coordination: Was the assistance well-coordinated, successful and, as much as possible, equitable, reaching all affected populations and avoiding duplication of assistance and gaps?	Commitment 6: Communities and people affected by crisis receive coordinated, complementary assistance. Quality criterion: Humanitarian response is coordinated and complementary.

Sector-wide guidance and standards

ALNAP

ALNAP's 2006 guide references the *Code of conduct for the International Red Cross and Red Crescent Movement and non-governmental organisations in disaster relief* (1994) to explain the importance of evaluating coverage: the humanitarian imperative to provide humanitarian assistance wherever it is needed, and the commitment to calculate aid priorities based on need alone without adverse distinction of any kind.

The guide recommends that the evaluation of coverage takes place at three levels: international, national, or regional and local. At regional and local levels, evaluators must assess both who has been included that should not have been (inclusion bias), and who has been excluded that should have received the assistance or protection (exclusion bias). It recommends that evaluators consider proportionality – whether aid was provided according to need at each level – as well as differing perspectives on what constitutes need. Equity is central to the analysis of coverage and requires both geographical analysis and disaggregation of data by socioeconomic categories including gender, socioeconomic grouping, ethnicity, age and ability. Political context should be considered to help determine why a particular group of people were included or excluded.

In 2016, ALNAP kept the definition largely unchanged. ALNAP's 2018 SOHS report adapted the definition to apply to the humanitarian system by focusing on the degree to which the system reaches all people in need. It found that coverage is increasingly getting worse. And that key informants were frequently able to identify areas that had not been reached by ongoing responses. Perhaps of use to evaluators as factors to consider, the 2018 SOHS report found that coverage was poorest in remote areas with low population densities, in areas where there was a high perceived risk to humanitarian staff and in areas under siege. Displaced people who were not resident in camps, irregular migrants and marginalised groups – particularly minority ethnic and cultural groups and the elderly – also experienced poor coverage.

OECD DAC criteria

The criterion of coverage is specific to humanitarian action and is not included in the OECD DAC guidance for the evaluation of development interventions. In its 1999 guidance for evaluating complex emergencies, the OECD DAC introduced the coverage criterion for humanitarian action given its significantly differing impacts on different population subgroups – including ethnicity, gender, socioeconomic status, occupation, location or family circumstance (e.g., single mother, orphan). The examples given at the time demonstrate the implications when coverage is not considered. For example, in one crisis, more than 90% of international assistance went to people located in government-controlled areas, essentially penalising those in areas controlled by insurgent movements (quoting Minear 1994).

Inter-Agency Standing Committee's IAHE 2018

The IAHE does not include coverage as a separate criterion in its IAHE process guidelines. It does, however, explore coverage under the criterion of coordination, exploring the key concepts of equity and reach to all affected populations.

The guidelines also mention that in order to enhance accountability to affected populations, IAHEs will 'endeavour to gain their perspectives on the ... coverage of the emergency response ...' (OCHA 2018: 15). 'Good coverage' is also listed as an outcome in the ideal model of coordinated humanitarian action (Ibid.).

Core Humanitarian Standard

The Core Humanitarian Standard does not emphasise coverage, but does consider it under Commitment 6, which focuses on coordinated, complementary assistance. In its 2018 guidance, the CHS notes as part of Action 6.3 that the purpose of organisations participating in relevant coordination bodies is to 'minimise demands on communities and maximise the coverage and service provision of the wider humanitarian effort' (CHS 2018: 20). When monitoring for Commitment 6, guiding questions include: 'are gaps and duplication in coverage identified and addressed?' (CHS 2018: 31).

Key issues arising

Coverage is among the least-used OECD DAC criteria for the evaluation of humanitarian action, but does still occur in approximately 40% of the 549 studies reviewed for the 2018 SOHS (Darcy and Dillon 2020). The question of how the criterion should be applied to evaluation is not widely discussed in the literature, and the contexts in which it is mentioned relate more to practice. For example, the World Disasters Report in 2018 (IFRC) focused on the substantive gaps in the humanitarian system's attempt to 'leave no one behind', noting that in some crises fewer than half the people in need were known to be reached by international assistance. It identifies five 'fatal flaws' that facilitate this lack of coverage: too many affected people are (1) *out of sight*, (2) *out of reach*, (3) *left out of the loop*, or find themselves in crises that are (4) *out of money* or deemed to be (5) *out of scope* because they are suffering in ways not considered the responsibility of the humanitarian sector. It notes tensions with efficiency and security considerations in reaching those furthest behind and hardest to reach – noting that they are also often the most expensive and risky to reach.

The pursuit of coverage has been identified as a driver of localisation (Barbelet et al. 2021) and, separately, local actors have been noted for their crucial role in improving coverage in South Sudan (Tanner et al. 2016).

Where it is included, the definition of coverage does not vary greatly across the various multisectoral guidance documents discussed above. Classified as a ‘humanitarian’ criterion, it is not included in guidance for evaluating development interventions. It is perhaps notable in its absence from the IAHE process guidelines. It is included as a component of the related CHS criterion of coordinated and complementary assistance, which may be explained by the focus of the CHS criterion on the role of organisations in being coordinated and complementary and, through coordinated and complementary actions, contributing to coverage more broadly.

Organisational guidance

Perhaps reflecting its varied inclusion in multisectoral guidance, some of the organisation guidelines reviewed did not elaborate on coverage. For example, it was not a criterion in USAID’s Bureau for Humanitarian Assistance’s *Technical guidance for monitoring, evaluation and reporting for emergency activities* (BHA 2022), and is referenced but not defined in the *CRS Guidance on monitoring and evaluation* (2012). While MSF elaborated on coverage in its guidance (2017), it did so under the effectiveness criterion.

Organisational guidance also varied in terms of how explicit it was about specifying geographic and socioeconomic coverage, and the importance of providing assistance and protection proportionate to need. The analysis in [Table 12](#) demonstrates that while most guidelines (eight of 10) consider geographical coverage, only IFRC and Save the Children specify the need to consider geographical coverage across the three geographical levels outlined by ALNAP in 2006: international, national/regional, and local. All organisation guidelines discussed population groups or target populations as important to coverage, with five of the 10 guideline documents making explicit the socioeconomic categories that might be important to consider.

The identification of differential coverage for geographic and socioeconomic groups helps evaluators to assess proportionality and equity – highlighted by ALNAP in 2006 as central to coverage. Of the guidelines reviewed, 60% incorporated the concept of proportionality. Four of the 10 organisation guidelines reviewed (those of UNFPA, Oxfam, IFRC and the WFP), specified the need to evaluate for inclusion bias, as well as for the more obvious exclusion bias.

Table 12: Organisational guidance on coverage

Organisation	Geographical coverage	Socioeconomic coverage	Proportionate need
ALNAP (2006)	✓ (across 3 levels)	✓	✓
ALNAP (2016)			✓
IFRC (2011)	✓	✓	✓
Save the Children (n.d.)	✓	Implicit – target groups	Not explicitly
ACT Alliance (2012)	Definition only	Definition only	Definition only
WFP (2021)	✓	✓	✓
Médecins Sans Frontières (2017)	✓	✓	Not explicitly
UNICEF (2014)	✓	X	✓
UNFPA (2019)	Not explicitly	✓	Not explicitly

Oxfam M&E Framework for WASH Market-Based Humanitarian Programming	✓	✓	Extent to which needs of groups were met
Humanitarian Cash Transfers M&E Guidance (2018)	Not explicitly	Implicit – the most vulnerable	Balance between number of people and amount of assistance
IOM M&E Guidelines: (2021)	Not explicitly	Implicit – major groups in need	✓

Application in evaluations

Despite the IAHE process guidelines not specifying coverage as a separate criterion, it clearly remains a priority in some contexts. It was evaluated separately for the IASC evaluation of the response to Cyclone Idai in Mozambique, which asked: *to what extent were different groups of affected people, in all locations affected by Cyclone Idai, reached with humanitarian emergency aid?* UNICEF focused an entire evaluation on the coverage and quality of its humanitarian response in complex emergencies in 2019. In it, UNICEF defines coverage with reference to ALNAP’s 2016 guidance, drawing on the WFP’s definition to add the provision ‘of impartial assistance and protection proportionate to need’, in line with the concerns that informed the design of the evaluation.

The UNICEF evaluation focused on coverage is comprehensive and hugely informative as to the issues likely to be faced by a range of humanitarian agencies when aiming for and evaluating coverage. For example, it identified challenges in:

- Identifying and distinguishing between people (a) affected by crisis; (b) in need; (c) targeted for planned interventions; and (c) ultimately reached. It referenced the 2015 SOHS finding that the ‘lack of solid data on people in need remains a major obstacle to understanding the success or failure of a humanitarian response’.
- Being transparent about the basis on which population and targeting figures are calculated.
- The trade-off between equity/quality and coverage – rarely explicitly documented by humanitarian actors. The evaluation found coverage is consistently prioritised over equity and quality – particularly at the onset of a crisis – with no common understanding about when to transition from prioritising coverage to increasing quality.

This final challenge, understanding the trade-off between equity and coverage, is of particular note, given ALNAP’s 2006 guidance that equity is central to the achievement of coverage.

While there is variation in terms of the way in which coverage is assessed in evaluations, the differences appear to be more related to the depth in which it is explored rather than the approach taken, although some approaches are quite different; see, for example, the evaluation of Pakistan’s IDP [internally displaced persons] Fund (GLOW Consultants and Humanitarian Advisory Group 2019).

The way in which coverage is assessed in evaluations varies – including large differences in the depth to which it is evaluated. It is most often assessed from the perspective of geographic coverage (nine of 10 evaluations reviewed). Of the evaluations reviewed, 70% considered different socioeconomic categories of people that were provided assistance, while only 40% explored proportionality.

Table 13: Organisational evaluations on coverage

Organisation	Title	Geography	Socioeconomic category	Proportionality
IAHE	Humanitarian Evaluation of the Response to Cyclone Idai in Mozambique (2020)	✓	✓	✓
UNICEF	Evaluation of the Coverage and Quality of the UNICEF Humanitarian Response in Complex Emergencies (2019)	✓	✓	✓
UNHCR	Evaluation of the UNHCR Regional Refugee Response to the Venezuela Situation (2020)	✓	✓	When discussing how well groups' needs were met
WFP	Corporate Emergency Evaluation of the WFP Regional Response to the Syrian Crisis (2018)	✓	Not explicitly (target population)	Unclear
Plan International	Plan International DEC-Funded Response to the Nepal Earthquakes, 2015 (As Coverage and Targeting) (2018)	Not explicitly (needs of marginalised members)	✓	Not explicitly
Concern Worldwide	Responding to Pakistan's IDP Fund (Responding to Pakistan's Internally Displaced (RAPID) Fund-II) (2019)	✓	Not explicitly	Not explicitly
IFRC	Evaluation of IFRC West Africa Ebola Viral Disease Appeal Response Sierra Leone and Liberia (2018)	✓	✓	✓
UNHCR	Independent Evaluation of the UNHCR South Sudanese Refugee Response in White Nile State, Sudan (2018)	✓	✓	Not explicitly
Oxfam	Missed Out: The Role of Local Actors in the Humanitarian Response in the South Sudan Conflict	✓	Not explicitly	X
DEC	East Africa Response Review: South Sudan (2016)	✓	✓	Not explicitly

Summary and questions to explore

Coverage is an additional criterion in ALNAP's 2006 guide; it is not an OECD DAC evaluation criterion. It is one of the least-used of the criteria (followed by coherence, sustainability and connectedness), occurring in approximately 40% of humanitarian evaluations (Darcy and Dillon 2020).

Coverage is highly relevant to issues identified in the performance of humanitarian action and is a priority criterion for some. It is central to the commitment to 'leave no one behind'. There is variance in terms of how explicitly evaluations consider geographic coverage, socioeconomic coverage and proportionality to need. A more recent definition of coverage (UNICEF 2019) drew on ALNAP and the WFP to define coverage as the extent to which major population groups facing life-threatening suffering are being (or were) reached by humanitarian action, including the provision of impartial assistance and protection proportionate to need. In practice, geographic and socioeconomic coverage is considered most often; proportionality is less common. Some guidelines specify the population

groups important to consider, along with the need to evaluate for inclusion bias as well as the more obvious exclusion bias.

- **Should the definition of coverage include geographic and socioeconomic coverage, and proportionality to need? Should they receive an equal level of focus?**

In 2019, UNICEF conducted an evaluation focused on coverage and quality. It identified challenges in reliably identifying and distinguishing between people (a) affected by crisis; (b) in need; (c) targeted for planned interventions; and (d) ultimately reached. It also identified challenges in understanding how these population and targeting figures are calculated.

- **Are these categories helpful to include in evaluations? Is it important for evaluations to review the source of calculations for targeting assistance?**

UNICEF's evaluation also highlighted trade-offs between equity, quality and coverage. The evaluation found that coverage is consistently prioritised over equity and quality – particularly at the onset of a crisis – with no common understanding of when to transition from prioritising coverage to increasing quality.

- **Should guidance support evaluators to assess coverage in the context of equity and in relation to quality? If so, how?**



COHERENCE



COHERENCE

SIXTH MOST-USED OECD DAC CRITERION

Definitions

ALNAP Guide 2006		OECD DAC Criteria 2019	
<p>Coherence: The need to assess 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.</p>		<p>Coherence: The compatibility of the intervention with other interventions in a country, sector or institution.</p>	
ALNAP EHA Guide 2016	ALNAP SOHS 2018	IASC's IAHE 2018	CHS 2018
<p>Coherence: The extent to which security, developmental, trade, and military policies as well as humanitarian policies, are consistent and take into account humanitarian and human rights considerations. (More focused on donor policy but can also be applied to individual agencies on their own policy coherence.)</p>	<p>Coherence: The degree to which actors in the international humanitarian system act in compliance with humanitarian principles and international humanitarian law (IHL), and the degree to which they are able to influence states and non-state armed groups to respect humanitarian principles and conform to IHL.</p>	<p>Coherence: Not a criterion.</p>	<p>Commitment 6: Communities and people affected by crisis receive coordinated, complementary assistance. Quality criterion: Humanitarian response is coordinated and complementary.</p>

Sector-wide guidance and standards

ALNAP

ALNAP included coherence as a criterion in its 2006 guide, reflecting the OECD DAC's 1999 *Guidance for evaluating humanitarian assistance in complex emergencies*. At the time, it was not an OECD DAC criterion for evaluating development assistance. It was introduced for complex emergencies in response to a 'notable lack of coherence' in the humanitarian response to the Rwanda genocide of 1994, among others (OECD 1999).

ALNAP's 2006 guide focuses on the extent to which policies of different actors are complementary or contradictory. It is very broad in that it includes any type of policy, from gender equality to environmental protection. It suggests the evaluation of coherence is particularly relevant in two main circumstances. First, when multiple humanitarian actors are involved in an intervention, due to the possibility of conflicting mandates or interests among the responding actors, and second, when political influences can be seen to foster the occurrence or continuation of a humanitarian emergency, and when military and civilian actors are involved in the same emergency.

ALNAP's 2006 guide finds that coherence is the most difficult of the OECD DAC criteria to evaluate, particularly for single-organisation, single-project evaluations. There is also a tendency to confuse it with coordination (noting that coherence should focus on policy-level coordination, while coordination should focus on operational issues, under the effectiveness criterion). The 2006 definition was reinforced in ALNAP's 2016 *Guide to evaluating humanitarian action*.

ALNAP has since taken a more targeted approach to the coherence criterion, defining it in the 2018 SOHS report entirely in relation to humanitarian principles and international

humanitarian law. ALNAP's 2018 *Evaluation of protection in humanitarian action* (ALNAP 2018b) combined its 2006 definition with its approach in the SOHS, framing the evaluation question for coherence as: 'does the intervention adhere to core humanitarian principles and align with broader peace and development goals?' It notes that while there may be broad consensus on core humanitarian principles and peace and development goals, the interpretations may vary – as do the mandates and areas of engagement of different agencies.

OECD DAC criteria

As part of the definition of coherence outlined at the beginning of this chapter, the OECD DAC guidance (2021: 45) includes the following note:

The extent to which other interventions (particularly policies) support or undermine the intervention and vice versa. This includes internal coherence and external coherence. Internal coherence addresses the synergies and interlinkages between the intervention and other interventions carried out by the same institution/government, as well as the consistency of the intervention with the relevant international norms and standards to which that institution/government adheres. External coherence considers the consistency of the intervention with other actors' interventions in the same context. This includes complementarity, harmonisation and co-ordination with others, and the extent to which the intervention is adding value while avoiding duplication of effort.

The OECD DAC introduced coherence for the first time in 2019 to better capture linkages, systems thinking, partnership dynamics and complexity (OECD 2019). It reflects the breadth of the 2006 ALNAP definition in that it considers 'the extent to which other interventions (particularly policies) support or undermine the intervention, and vice versa'. The guidance distinguishes between internal and external coherence, as outlined in the note above.

The OECD DAC's 2021 guidance likens the assessment of coherence to taking a 'zoomed out' approach that allows evaluators to explore where an intervention sits within the broader context of a particular response. It argues that assessing coherence allows for an integrated approach that considers the trade-offs, tensions and inconsistencies between governments and institutions that can undermine overall progress.

Inter-Agency Standing Committee's IAHE 2018

The IAHE process guidelines (OCHA 2018) do not include coherence as a criterion. It has, however, been applied in one of the four inter-agency evaluations completed since the publication of the process guidelines: the 2020 Inter-Agency Humanitarian Evaluation on the Gender Equality and the Empowerment of Women and Girls (IASC 2020a).

The IAHE indicates that humanitarian principles are to be given due consideration in the conduct of the evaluation, and IAHE evaluators are to pay systematic attention to access and protection and collective responses for collective outcomes (under the three criteria: effectiveness, sustainability and partnerships). Coordination is a separate criterion in the IAHE process guidelines, noting ALNAP's distinction from coherence.

Core Humanitarian Standard

Although there is no single CHS commitment that aligns perfectly with the coherence criterion, Commitment 6 (as defined above) provides some crossover with existing guidance on evaluating coherence, particularly regarding coordination and complementarity.

Key issues arising

Coherence is one of the least-applied criteria. This is possibly due to its being subsumed under other criteria, and the difficulties evaluating it. This may be changing. There has long been a global commitment to working more coherently, including the UN's Deliver as One agenda (2006). The OECD DAC added coherence as a criterion for the first time in 2019. Its definition of coherence aligns with the ALNAP 2006 definition but is broader in scope, including concepts of complementarity and coordination (OECD 2021). Drew (2021) also finds the OECD DAC definition to be inclusive of the connectedness criterion.

There is relatively limited literature available on coherence and that which is available mostly examines the challenges with its application. In a comprehensive scoping study, Drew (2021: 18) identified considerable heterogeneity in both the type of coherence being evaluated and the elements evaluated under the coherence criterion. He also found that coherence is often not assessed separately, but rather included under other criteria, perhaps due to its relationship with relevance, effectiveness, impact, sustainability and efficiency. Others have noted relationships with connectedness (ALNAP 2018a) and coordination (ALNAP 2006).

There is agreement that the main sub-division of coherence is internal and external coherence. Internal coherence is measured by an organisation's alignment with its own policies, strategies and commitments to standards and principles. External coherence is assessed by the organisation's alignment or fit with other actors in a humanitarian intervention context.

Challenges in evaluating coherence can be linked to challenges in implementation. In the context of peace and stability operations, de Coning and Friis (2011) argue that there are greater limits to external coherence than is acknowledged in the policy debate. Their paper questions the assumed cause-and-effect relationship between coherence, effectiveness, efficiency and sustainability. It emphasises that coherence should not be aimed equally across all interventions and stakeholders but scaled to the most appropriate and realistic level of coherence for each relationship, which varies in every context.

Humanitarian principles have become more prominent in ALNAP's guidance and application of the coherence criterion (see above discussion on sector-wide standards and guidance). This shift appears to have served the SOHS well; it reflects the origin of the criterion and focuses the evaluation. As with coherence more broadly, humanitarian principles are both debated (see, for example, Slim 2020) and challenging to implement (see, for example, Buchanan-Smith 2022).

A United Nations Evaluation Group working paper (UNEG 2016a) advocated for humanitarian principles to be embedded systematically as core elements of the evaluation of humanitarian action. It found that the evaluation of humanitarian principles has not been a priority, having only been assessed well in 4% of the evaluations reviewed. It also noted the significant challenges in evaluating humanitarian principles, including a lack of a common understanding of the implementation of humanitarian principles, and sensitivity to increased attention on humanitarian principles in evaluations due to the potential for impact on security, ongoing access negotiations and so on.

Organisational guidance

Of the individual organisational guidance documents analysed, coherence is the criterion least often covered. IAHEs, UNFPA, CRS, Save the Children and OXFAM (for water, sanitation and hygiene [WASH] market-based humanitarian programming) did not include coherence as a criterion. Those that did include coherence focused on policies and principles (Table 14). The analysis also suggests that the importance of this criterion may be higher for some organisations, such as UN agencies, than others, such as NGOs, who may be more likely to evaluate single-agency, single-project interventions. Coherence is treated more comprehensively in guidance documents from UNICEF, the WFP and IFRC. Only IFRC and the WFP include humanitarian principles as central to coherence. Humanitarian principles are otherwise not mentioned, or only tangentially mentioned, in organisational guidance for evaluation.

Table 14: Organisational guidance on coherence

Organisation	Focus on policies and principles	Humanitarian principles
ALNAP (2006)	The extent to which policies of different actors were complementary or contradictory	Not mentioned
OECD (2019)	The extent to which other interventions (particularly policies) support or undermine the intervention and vice versa. Internal and external coherence	Not mentioned
IFRC (2011)	Adopts the ALNAP 2006 definition	Yes – central to coherence
Save the Children (n.d)	Coherence not a criterion, however definition of coverage and non-discrimination asks if the work resulted in more equitable and non-discriminatory projects, programmes and/or policies for children and young people at national, regional and global levels	Not mentioned
ACT Alliance (2012)	Adopts the ALNAP 2006 definition	Yes, but tangential (not under coherence)
WFP (2021)	Adopts the OECD DAC 2019 definition. Considers adherences to internal and external policies and principles; consistency and alignment with other humanitarian interventions by government or other actors	Yes, as evaluation question for coherence
Médecins Sans Frontières (2017)	Reference to coherence as a criterion only, no definition provided	Yes, but under relevance and effectiveness
UNICEF (2014)	Adopts the ALNAP 2006 definition. Considers how UNICEF has leveraged its position to advocate for more effective policies. Includes collaboration with non-traditional humanitarian actors	Not mentioned
USAID Bureau for Humanitarian Assistance (2022)	Adopts the OECD DAC 2019 definition. Less explicit about policies and principles. Includes how the activity considers gender equity, protection, age, physical and emotional challenges of participants, and risks to participation, as well as programme adaptation	Yes, but in relation to methods, not criteria
IOM (2021)	Adopts the OECD DAC 2019 definition and the ALNAP 2006 definition	Not mentioned
UN Women (n.d)	Adopts the OECD DAC 2019 definition. Coherence is also a principle for how the evaluation should be conducted	Not mentioned

Application in evaluations

A reading of 10 evaluations conducted on humanitarian responses (as shown in Table 15) demonstrates the focus of coherence to be on internal and external policy and programme/planning coherence. Coherence was represented mainly in UN and IFRC evaluations or in donor-led evaluations. This supports the literature findings on the complexity of assessing coherence in a single-organisation/single-intervention evaluation, particularly for NGOs.

Three of the evaluations reviewed considered humanitarian principles, with two of these considering humanitarian principles under questions related to coherence. These were for evaluations of humanitarian interventions implemented by UNICEF, the WFP and the European Union. This is a higher proportion than the 4% of evaluations found to consider humanitarian principles well in the United Nations Evaluation Group paper discussed above (2016) but this is likely due to the small sample size and its composition.

Table 15: Organisational evaluations on coherence

Organisation	Title	Focus Area	Humanitarian principles	Other Areas
UNHCR	Evaluation of UNHCR's Level-3 Emergency Response to Cyclone Idai (Mozambique, Malawi, Zimbabwe) (2021)	Internal and external policy coherence	Not mentioned	Notes challenges of coherence across countries/regions
UNICEF	Review of the L2 Response in Venezuela (2020)	Internal and external activity coherence	Evaluated (under coherence +)	
AHP	Response to the Rohingya Humanitarian Crisis Phase II Evaluation Final Report (2021)	External programme and policy coherence	Not mentioned	Assesses engagement with coordination mechanisms
IFRC	Real-Time Evaluation Indonesia: Earthquakes and Tsunami (Lombok, Sulawesi) (2019)	External programme coherence	Not mentioned	Assesses engagement with coordination mechanisms
UNHCR	Evaluation of the UNHCR Regional Response to the Venezuela Situation (2020)	Internal programme and policy coherence	Not mentioned	-
WFP	Corporate Emergency Evaluation of the WFP Regional Response to the Syrian Crisis (2018)	External policy and programme coherence	Evaluated (under results delivered)	Notes complexity and sensitivity of political context
European Union	Evaluation of the European Union's Humanitarian Interventions in India and Nepal, 2013-2017 (2018)	Internal and external policy coherence with other actors	Evaluated (under coherence +)	-
DFID	Evaluation of DFID's Humanitarian Response to Typhoon Haiyan (Yolanda) (2015)	External programme coherence	Mentioned (very light touch)	Assesses capacity to influence priority sectors
European Union	Evaluation of the European Union's Humanitarian Response to the Refugee Crisis in Turkey Final Report (2019)	Internal and external policies and planning coherence	Mentioned (very light touch)	-
IFRC	Evaluation of IFRC West Africa Ebola Viral Disease Appeal Response Sierra Leone and Liberia (2018)	External policy and planning coherence	Not mentioned	-

Summary and questions to explore

Coherence is one of the least-applied criteria in humanitarian evaluations and is sometimes assessed as part of other criteria (Darcy and Dillon 2020; ALNAP 2006 and ALNAP 2018a). While it has long been a criterion for evaluating humanitarian action, it has only recently been added to the OECD DAC evaluation criteria for development. In the evaluations reviewed, coherence as a criterion was represented mainly in UN, IFRC or donor-led evaluations, supporting literature findings on the complexity and challenge of assessing coherence for a single organisation or programme.

- **What are some of the reasons that might account for why coherence is among the least applied, and least included in guidance?**
- **When is it most useful or necessary? Are there circumstances or types of evaluations where it is less applicable?**

The literature and other guidance focus on the sub-division of coherence into internal coherence with organisations' own policies and standards, and external coherence with other actors and standards, with varying levels of focus on coherence with humanitarian principles.

- **Do these elements of coherence resonate? Are they equally important?**
- **Could the definition be more precise or should it remain broad?**

Humanitarian principles have become more prominent in ALNAP's guidance and application of the coherence criterion. As with coherence more broadly, humanitarian principles are both debated and challenging to implement (see, for example, Slim 2020 and Buchanan-Smith 2022 in relation to the Ukraine conflict). A United Nations Evaluation Group working paper (2016) advocated for humanitarian principles to be embedded systematically as core elements of the evaluation of humanitarian action.

- **How important is it that humanitarian principles are evaluated as part of coherence? Should guidance recommend it?**



**CONNECTNESS
AND
SUSTAINABILITY**



CONNECTEDNESS AND SUSTAINABILITY

LEAST USED OECD DAC CRITERIA

Definitions

ALNAP Guide 2006		OECD DAC Criteria 2019	
<p>Connectedness: Refers to the need to ensure that activities of a short-term emergency nature are carried out in a context that takes longer-term and interconnected problems into account.</p>		<p>Sustainability: Will the benefits last? The extent to which the net benefits of the intervention continue or are likely to continue.</p>	
ALNAP EHA Guide 2016	ALNAP SOHS 2018	IASC's IAHE 2018	CHS 2018
<p>Connectedness: The extent to which activities of a short-term emergency nature are carried out in a context that takes longer-term and interconnected problems into account. Replaces the sustainability criterion used in development evaluations.</p>	<p>Connectedness: The degree to which the international humanitarian system articulates with development, resilience, risk reduction and peacebuilding.</p>	<p>Sustainability: What were the positive and negative, intended and unintended effects of the IASC humanitarian system's assistance for people affected by the crisis?</p>	<p>Commitment 3: Communities and people affected by crisis are not negatively affected and are more prepared, resilient and less at-risk as a result of humanitarian action. Quality criterion: Humanitarian response strengthens local capacities and avoids negative effects.</p>

Sector-wide guidance and standards

ALNAP

In line with OECD DAC's 1999 *Guidance for evaluating humanitarian assistance in complex emergencies*, ALNAP's 2006 guide replaced the 1991 OECD DAC criterion of sustainability with connectedness. ALNAP's guidance noted both the lack of consensus on the extent to which humanitarian action should support longer-term needs as required by the sustainability criterion, and the challenge in achieving sustainability with short-term interventions. The connectedness criterion was intended to focus on linkages between relief and recovery, such as exit strategies and linkages between programme outputs and outcomes. The ALNAP guidance further highlights the importance of partnerships and the development of local capacity in ensuring the effects of interventions are not lost. ALNAP's 2016 guidance did not further update the definition.

In 2018, ALNAP's SOHS defines connectedness as the degree to which the humanitarian system links with development and specifies linkages to resilience, risk reduction and peacebuilding. It intentionally separated out the relationship between international humanitarian action and national and local capacities from the criterion of connectedness, and explored this under an additional criterion – complementarity. It further specified that alignment with humanitarian principles, international humanitarian law and refugee law would be explored under the criterion of coherence.

Also in 2018, ALNAP's *Evaluation of protection in humanitarian action* (ALNAP 2018b) takes a slightly different approach. It explains connectedness as humanitarian interventions that take account of other key actors and efforts. It splits this in to two parts: coordination within 'the system', to the extent that it is appropriate, feasible and desirable; and coordination with the broader protection environment.

OECD DAC criteria

As part of the definition of sustainability outlined at the beginning of this chapter, the OECD DAC guidance (2021: 45) includes the following note:

Includes an examination of the financial, economic, social, environmental and institutional capacities of the systems needed to sustain net benefits over time. Involves analyses of resilience, risks and potential trade-offs. Depending on the timing of the evaluation, this may involve analysing the actual flow of net benefits or estimating the likelihood of net benefits continuing over the medium and long term.

The OECD DAC's update to the criterion in 2019 intentionally removed the previous emphasis on donors and external funding (by removing reference to the question: 'To what extent did the benefits of a programme or project continue after donor funding ceased?').

Of perhaps most relevance to humanitarian action, its 2021 guidance emphasises exit planning as a key aspect of sustainability.

Inter-Agency Standing Committee's IAHE 2018

The IAHE definition of sustainability asks the evaluator to establish the effects of the assistance for people affected by crisis, rather than requiring the net benefits to continue. Its explicit reference to negative as well as positive, and unintended as well as intended, effects, is unique compared with the ALNAP and OECD DAC definitions, which more explicitly explore these elements under the impact criterion.

Core Humanitarian Standard

CHS Commitment 3 provides some crossover with other definitions of connectedness, particularly the IAHE guidelines. It more explicitly emphasises the perspective of communities and people affected by crisis.

Key issues arising

The humanitarian sector uses both connectedness and sustainability as criteria when evaluating humanitarian action. Perhaps surprisingly, sustainability has been used slightly more often than connectedness (Darcy and Dillon 2020). ALNAP's working paper on evaluating humanitarian innovation for example, uses the criterion of sustainability rather than connectedness (Obrecht, Warner and Dillon 2017). Considered separately, connectedness and sustainability appear to be the least used of the OECD DAC criteria at around 35% and 30% respectively (Darcy and Dillon 2020). However, it is likely that some of these evaluations used either sustainability or connectedness as criteria and that, combined, the regularity of their use could be much higher.

Connectedness is one of the less settled of the OECD DAC criteria for evaluating humanitarian action. The different components emphasised in definitions of connectedness (and sustainability) are apparent in the sector-wide guidance and standards. While ALNAP in 2006 and 2016 focuses on taking long-term issues into account, in the 2018 SOHS report its definition specifies linkages with development, resilience, risk reduction and peacebuilding. The IAHE guidelines and CHS included intended and unintended effects. The CHS differs again in its focus on the perspective of communities and affected populations. The CHS definition is perhaps closest to the current OECD DAC definition in that it includes components of lasting benefit: increased preparedness and reduced risk.

The literature identifies a number of issues relating to the evaluation of connectedness and sustainability. In development contexts, for example, Heider (2017) finds that it is both difficult and costly to evaluate all dimensions of sustainability systematically. Similarly, a meta-evaluation of German development cooperation evaluations found that sustainability is being assessed unsystematically and inconsistently, suggesting this may be due to the absence of a common conceptual framework (Noltze et al. 2018).

Debates on the OECD DAC evaluation criteria in the lead-up to the 2019 revision (OECD 2018) included similar critiques regarding a lack of systematic analysis, as well as the need increase linkages to impact, and distinguish between sustainability during implementation and future sustainability. Discussions (OECD 2018) also raised the possibility of integrating sustainability into other OECD DAC evaluation criteria. While originally made in the context of development assistance, this comment reflects one of the more consequential issues regarding application of the criteria when evaluating humanitarian action: that the humanitarian sector continues to debate whether connectedness and sustainability are critical to all humanitarian action.

This paper, therefore, focuses on the inclusion of connectedness or sustainability as separate criteria, and on the unsystematic and variable components of connectedness and sustainability apparent in definitions and application of the criteria.

The humanitarian sector continues to debate whether connectedness and sustainability are critical to all humanitarian action

The OECD DAC adapted the sustainability criterion to connectedness in the context of complex emergencies because 'many humanitarian interventions, in contrast to development projects, are not designed to be sustainable' (OECD 1999). The criterion remains problematic. ALNAP's 2018a SOHS report was clear: connectedness is 'unusual among the OECD DAC performance criteria in that there is disagreement as to whether it should be used as a measure of humanitarian performance at all'.

The 2018 SOHS noted both the long history and continued debate regarding the relationship between humanitarian action and development, including its prominence during the 2016 World Humanitarian Summit. The International Committee of the Red Cross (ICRC) requested the Summit to recognise the role of humanitarian action in 'meeting urgent needs, rather than ending needs' (ICRC 2016). One of the concerns is the tension with humanitarian principles. Involvement in development processes, usually driven by government priorities, could challenge the humanitarian principles of impartiality and of independence (ALNAP 2018a).

When selecting criteria for evaluating humanitarian, action however, ICRC and MSF allow for similar approaches to those of the OECD DAC and ALNAP. MSF uses a definition for connectedness similar to ALNAP's 2006 guide, though calls it 'continuity' (MSF 2017), while ICRC includes the option to evaluate the sustainability criterion, depending on the length and objective of the intervention (ICRC 2022).

Variable components of the definition

When outlining the rationale for the connectedness criterion in humanitarian action, the OECD DAC (1999) provided examples of humanitarian interventions that had *unintended negative effects*: large refugee camps resulting in severe environmental impacts, food distribution damaging roads and private employers offering high salaries poaching staff from government clinics and schools leaving the local population with reduced levels of service. Consequently, just

one year after the publication of the ALNAP 2006 guide, the OECD DAC discussed replacing the term 'connectedness' with 'linkages' (OECD 2007) in the context of conflict and peace, to focus the meaning on linkages between activities and policies at different levels and across sectors, including with longer-term development processes.

While such differences in the definition of connectedness and sustainability do not appear to be debated extensively in the literature, they are quite apparent in contemporary guidance and their application in evaluations. There are four components to connectedness and sustainability that are variably applied:

- a. Consideration of longer-term and interconnected problems (ALNAP 2006 and 2016 definition).
- b. Links with development, resilience, risk reduction and peacebuilding (ALNAP 2018a) lasting benefits (OECD 2019).
- c. Positive and negative, intended and unintended effects (IAHE; similarities with CHS Commitment 3).
- d. The relationship between international humanitarian action and national and local capacities (ALNAP 2018a).

Organisational guidance

An analysis of guidance for evaluating humanitarian action found that, of the 10 guidance documents reviewed, most included both sustainability and connectedness as criteria. The majority tend to follow ALNAP's 2006 approach to defining the terms (connectedness: considering longer-term and interconnected problems), the OECD DAC's 2019 approach (sustainability: focusing on lasting benefits) or the ALNAP (2018a) SOHS approach (connectedness: considering linkages with development etc.).

The CHS and IAHE approach to emphasising the effects on the affected population, including both intended and unintended effects, is less common. No agency guidelines emphasised the effects on the affected population, and just two of the 10 reviewed included reference to intended and unintended effects (the USAID Bureau for Humanitarian Assistance and UNICEF's 2011 guidance for equity-focused evaluations).

Local capacities or partnerships are important components of connectedness and sustainability across most guidelines, being referenced in seven of the 10 guidelines reviewed, while exit plans or strategies were referenced in just three of the 10. For comparison, the IAHE guidelines (OCHA 2018) separated out partnerships, localisation and coordination, all as separate criteria.

Table 16: Organisational guidance on connectedness/sustainability

Organisation	Connectedness or sustainability	Component: (a) considers longer-term and interconnected problems, (b) links with development etc., (c) lasting benefits or (d) un/intended effects	Explicit emphasis on affected populations in definition	Local capacity, partnerships, resilience, exit strategies
ALNAP (2006)	Connectedness	(a) considers longer-term and interconnected problems	No	Local partnerships, local capacity, exit strategies
OECD (2019)	Sustainability	(c) lasting benefits	No	Exit planning, local partner capacity
IAHE (2019)	Sustainability	(d) un/intended effects	Yes	N/A
IFRC (2011)	Connectedness and sustainability	(a) considers longer-term and interconnected problems (b) links with development etc. (specifically recovery) (c) lasting benefits	No	Local capacity and ownership, exit strategies
Save the Children (n.d.)	Sustainability	(c) lasting benefits	No	Local partner capacity
ACT Alliance (2012)	Connectedness and sustainability	(a) considers longer-term and interconnected problems (c) lasting benefits	No	Community ability to sustain, project support to community
Catholic Relief Services (2012)	Connectedness and sustainability	N/A	N/A	N/A
WFP (2021)	Connectedness and/or sustainability	(a) considers longer-term and interconnected problems (b) links with development etc. (c) lasting benefits	No	Exit strategies, capacity building of national and local partners and communities
Médecins Sans Frontières (2017)	Continuity (also referred to as connectedness)	(a) considers longer-term and interconnected problems	No	Exit strategies, connections with local capacities
UNICEF (2014)	Itainability	(c) lasting benefits	No	Stakeholder ownership and capacity, local capacity
USAID BHA (2022)	Sustainability	(b) links with development etc. (c) lasting benefits (d) un/intended effects for targeted communities	No	Exit strategies, integrate host government services
UNICEF – Equity Evaluations (2019)	Connectedness and sustainability	(a) considers longer-term and interconnected problems (b) links with development etc. (c) lasting benefits (d) un/intended effects for targeted communities (focus on inequities)	No	Specifies 'worst-off groups'
UNFPA Evaluation Handbook (2019)	Sustainability	(c) lasting benefits	No	No
	Connectedness	(a) considers longer-term and interconnected problems	No	Yes

Application in evaluations

This paper reviews 10 evaluations that assessed either the connectedness or sustainability of humanitarian action. Of the selection, sustainability was applied in five of the evaluations, connectedness in three of the evaluations and both sustainability and connectedness in two of the evaluations. Where sustainability was used as the criterion, either together with or instead of connectedness, the definition applied by the evaluation is interpreted in accordance with the programme that is being evaluated. For example, in the sample, resilience appears to be more of a focus in responses to disasters caused by natural hazards, as compared to crises caused by conflict.

In terms of the components of the definitions of connectedness and sustainability outlined above, the most common application of the criteria included a version of ‘lasting benefits’ (60%), similar to the OECD DAC criteria 2019, although not necessarily using those exact words. Links with development, resilience, risk reduction and peacebuilding were less common (40%), followed by consideration of longer-term and interconnected problems (30%) and an assessment of positive and negative and intended and unintended effects (20%).

Even where evaluations focused on components of the definition other than ‘lasting benefits’, they were often assessed from that perspective. For example, the IAHE evaluation of the drought response in Ethiopia focused on whether the response helped to build resilience, which also contributed to lasting benefits. The evaluation questions included whether the response helped affected people to cope better with subsequent droughts.

In line with the guidance, most evaluations incorporated elements of partnership, local capacity and exit strategies as part of their assessment of connectedness or sustainability. IAHE evaluations added separate criteria for partnerships and localisation, reflecting the additional criteria in its guidance. Similar to ALNAP’s approach to the SOHS in 2018, the WFP and DEC evaluations separated out the criterion of complementarity. This is discussed further in Chapter 9: Additional criteria and cross-cutting themes.

Table 17: Organisational evaluations on connectedness and sustainability

Organisation	Title	Criterion	Emphasis on component of definition	Local capacity, partnerships, resilience, exit strategies
IAHE	Evaluation of the Drought Response in Ethiopia 2015–2018 (2019)	Sustainability	(b) links with development etc. Evaluation question: did the response help to build resilience?	Yes. Focused on resilience. Includes strengthening government service provision, supporting affected people to cope and links to recovery and development actors. Evaluated local capacity & partnerships and coordination as separate criteria
IAHE	Humanitarian Evaluation of the Response to Cyclone Idai in Mozambique (2020)	Connectedness	(b) links with development etc.	Focus on transition to development; resilience. Partnerships, localisation and coordination separate criteria
CARE	Multisectoral Protection Response for Vulnerable Populations in Ecuador Affected by the Humanitarian Crisis (2020)	Sustainability	(c) lasting benefits (medium and long-term impact)	Institutional and social capacities built by the project
DEC	Real-Time Review of DEC’s Response to Cyclone Idai (2019)	Sustainability and connectedness	(a) connected to longer-term issues (c) lasting benefits (maintain positive effects) (d) unintended effects (any negative impacts)	Linking to longer-term programmes, and ‘build back better’. Strong focus on the environment. Coordination and complementarity separate criteria

WFP	Evaluation of the WFP Regional Response to the Syrian Crisis (2015–2018) (2018)	Sustainability	(b) links with development, etc. (c) lasting benefits (considers sustainable gains for communities, risk management and future planning)	Yes, implicitly e.g., exit strategies in terms of integration into local safety nets; complementarity a separate criterion
European Union	Evaluation of the European Union's Humanitarian Response to the Refugee Crisis in Turkey (2019)	Sustainability/ Connectedness	(b) links with development etc. (humanitarian development nexus) and (c) lasting benefits (implicitly in findings; exit strategy and handover; integration with government)	Exit strategies, government capacity
The National Health Cluster in Yemen	Assessing the Coordination of Health Response During Humanitarian Crises (2022)	Connectedness	Defined as linking with long-term plans, and findings included (d) unintended effects	Exit strategies. Focused on local capacities of MoH to continue. Participation evaluated as separate criterion
Oxfam	Missed Out: The Role of Local Actors in the Humanitarian Response in the South Sudan Conflict (2016)	Connectedness	(a) takes longer-term and interconnected problems into account	Evaluation focused on the role of local actors. Findings note importance of longevity of partnerships; community requests for DRR and recovery
Plan International UK	Plan International UK's DEC-Funded Response to the Rohingya Refugee Crisis in Bangladesh (2018)	Sustainability	(a) takes longer-term and interconnected problems into account (in sub-question) (c) lasting benefits after donor funding is withdrawn. Also looked at external factors influencing sustainability	Yes, CHS 3. Capacities of affected populations and groups; service providers and government institutions
OFDA	Final Evaluation of OFDA Response Program July to October 2021 (CARE Turkey and partners in NorthWest Syria) (2021)	Sustainability (reported with impact)	(c) lasting benefits (e.g., sustainable infrastructure)	To some extent, in terms of ability/capacity to meet basic needs

Summary and questions to explore

Connectedness is among the least-used criterion in evaluating humanitarian action. A 2020 review by ALNAP found that sustainability has been applied slightly more often than connectedness in humanitarian evaluations (Darcy and Dillon 2020).

Application of this criterion is complicated by the related terms used (connectedness or sustainability), and differences in the definitions thereof. The more common definitions reflect either the ALNAP and OECD DAC definitions listed above, or the definition provided by ALNAP in the 2018 SOHS report: the degree to which the international humanitarian system articulates with development, resilience, risk reduction and peacebuilding (35).

- **Given that sustainability has been applied more often than connectedness in humanitarian evaluations, should the term ‘sustainability’ be adopted for this criterion, along with the updated OECD DAC definition?**
- **What elements of connectedness are most useful to evaluating humanitarian action or essential to keep or emphasise further? (a) lasting benefits, (b) considers longer-term and interconnected problems, (c) links with development, including recovery, resilience, risk reduction and peacebuilding, (d) contribution to stakeholder ownership, local capacities and local partnerships.**

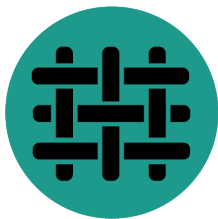
Some organisations include local capacities and external partnerships as foundational to connectedness/sustainability, while others (ALNAP 2018a, and in evaluations by the WFP and the Disasters Emergency Committee) found this element important enough to be considered under a separate criterion – complementarity.

- **Should local capacities and partnerships be highlighted under the criterion of connectedness? Is it sufficiently critical to effective humanitarian action that it should be further elevated to an additional or cross-cutting criterion, such as complementarity?**

Connectedness and sustainability are much debated. It is ‘unusual among the OECD DAC performance criteria in that there is disagreement as to whether connectedness should be used as a measure of humanitarian performance at all’ (ALNAP 2018a: 239). This disagreement is anchored in the long history and continued debate regarding the relationship between humanitarian action and development (ALNAP 2018a).



**ADDITIONAL
CRITERIA AND
CROSS-CUTTING
THEMES**



ADDITIONAL CRITERIA AND CROSS-CUTTING THEMES

This section discusses the approaches taken to additional criteria and cross-cutting themes when evaluating humanitarian action. It begins by identifying the approaches taken by sector-wide guidance and standards and exploring how this has been applied in individual organisational guidance and evaluations. It identifies options for additional criteria and cross-cutting themes to be considered when developing future guidance. It includes a more detailed analysis of two of the more common cross-cutting themes/additional criteria – inclusion and accountability to affected populations – and the emerging theme of adaptive management.

As approaches and terminology vary widely, this paper applies the following terminology to assist with the discussion:

- **Additional criteria:** where guidance or evaluations identify the theme formally as an additional criterion. For example, IAHE process guidelines includes coordination and localisation as stand-alone criteria in addition to criteria derived from the OECD DAC 2019 and ALNAP 2006 guidance.
- **Cross-cutting themes:** where guidance or evaluations formally identify the issue as one which should be addressed under multiple criteria. For example, ALNAP's 2006 guide provides a list of eight cross-cutting themes.
- **Additional considerations:** where guidance or evaluations formally identify an issue as important to the evaluation, but do not direct users to apply the issue across all criteria. For example, the OECD DAC emphasises the importance of the Sustainable Development Goals, but they are not included as cross-cutting themes in the OECD's 2019 *Better criteria for better evaluation revised evaluation criteria definitions and principles for use*

Many of the themes and considerations identified in this section are incorporated under one or two specific criteria in the guidance documents or evaluations. When used in that way, they are not considered a cross-cutting theme for the purpose of this paper.

Approach

ALNAP Guide 2006		OECD DAC Criteria 2019	
Cross-cutting themes: Local context, human resources, protection, participation of primary stakeholders, coping strategies and resilience, gender equality, HIV/AIDS and the environment.		Additional considerations: Gender, equity, inclusion, the 2030 Agenda and Sustainable Development Goals (OECD DAC 2021).	
ALNAP EHA Guide 2016	ALNAP SOHS 2018	IASC's IAHE 2018	CHS 2018
Additional criterion: Coordination	Additional criteria: Sufficiency, complementarity, accountability and participation	Additional criteria: Partnerships, localisation and coordination	Additional criteria: (In brief) coordination, complaints, participation and feedback, learning and improvement, staff support, strengthened local capacities and avoids negative effects
Additional considerations: gender equality, advocacy, consideration of vulnerable and marginalised groups		Additional considerations: Access and protection, collective responses for collective outcomes, gender, inclusiveness and accountability to affected populations	Additional considerations: Principled humanitarian action, gender, protection from sexual exploitation, harassment and abuse

Sector-wide guidance and standards

ALNAP

ALNAP's approach to cross-cutting themes has varied over time. The eight possible cross-cutting themes identified in ALNAP's 2006 guide were: local context, human resources, protection, participation of primary stakeholders, coping strategies and resilience, gender equality, HIV/AIDS and the environment. It suggests that many of these can be considered under the relevance/appropriateness criterion. While it acknowledges that not every cross-cutting theme needs to be included, it recommends providing a rationale for those that are excluded.

Although ALNAP has not formally updated its list of cross-cutting themes since 2006, many of the original eight mentioned above were not included in subsequent ALNAP guidance. ALNAP's 2009 guidance on Real-Time Evaluations lists coping strategies, climate and environment, gender and inclusion as additional considerations. ALNAP's 2016 *Evaluating humanitarian action guide* provides tools to evaluate gender and vulnerability as cross-cutting issues.

In 2018, ALNAP's SOHS report included three additional criteria: sufficiency, complementarity and accountability & participation. It focused on three cross-cutting themes under the effectiveness criterion – disability, gender and age – highlighting protection considerations for each.

OECD DAC

Although the OECD DAC guidance does not use the terminology ‘cross-cutting themes’, it explicitly aims to better respond to equity, gender equality and the ‘leave no one behind’ imperative. The definitions of both relevance and effectiveness in particular are designed to encourage more in-depth analysis of these themes (2019). In its 2021 guidance, the OECD DAC includes a section on how to integrate inclusion under each criterion. It provides advice on applying a gender lens and encourages evaluators to ‘work in ways that thoughtfully consider differential experiences and impacts by gender, and the way they interact with other forms of discrimination in a specific context (e.g., age, race and ethnicity, social status)’ (OECD 2021: 32).

Inter-Agency Standing Committee’s IAHE 2018

The IAHE process guidelines (2018) propose quite different criteria from ALNAP and the OECD DAC. It applies six criteria for crisis-specific evaluations. These six criteria do not include the ALNAP and OECD DAC criteria of efficiency, coherence or impact. It adds the criteria of partnerships, localisation and coordination.

In terms of cross-cutting considerations, the IAHE process guidelines require that evaluations ‘systematically give particular attention to’ access and protection, and collective responses for collective outcomes (OCHA 2018: 17). It includes as ‘special considerations’ requirements for gender, inclusiveness and accountability to affected populations, for example:

- gender analysis is to be applied in all phases of the evaluation, and
- the evaluation process must seek to understand how the crisis response enhanced equitable and effective inclusion, access and participation in humanitarian activities and decision-making, and
- how various segments of the affected population are consulted in the prioritisation of needs and decision-making processes.

Core Humanitarian Standard

The CHS, with its nine commitments and quality standards, incorporates a number of additional commitments beyond the ALNAP 2006 criteria, including (in brief): coordination, complaints, participation and feedback, learning and improvement, staff support, strengthened local capacities and avoiding negative effects. It specifies the additional consideration of principled humanitarian action; that humanitarian action is guided by the humanitarian principles (humanity, neutrality, impartiality and independence) and that the humanitarian principles are integrated throughout the CHS (2014). Gender is mentioned across four of the nine commitments (CHS 2018). Protection from sexual exploitation, harassment and abuse is included across each of the nine commitments in the CHS verification framework (CHS 2020).

Key issues arising

Sector-wide guidance and additional criteria

Each of the sector-wide guidance documents takes a different approach to both criteria and cross-cutting themes. In terms of **additional criteria**, the greatest deviation from the ALNAP 2006 and OECD DAC 2019 criteria are the three additional criteria in the 2018 SOHS, the three additional criteria in the IAHE guidelines (OCHA 2018) and the five additional criteria

in the CHS (2018). There is greatest similarity across themes of coordination, localisation/ complementarity with national and local actors, and accountability and participation. See [Table 18](#) for more detail. It is notable that inclusion, equity and gender equality were not included as additional criteria in any of the guidance.

Table 18: Definitions of additional criteria in sector-wide guidance

Criterion	Source	Definition
Coordination	ALNAP (2016)	The extent to which the interventions of different actors are harmonised with each other, promote synergy, avoid gaps, duplication and resource conflicts (often folded into effectiveness)
	IAHE (2018)	Was the assistance well-coordinated, successful and, as much as possible, equitable, reaching all affected populations and avoiding duplication of assistance and gaps?
	CHS (2018)	Communities affected by crisis receive coordinated, complementary assistance. Quality criterion: humanitarian response is coordinated and complementary
Partnerships	IAHE (2018)	To what extent have adequate partnerships been established (with international, national and/or local stakeholders) to deliver assistance to affected people?
Localisation	IAHE (2018)	Have national and local stakeholders been involved in the response design and have their capacities and systems to respond in the future been strengthened through the response?
Complementarity	SOHS (2018)	The degree to which the international humanitarian system recognises and supports the capacities of national and local actors, in particular governments and civil society organisations
Accountability & participation	SOHS (2018)	The degree to which actors within the international humanitarian system can be held to account by crisis-affected people, and the degree to which crisis-affected people are able to influence decisions related to assistance and protection
	CHS (2018)	Communities and people affected by crisis know their rights and entitlements, have access to information and participate in decisions that affect them. Quality criterion: humanitarian response is based on communication, participation and feedback
	CHS (2018)	Communities and people affected by crisis have access to safe and responsive mechanisms to handle complaints. Quality criterion: complaints are welcomed and addressed
Sufficiency	SOHS (2018)	The degree to which the resources available to the international humanitarian system are sufficient to cover humanitarian needs
Improvement	CHS (2018)	Communities and people affected by crisis can expect delivery of improved assistance as organisations learn from experience and reflection. Quality criterion: humanitarian actors continuously learn and improve
Competent staff	CHS (2018)	Communities and people affected by crisis receive the assistance they require from competent and well-managed staff and volunteers. Quality criterion: staff are supported to do their job effectively and are treated fairly and equitably

Sector-wide guidance and cross-cutting themes

The approach to cross-cutting themes across the sector-wide guidance is inconsistent. ALNAP in 2006, for example, nominated and explained eight cross-cutting themes. In contrast, the update to the OECD DAC criteria took a minimalist approach, not formally nominating 'cross-cutting themes', but applying a gender lens and integrating inclusion throughout (2019). The IAHE guidelines and the CHS also do not nominate 'cross-cutting themes' but use alternate language to identify a select list of priority themes.

Sphere for monitoring and evaluation (Mountfield 2015) identifies cross-cutting themes that

address specific needs or considerations – children, gender, people living with HIV and AIDS, older people and people with disabilities – and cross-cutting themes related to external factors – disaster risk reduction, including climate change issues, the environment and psychosocial support.

Cross-cutting themes and considerations are often addressed in guidance under individual criteria. ALNAP's 2006 guide, for example, suggests that many of the cross-cutting themes it identified can be covered under the relevance criterion. Cross-cutting themes are also less likely to be defined than evaluation criteria. ALNAP's 2006 guide does provide an explanation of each of the cross-cutting themes it nominates.

Organisational guidance

Gender, inclusion, protection, accountability and participation are common evaluative cross-cutting themes in individual organisational guidance. Guidance from UN agencies primarily focuses on gender, inclusion and accountability, with the United Nations Evaluation Group (UNEG) Norms and standards for evaluations (UNEG 2016) recognising that human rights and gender equality (Norm 8) must be integrated into all stages of an evaluation. UNEG (2016) also notes that evaluators and evaluation managers are responsible for ensuring human rights and gender equality are respected, addressed and promoted, in line with the principle of 'leave no one behind'.

UNICEF's *How to design and manage equity-focused evaluations* (Bamberger and Segone 2011) offers advice on ensuring equity and inclusion as cross-cutting themes in evaluations, while also recommending the addition of equity and inclusion as separate criteria. UNICEF's *Evaluative criteria* (Peersman 2014) extends the equity narrative to 'promoting the equal rights of women and girls, and supporting their full participation in the political, social and economic development of their communities' (2). The WFP's *Evaluation policy* (2022) does not provide any specific guidance on or definitions of cross-cutting themes but lists some to be considered across the evaluation process: accountability to affected populations, climate/environment, gender and protection.

An analysis of guidance from NGOs revealed similar trends. While the Save the Children *Evaluation handbook* (Save the Children 2012) has a development focus, it highlights cross-cutting themes applicable to humanitarian evaluations: accountability, gender, inclusion and participation. Considerable attention is paid to non-discrimination as a cross-cutting action that strengthens inclusion, participation and accountability. Catholic Relief Services' *Guidance on monitoring and evaluation* (2012) provides substantive advice on assessing accountability, gender and participation as cross-cutting themes in humanitarian evaluation. Inclusion is a part of coverage according to the MSF *Evaluation manual* (2017), which interprets it as an examination of who has been included or excluded from a programme in terms of ethnicity, gender, socioeconomic status, occupation, location or family circumstances. Participation is noted, mainly in relation to involving all relevant stakeholders in the evaluation process as far as feasible and appropriate.

The IFRC *Framework for evaluations* (2011) references inclusion, defining it as the consultation and involvement of relevant stakeholders in the evaluation process, with attention paid to beneficiaries, particularly marginalised or vulnerable groups. The framework suggests a rights-based approach to the treatment of equity in evaluations, with close attention given to the potential effects of differences and inequalities in society related to race, age, gender, sexual orientation, physical or intellectual ability, religion and socioeconomic or ethnic background.

Application in evaluations

Forty evaluations were analysed for the occurrence of additional criteria and cross-cutting themes. Of those 40 evaluations, the most common additional criterion used was coordination, likely reflective of its inclusion in the IAHE guidelines, CHS and ALNAP 2016 guidance. Aside from coordination (occurring in seven of 40 evaluations), it was relatively uncommon for additional criteria to be applied. At just two instances each, the only additional criteria applied in the 40 evaluations were: added value, adherence to the CHS, inclusion, localisation, partnerships and protection.

In terms of the most consistent cross-cutting themes applied in the 40 evaluations, the most common were gender (in 14 evaluations) and inclusion (13), followed by protection (12), accountability to affected populations (9) and participation (7). Gender, protection and participation were originally cross-cutting themes in ALNAP's 2006 guide.

Of the remaining eight original cross-cutting themes identified by ALNAP, there were occasional references to coping strategies and resilience (5) as cross-cutting themes, as well as human resources/adaptive management (5) and communication with communities (3). Evaluations only infrequently applied the local context (1) or the environment (2) as cross-cutting themes, and no reference was found to HIV/AIDS. See [Table 19](#) for more detail.

The sample of evaluations analysed for this paper did not include evaluations focused on cross-cutting and thematic issues. It is not uncommon for larger agencies to conduct evaluations focused on such themes. Examples include:

- *Evaluation of diversity, inclusion and accountability to affected populations in ICRC operations* (Meier et al 2018).
- *Inter-agency humanitarian evaluation on gender equality and the empowerment of women and girls* (IASC 2020a).
- *Evaluation of the UNFPA capacity in humanitarian action (2012–2019)* (UNFPA Evaluation Office 2019).
- *Evaluation of the FAO strategy for partnerships with the private sector* (FAO 2019).
- *Evaluation of the WFP People Strategy (2014–2017)* (WFP Office of Evaluation 2020).

Table 19: Application of additional criteria and cross-cutting themes in 40 evaluations

	Included as own criteria	Included as cross-cutting theme	Included under DAC criteria	Included in some other form	Total
Inclusion	2	13	10	5	30
Accountability to affected populations	1	9	6	5	21
Protection	2	12	6	1	21
Gender	1	14	4	1	20
Coping strategies and resilience		5	10	3	18
Adaptive management/HR		5	8	4	17
Coordination	7	3	1	1	12
Communication with communities		3	6	2	11
Participation		7	3		10
Core Humanitarian Standard	2		2	1	5
Localisation	2	2	1		5
Climate/environment		2		2	4
Equity		1	1	1	3
Humanitarian principles			1	2	3
Local capacity/capacity building	1	2			3
Added value	2				2
COVID-19	1	1			2
Partnerships	2				2
Value for money	1	1			2
Leave no one behind		1			1
Local context		1			1
Complementarity	1				1
Humanitarian–development nexus			1		1
Durable solutions	1				1
Timeliness	1				1
Transparency	1				1
Triple nexus (HDP)	1				1
HIV/AIDS					0
Transformation (transformational evaluation)					0

Options for future guidance

There is a wide diversity in approach to additional criteria, cross-cutting themes and additional considerations. It is here that the observation of Cosgrave, Ramalingam and Beck (2009) ring particularly true: any given set of evaluation criteria reflects recurring areas of weakness and the main concerns of the humanitarian sector at the time when it is being developed. The guidance and evaluations reviewed span more than two decades, with various concerns of the humanitarian sector receiving greater attention at different stages and by different actors. For example, the consideration of HIV/AIDS did not occur once across the 40 evaluations reviewed for this paper (see [Table 19](#)), though must have been a prominent concern in 2006 when it was included as a cross-cutting theme for humanitarian action.

There are, however, some clear trends that future guidance may wish to consider incorporating as additional criteria or cross-cutting themes: 1) ways of working with national and local actors (including coordination), 2) ways of working with communities (such as accountability to affected populations) and 3) diversity, equity and inclusion. Protection, as the third most-applied cross-cutting theme will also need to be addressed, given the centrality of protection and the challenges in its evaluation (ALNAP 2018b).

Criteria

There is a trend in multisectoral guidance towards additional criteria that evaluate a way of working that is coordinated and complementary, operates in partnership, is accountable and strengthens national actors and local communities. This is reflected in the additional criteria in 2018 guidance published by the OCHA (partnerships, localisation and coordination), ALNAP (sufficiency, complementarity, accountability and participation) and CHS (in brief: coordination, complaints, participation and feedback, learning and improvement, staff support, strengthened local capacities and avoids negative effects). In practice, coordination is the additional criterion most often applied in evaluations – see [Box 6](#) below.

Box 6: Coordination

The original adaptation of the OECD DAC criteria for evaluating humanitarian assistance in complex emergencies (1999) highlighted the importance of explicitly considering coordination. ALNAP (2006 and 2018) includes coordination under the effectiveness criterion, as do MSF (2017) and IFRC (2011). Others, including the IAHE (OCHA 2018), CHS (2018), UNFPA (2019) and UNICEF's Office of Research (Peersman 2014) include coordination as its own criterion – as did 18% of the evaluations reviewed (see [Table 19](#)). Darcy and Dillion similarly found that around 20% of evaluations included coordination (2020). Refer to *Chapter 2 on the effectiveness criterion for further discussion of coordination*.

Future guidance will need to be explicit about how to best address coordination. There is merit in either emphasising coordination as a critical component of effectiveness or including it as an additional criterion. A preference for fewer, rather than more, criteria may be the determining factor.

Coordination is not sufficient on its own to reflect the trends towards working in ways that support and strengthen local and national actors, accountability and partnership. This could be explored through ways of working with local and national actors, and ways of working with communities.

Ways of working with communities is well-established through the concepts of accountability to affected populations, participation, and communication with communities. These concepts were the second most-applied as a cross-cutting theme in evaluations – in almost 50% of the evaluations reviewed. They were also included in additional criteria in the SOHS report (ALNAP 2018a) and the CHS (2018). Future guidance should explore how to best incorporate these concepts. This is explored in more detail below.

Ways of working with local and national actors is perhaps less defined. Following years of attempts to reform the humanitarian sector, the more recent discussion about decolonising aid is pushing the sector to consider ways of working that are even more transformative. The criteria of localisation, complementarity and partnership added by the IAHE (OCHA 2018) and in the SOHS (ALNAP 2018a) could be reviewed to find a concept or single criterion that best captures the intent of the sector. Any additional criteria or cross-cutting themes on this topic would benefit from substantive consultation.

The most-commonly applied cross-cutting theme found was a version of **gender, diversity, equity, equality and inclusion**, assessed in 80% of the evaluations reviewed. Gender and equity are integrated across all criteria in the updated OECD DAC guidance (2021). It will be essential for future humanitarian guidance to incorporate gender, diversity, equity, equality or inclusion in some form. If existing practice is followed, it will be included as a cross-cutting theme, as it is rarely included as an additional criterion in guidance or practice. See below for further exploration of this topic.

Protection was the third most-common cross-cutting theme in evaluations, assessed in 30% of evaluations reviewed. Future guidance will need to consider how to best approach protection, particularly given the IASC statement that ‘all humanitarian actors have a responsibility to place protection at the centre of humanitarian action’ (2013). It is substantively explored in ALNAP’s 2018 *Evaluation of protection in humanitarian action* (ALNAP 2018b). This highlights the importance of evaluating protection, while acknowledging the complexity of doing so. Usefully for future guidance, it provides protection considerations for each of the OECD DAC criteria for humanitarian action to assist when designing questions for protection evaluations. It emphasises the importance of starting with an understanding of the types of protection included in the intervention, and how the agency is using the concept of protection. Where protection is a more implicit focus of an intervention, evaluators may need to ‘tease out the protective features that can be inferred’ (ALNAP 2018b: 17).

Although less commonly applied in evaluations, **adaptive management** is emerging as a priority for some evaluators and donors. It has been applied as a cross-cutting theme and also under existing criteria. It is explored further below.

Approach

In terms future guidance, the options that arise in relation to additional or cross-cutting criteria are as follows:

- To consider one or two additional criteria. Noting, however, that there is already an additional criterion for evaluating humanitarian action (i.e., a seventh criterion as compared to the six standard 2019 OECD DAC criteria). The most likely candidates for additional criteria would be coordination and/or a variation on complementarity and localisation.

- To recommend cross-cutting themes, as ALNAP did in 2006 when it recommended eight cross-cutting themes. The current trend appears to skew to fewer, rather than more, cross-cutting themes. A limited list of recommended, and optional, cross-cutting themes could help to focus future guidance if more were considered necessary. The most prominent themes to consider are gender, diversity, equity, equality and inclusion, protection, and adaptive management.
- To follow the OECD DAC's approach and integrate two themes across the guidance, without labelling them as 'cross-cutting': 1) equity and gender equality and 2) the 'leave no one behind' imperative.

Regardless of the approach taken, it would be useful for future guidance to indicate how an evaluation could address the most commonly occurring themes – if not as additional criteria or as a cross-cutting theme, then under other criteria. For example, MSF (2017) recommend that inclusion should be addressed as part of coverage. This would facilitate greater consistency and therefore comparison of common themes across evaluations.

Gender, diversity and inclusion, equity and equality

The most-applied cross-cutting theme or consideration is the intent to ensure at-risk and marginalised groups receive the assistance they need. It is also the primary focus of the OECD DAC's 2021 guidance, which applies a gender lens and explains how to integrate inclusion across each of the evaluation criteria.

ALNAP includes gender equality in its 2006 guidance and adds consideration of vulnerability and marginalised groups in its 2016 guidance. The IAHE also includes gender and inclusiveness in the IAHE guidance (OCHA 2018) and addresses equity and inclusion in its definition of coordination, which it includes as an additional criterion. Indications of the prominence of cross-cutting themes related to gender and inclusion include UNICEF's guide for equity-focused evaluations (Bamberger and Segone 2011) and recent thematic evaluations, including the IAHE on gender equality and the empowerment of women and girls (IASC 2020) and ICRC's evaluation of its diversity, inclusion and accountability to affected populations (2018).

Of the evaluations analysed for this report, 80% considered gender, equity or inclusion in their approach, with 50% of these including a combination of the three. While inclusion was the most frequently cited, it often incorporated gender and/or equity, as well as disability status and age. Inclusion, gender and equity were primarily considered as a cross-cutting theme. Where they were considered under individual OECD DAC evaluation criteria, they were most commonly considered under effectiveness (seven times) and relevance (five times).

- 30 out of 40 evaluations considered **inclusion**: 13 as a cross-cutting theme, 10 under OECD DAC criteria, twice as a separate criterion and five times in another form.
- 50% of the evaluations included an assessment of **gender**-related issues: 14 as a cross-cutting theme, four under the OECD DAC criteria and once as a separate criterion (with inclusion), and once in another form.
- Three of the 40 evaluations analysed included the assessment of **equity**: once as a cross-cutting theme (with gender and inclusion) and once under the OECD criterion of effectiveness. A third evaluation included equity as an indicator of value for money.

Both terminology and definitions vary widely. Guidance and evaluations refer to gender, diversity, inclusion, equity, equality, 'leave no one behind' and more. The Humanitarian Policy Group (HPG)

has focused on inclusion (Lough et al. 2022); in the absence of a common definition, the HPG has provided a working definition of 'inclusion': ensuring equitable access, meeting diverse needs, enabling participation and focusing on the most urgent cases (impartiality) (ibid.). It found that understandings of inclusion within the humanitarian sector are heavily fragmented, that different aspects of inclusion are made discrete from one another and rarely considered within an overarching approach to understanding who is falling through the cracks and why.

In practice, humanitarian action does not reach all sections of society in need. IFRC's 2018 *World disasters report* found that millions of people are left behind in humanitarian crises, with many people unintentionally excluded by humanitarian organisations (Fisher et.al 2018). HPG similarly found that the humanitarian system is lacking in its ability to deliver inclusive humanitarian action. One issue, highlighted by both the HPG and ALNAP (Lough et al. 2022 and ALNAP 2020) is the tendency to address 'vulnerable' groups individually and programmatically, missing systemic issues and important sections of society. The HPG cites, among other causes for poor performance on inclusion, a lack of high-level commitment from leadership and a lack of tools to assess, track and evaluate progress on inclusion (Lough et al. 2022).

Evaluation criteria has the potential to support improved practice and, by many accounts, needs to better support inclusion. Peersman, for example, found gender equality and equity dimensions to be among common weaknesses relating to addressing evaluative criteria (2014). Indeed, an Equitable Evaluation Initiative has been established to advance equity in the purpose and practice of evaluation (www.equitableeval.org/about).

Despite incorporating a gender lens and integrating inclusion across criteria, the OECD DAC's 2019 update has been criticised for not adding a criterion focused on equity. Bitar (2021: 4) argues that the way in which the OECD DAC criteria address equity 'does not allow for systematic and intersectional consideration of these issues'. He proposes a social equity assessment tool (SEAT) to assist evaluators with the assessment of equity (2021). Patton (2020), in the context of evaluating transformation, has also argued for a diversity/ equity/inclusion criterion. He argues that, together with sustainability, equity is a foundation for transformation. Updated humanitarian guidance could choose to take a more defined approach and specifically incorporate gender, equity and/or inclusion as a cross-cutting theme.

Equitable evaluation encourages evaluators to consider four aspects in their evaluation practice, all at once: diversity of evaluation teams (beyond ethnic and cultural), cultural appropriateness and validity of evaluation methods, ability of evaluation designs to reveal structural and systems-level drivers of inequity, and the degree to which those affected by what is being evaluated have the power to shape and own how evaluation happens. Patton (2020: 73)

Accountability to affected populations, participation and communication with communities

The second most-applied cross-cutting theme or consideration in the evaluations analysed is that of accountability to affected populations, alongside participation and communication with communities. These are considered together as closely related concepts. The ODI's Humanitarian Policy Group has also noted their importance in supporting inclusive humanitarian action – particularly participation and local leadership (including the decision-making power of local and national actors) (Lough et al. 2022). Further, and as noted above, participation is a key component of the HPG's working definition of inclusion.

ALNAP included participation as a cross-cutting theme in its 2006 guide and made accountability and participation an additional criterion in the 2018 SOHS. Accountability to affected populations is an additional consideration in the IAHE guidelines and an additional criterion in the CHS.

Together, accountability to affected populations, participation and communication with communities were applied as cross-cutting themes 19 times in the 40 evaluations analysed, with accountability to affected populations being the most common. The topics were also considered under existing OECD DAC criteria 15 times, most often under relevance/appropriateness (seven times), followed by effectiveness (five times) and coherence (three times):

- 21 of 40 evaluations included **accountability to affected populations**: nine times as a cross-cutting theme, six times under the existing OECD DAC criteria, once as its own criterion, and five times in another form.
- 11 of 40 evaluations considered **communication with communities**: three times as a cross-cutting theme, six times under the existing OECD DAC criteria, and twice under CHS commitments.
- 10 of the 40 evaluations reviewed for this report included the assessment of **participation**: seven times as a cross-cutting theme and three times under existing OECD DAC criteria.

While definitions vary, ALNAP defined accountability and participation in the context of evaluation criteria for the 2018 SOHS as: ‘the degree to which actors within the international humanitarian system can be held to account by crisis-affected people, and the degree to which crisis-affected people are able to influence decisions related to assistance and protection’ (ALNAP 2018a: 35). It justified the inclusion of accountability to affected populations as a separate criterion by finding ‘it is not possible to say the system has performed satisfactorily unless aid is provided in a way that is accountable to those who receive it and allows them some measure of influence in decisions over the aid they receive’ (34).

In the 2018 SOHS, ALNAP found that, while there was significant activity related to accountability and participation, this seldom resulted in major changes. An independent review of the Grand Bargain (Metcalf-Hough et al. 2022: 15) also found that, despite efforts to elevate and improve participation (also referred to as accountability to affected populations), there has not yet been ‘any substantive impact’. The humanitarian system has also been criticised for its consistently poor performance in ensuring the participation of affected populations in evaluation processes (Rossignoli et al. 2017). The importance of not only evaluating the accountability of humanitarian action to affected populations but being accountable for the ways in which the evaluation is conducted is emphasised by many (see, for example, HAG 2021).

Future guidance could choose to strengthen the trend of evaluating accountability to affected populations and consider including it as a cross-cutting theme.

Adaptive management

One cross-cutting theme that has had attracted attention more recently is adaptive management, viewed as being central to improved practice. The global scale of the COVID-19 pandemic, and its dramatic impact on ways of operating, lent an urgency to learning, adaptive management and innovation (Buchanan-Smith and Morrison-Métois 2021). Many Real-Time Learning exercises in the early response to the pandemic focused on adaptive management, with UNICEF, the Danish Refugee Council and the Norwegian Refugee Council (NRC) all conducting adaptive management reviews (Buchanan-Smith and Morrison-Métois 2021).

It has also been a priority for donors, with the United Kingdom's Department for International Development and the United States's Agency for International Development establishing the Global Learning for Adaptive Management (GLAM) programme. GLAM's goal is to strengthen evidence-informed adaptive management through enhanced monitoring, evaluation and learning (Hernandez et al. 2019). GLAM has further introduced the concept of adaptive rigour, aiming to make adaptive decisions and practices more evidence-based, improving openness and accountability (Ramalingam et al. 2019). Both GLAM and ALNAP, among others, have published a number of resources supporting adaptive management in evaluation.

The updated OECD DAC guidance for evaluation criteria (2021) incorporates 'adapting over time' under the relevance criterion. It recommends an adaptive management analysis to inform assessments of relevance over time (OECD 2021: 38).

Some prominent humanitarian evaluations have assessed adaptive management. In the *Evaluation of diversity, inclusion and accountability to affected populations in ICRC operations* (Meier et al. 2018), the evaluation considered ICRC's ability to adapt its strategic orientation, as well as its practice of adapting the details of its programmes. The *Evaluation of OCHA's country-based pooled fund* (OCHA 2019) considered the fund's flexibility to adapt to local contexts and to accommodate changes in humanitarian priorities and programme approaches, although it was not mentioned specifically in the evaluation framework.

It has also been apparent in evaluations led by national consultants. For example, Low et al. (2019) assessed the ability of IFRC's Building Resilient Communities in Fiji project to adapt to the changing context under 'relevance'. The final evaluation of World Vision's cyclone response programme (World Vision 2020) in Zimbabwe assessed how well the programme adapted to changing context under CHS Commitment 7: Communities and people affected by crisis can expect delivery of improved assistance as organisations learn from experience and reflection.

Given that the approach to adaptive management is still emerging, it would seem most appropriate that future guidance follow the OECD DAC's approach to incorporating adaptation under the relevance criterion. Alternative approaches could be considered through consultation.

Summary and questions to explore

The research identified at least 30 additional and cross-cutting themes in a sample of 40 humanitarian evaluations. ALNAP itself listed eight cross-cutting themes in its 2006 guide and three additional criteria in its 2018 and 2022 SOHS reports. The OECD DAC (2021) has taken the approach of not explicitly including cross-cutting themes. It does, however, explain how to incorporate inclusion and equity across each criterion and encourages the application of a gender lens to evaluations. It also explains how to consider the 2030 Agenda and Sustainable Development Goals in evaluations.

There are some apparent trends that future guidance may wish to consider incorporating as additional criteria or cross-cutting themes: ways of working with national and local actors (including coordination), ways of working with communities (such as accountability to affected populations) and diversity, equity and inclusion. Protection, as the third most-applied cross-cutting theme, will also need to be addressed, given the centrality of protection and the challenges in its evaluation (ALNAP 2018b). Adaptive management, as an emerging criterion, could also be explored.

- **Should guidance specify cross-cutting themes, similar to the approach taken by ALNAP in 2006, or explain how to incorporate selected themes (not explicitly labelled as ‘cross-cutting’), similar to the approach taken by the OECD DAC (2021)?**

Future guidance will need to be explicit about how to best address coordination. There is merit in reflecting current guidance and practice, which both emphasises coordination as a critical component of effectiveness and, alternately, includes it as an additional criterion. A preference for fewer, rather than more, criteria may be the determining factor.

- **Should coordination be included as an additional criterion, or incorporated under other criteria?**

Coordination is not sufficient on its own to reflect the trends towards working in ways that support and strengthen local and national actors, accountability and partnership. This could be explored through ways of working with local and national actors, and ways of working with communities.

Gender, diversity and inclusion, equity and equality was the most-applied cross-cutting theme or consideration. Of the evaluations analysed, 80% considered gender, equity or inclusion in their approach. While inclusion was the most frequently cited, it often incorporated gender and/or equity, as well as disability status and age. Where these were considered under individual OECD DAC evaluation criteria, it was most commonly under effectiveness (seven times) and relevance (five times).

Terminology and definitions vary widely. Guidance and evaluations refer to gender, diversity, inclusion, equity, equality, leaving no one behind and more. The HPG has focused on inclusion, finding that the humanitarian system is lacking in its ability to deliver inclusive action (Lough et al. 2022). One issue, highlighted by both HPG and ALNAP (Lough et al. 2022 and ALNAP 2020) is the tendency to address ‘vulnerable’ groups individually and programmatically, thereby missing systemic issues and important sections of society. The HPG cites, among other causes for poor performance on inclusion, a lack of high-level commitment from leadership and a lack of tools to assess, track and evaluate progress in this area (Lough et al. 2022).

Updated humanitarian guidance could choose to address this poor performance on inclusion by incorporating gender, equity and/or inclusion as a cross-cutting theme.

Accountability to affected populations, participation and communication with communities were collectively the second most-applied considerations in the evaluations analysed. Considered together as closely related concepts, together they were applied as cross-cutting themes 19 times in the 40 evaluations analysed, with accountability to affected populations being the most common.

An independent review of the Grand Bargain (Metcalf-Hough et al. 2022: 15) found that, despite efforts to elevate and improve participation (also referred to as accountability to affected populations), there has not yet been ‘any substantive impact’. ALNAP justified the inclusion of accountability to affected populations as a separate criterion by finding ‘it is not possible to say the system has performed satisfactorily unless aid is provided in a way that is accountable to those who receive it and allows them some measure of influence in decisions over the aid they receive’ (2018a: 34).

Future guidance could choose to strengthen the current trend towards evaluating accountability to affected populations and consider including it as a cross-cutting theme.

Protection was the third most-common cross-cutting theme in evaluations, assessed in 30% of evaluations reviewed. Future guidance will need to consider how to best approach protection, particularly given the IASC statement that 'all humanitarian actors have a responsibility to place protection at the centre of humanitarian action' (2013). It is substantively explored in ALNAP's (2018b) *Evaluation of protection in humanitarian action*. ALNAP (2018b) highlights the importance of evaluating protection, while acknowledging the complexity of doing so. Usefully for future guidance, it provides protection considerations for each of the OECD DAC criteria for humanitarian action to assist when designing questions for protection evaluations.

Adaptive management has attracted attention more recently. The global scale of the COVID-19 pandemic, and its dramatic impact on ways of operating, lent an urgency to learning, adaptive management and innovation (Buchanan-Smith and Morrison-Métois 2021). Many Real-Time Learning exercises in the early response to the pandemic focused on adaptive management, as have some prominent humanitarian evaluations. It has also been a priority for some donors. Evaluating adaptive management challenges organisations to look at the way in which their organisational processes support or hinder the ability to learn, adapt and improve delivery. Applying an adaptive management lens may also assist evaluators to test whether the objectives being assessed are the most relevant to the changing context over time.

Given that the approach to adaptive management is still emerging, it would seem most appropriate that future guidance follow the OECD DAC's approach to incorporating adaptation under the relevance criterion. Alternative approaches could be considered through consultation.

Ways of working with local and national actors is perhaps less defined. Following years of attempts to reform the humanitarian sector, the more recent discussion about decolonising aid is pushing the sector to consider ways of working that are even more transformative. The criteria of localisation, complementarity and partnership added by the IAHE (OCHA 2018) and in the SOHS (ALNAP 2018a) could be reviewed to find a concept or single criterion that best captures the intent of the sector. Any additional criteria or cross-cutting themes on this topic would benefit from substantive consultation.

- **If cross-cutting themes are specified, should they be (a) gender, equity and inclusion, (b) accountability to affected populations or (c) protection? If not, should these be addressed as additional criteria or integrated across existing criteria?**
- **Is it important to more explicitly incorporate ways of working with local and national actors (such as localisation or complementarity) into guidance for evaluating humanitarian action? If so, should this be as a cross-cutting theme, additional criteria or integrated across existing criteria?**
- **Is there an ideal number of cross-cutting themes or additional criteria? Should they be limited?**

CONCLUSION

The OECD DAC criteria remain the most used criteria for evaluating humanitarian action, despite other options. Guidance for applying the criteria supports the sector to compare and understand broader performance and trends. Any updates to the criteria and guidance must support improved use and help evaluators to measure what is most meaningful to improved humanitarian action.

The criteria have been defined and applied differently across organisations and over time. Some criteria are used much more often than others. The review of organisational evaluation guidance and recent humanitarian evaluations demonstrated that organisations and evaluators have chosen to interpret and adapt the criteria differently in practice from what is outlined in ALNAP's 2006 guide. This reflects contestations of individual criteria and the broader set of OECD DAC criteria in the literature. ALNAP's own application of the criteria has evolved, as demonstrated by their separating out the use of complementarity, sufficiency, and accountability and participation in assessing the SOHS (ALNAP 2018a).

There have been various proposals over time to include additional criteria and cross-cutting themes to the OECD DAC criteria and guidance. Many organisations and evaluations are already using a wide range of additional criteria and cross-cutting themes, in line with their individual priorities. This review has highlighted some of the more common: gender equity, diversity and inclusion; accountability to affected populations and participation; coordination; protection; ways of working with local and national actors; and the newly emerging adaptive management.

There are advantages to the flexible interpretation and application of the criteria, allowing evaluators to adapt to the objectives of the evaluation, the context and the intervention being assessed. Flexibility may also better support calls to decolonise evaluation approaches and practices. To more substantively contribute to decolonisation, guidance could also support users to consider the positionality of evaluators, whose world view inevitably determines their interpretation of the criteria, evaluation methodologies and findings.

The endurance of the evaluation criteria, their extensive application and the utility of having a common basis or framework for evaluating and understanding performance in the humanitarian sector, are all arguments in favour of ensuring there is adequate guidance for practitioners on using the OECD DAC criteria in humanitarian settings. The contemporary critiques; changes in concepts, topical issues and language over time; and the varied application of the criteria in practice suggest that updated and additional guidance would be helpful. The OECD DAC's own update to the criteria (2019) and guidance (2021) provide an anchor for ALNAP to update its guidance for humanitarian evaluators and a foundation upon which to build.

Your views matter!

This paper is meant to inform further exchanges and discussions in the humanitarian evaluation community, as ALNAP seeks to update its existing guidance.

To share your views, visit our [website](#) to find more information about participating in a survey of evaluation practitioners and upcoming consultation events.

Want to get in touch?

You can contact ALNAP's Senior Research Fellow, Susanna Morrison-Métois at: evalcriteria@alnapp.com

BIBLIOGRAPHY

- Abdelmagid, N, Checchi, F, Garry, S. et al. (2019) 'Defining, measuring and interpreting the appropriateness of humanitarian assistance'. International Journal of Humanitarian Action, 4(14). (<https://www.alnap.org/defining-measuring-and-interpreting-the-appropriateness-of-humanitarian-assistance>).
- ACT Alliance. (2012) ACT evaluation policy and guidelines. Geneva: ACT Alliance. (www.alnap.org/help-library/act-evaluation-policy-and-guidelines).
- ALNAP. (2006) Evaluating humanitarian action using the OECD DAC criteria: An ALNAP guide for humanitarian agencies. ALNAP Guide. London: ALNAP/ODI. (www.alnap.org/help-library/evaluating-humanitarian-action-using-the-oecd-dac-criteria).
- ALNAP. (2014) Evaluating protection in humanitarian action: Issues and challenges. London: ALNAP/ODI. (www.alnap.org/help-library/evaluating-protection-in-humanitarian-action-issues-and-challenges).
- ALNAP. (2015) The State of the Humanitarian System. ALNAP Report. London: ALNAP/ODI. (<https://www.alnap.org/the-state-of-the-humanitarian-system-report-2015>).
- ALNAP. (2016) Evaluation of humanitarian action guide. ALNAP Guide. London: ALNAP/ODI. (www.alnap.org/evaluation-of-humanitarian-action-eha-guide).
- ALNAP. (2018a) The State of the Humanitarian System. ALNAP Report. London: ALNAP/ODI. (www.alnap.org/help-library/the-state-of-the-humanitarian-system-sohs-2018-full-report).
- ALNAP. (2018b) Evaluation of Protection in Humanitarian Action. ALNAP Guide. London: ALNAP/ODI. (www.alnap.org/help-library/alnap-guide-evaluation-of-protection-in-humanitarian-action).
- ALNAP. (2020) ALNAP 32nd Annual Meeting study: More relevant? 10 ways to approach what people really need. London: ALNAP/ODI. (www.alnap.org/help-library/alnap-32nd-annual-meeting-study-more-relevant-10-ways-to-approach-what-people-really).
- ALNAP. (2022) The State of the Humanitarian System. ALNAP Report. London: ALNAP/ODI. (<https://www.alnap.org/2022-the-state-of-the-humanitarian-system-sohs-%E2%80%93-full-report>).
- Bakrania, S., Balvin, N., Daidone, S. and de Hoop, J. (2021) Impact evaluation in settings of fragility and humanitarian emergency. Florence: UNICEF. (www.alnap.org/impact-evaluation-in-settings-of-fragility-and-humanitarian-emergency).
- Bamberger, M., and Segone, M. (2011) How to design and manage equity-focused evaluations. York: UNICEF. (www.alnap.org/help-library/how-to-design-and-manage-equity-focused-evaluations).
- Barbelet, V., Davies, G., Flint, J. and Davey, E. (2021) Interrogating the evidence base on humanitarian localisation: a literature study. London: ODI. (<https://www.alnap.org/interrogating-the-evidence-base-on-humanitarian-localisation-a-literature-study>).

Barber, R. (2016) Did the humanitarian response to the Nepal earthquake ensure no one was left behind? A case study on the experience of marginalised groups in humanitarian action. London: Save the Children International. (<https://www.alnap.org/did-the-humanitarian-response-to-the-nepal-earthquake-ensure-no-one-was-left-behind-a-case-study-on>).

BHA. (2022) Bureau for Humanitarian Assistance technical guidance for monitoring, evaluation, and reporting for emergency activities. Washington D.C.: USAID Bureau for Humanitarian Assistance. (www.alnap.org/bureau-for-humanitarian-assistance-technical-guidance-for-monitoring-evaluation-and-reporting-for).

Bitar, K. (2021) A social equity assessment tool (SEAT) for evaluation. Erfurt: Willy Brandt School of Public Policy, University of Erfurt. (www.alnap.org/a-social-equity-assessment-tool-seat-for-evaluation).

Bizzarri, M., Cano Vinas, M., Khogali, H. and Jantzi, T. Inter-agency humanitarian evaluation on gender equality and the empowerment of women and girls. New York: IASC. (<https://www.alnap.org/inter-agency-humanitarian-evaluation-on-gender-equality-and-the-empowerment-of-women-and-girls>).

Brusset, E., Cosgrave, J. and MacDonald, W. (2010) 'Real-Time Evaluation in humanitarian emergencies'. New Directions for Evaluation, 126: 9-20. (www.alnap.org/real-time-evaluation-in-humanitarian-emergencies).

Buchanan-Smith, M. (2022) Margie Buchanan-Smith on the importance of being guided by humanitarian principles and sector standards. ALNAP Blog. London: ALNAP/ODI. (www.alnap.org/blogs/margie-buchanan-smith-on-the-importance-of-being-guided-by-humanitarian-principles-and-sector).

Buchanan-Smith, M. and Morrison-Métois, S. (2021) From Real-Time Evaluation to Real-Time Learning: exploring new approaches from the COVID-19 response. ALNAP paper. London: ALNAP/ODI. (www.alnap.org/help-library/from-real-time-evaluation-to-real-time-learning).

Chilisa, B. and Mertens, D. (2021) 'Indigenous Made in Africa evaluation frameworks: Addressing epistemic violence and contributing to social transformation'. American Journal of Evaluation, 42(2): 241-253. (www.alnap.org/indigenous-made-in-africa-evaluation-frameworks-addressing-epistemic-violence-and-contributing-to).

CHS. (2014) Core humanitarian standard on quality and accountability. Geneva: Core Humanitarian Standard. (<https://www.alnap.org/core-humanitarian-standard-on-quality-and-accountability>).

CHS. (2018) Core Humanitarian Standard on quality and accountability: Updated guidance notes and indicators 2018. Geneva: CHS. (www.alnap.org/core-humanitarian-standard-on-quality-and-accountability-updated-guidance-notes-and-indicators-2018).

CHS. (2020) Verification Framework: Core Humanitarian Standard on quality and accountability. Geneva: CHS. (www.alnap.org/verification-framework-core-humanitarian-standard-on-quality-and-accountability).

Clements, P. (2020) 'Improving learning and accountability in foreign aid'. World Development, 125. (www.alnap.org/improving-learning-and-accountability-in-foreign-aid).

- Coghlan, D. and Brydon-Miller, M. (eds) (2014) *The SAGE Encyclopaedia of Action Research*. Thousand Oaks: SAGE Publications Ltd. (<https://www.alnap.org/the-sage-encyclopedia-of-action-research>).
- Cosgrave, J., Ramalingam, B., and Beck, T. (2009) *Real Time Evaluations of humanitarian action – an ALNAP guide*. ALNAP Guide. London: ALNAP/ODI. (<https://www.alnap.org/real-time-evaluations-of-humanitarian-action-an-alnap-guide>).
- Darcy, J. and Dillon, N. (2020) *Missing the point? Reflections on current practice in evaluating humanitarian action*. London: ALNAP/ODI. (<https://www.alnap.org/missing-the-point-reflections-on-current-practice-in-evaluating-humanitarian-action>).
- De Coning, C. and Friis, K. (2011) 'Coherence and coordination: The limits of the comprehensive approach'. *Journal of International Peacekeeping*, 15: 243-272. (www.alnap.org/coherence-and-coordination-the-limits-of-the-comprehensive-approach).
- DEval. (2018) *DEval Policy Brief 9/2018: The OECD DAC evaluation criteria: To reform or to transform?* Bonn: DEval. (www.alnap.org/deval-policy-brief-92018-the-oecd-dac-evaluation-criteria-to-reform-or-to-transform).
- DFID. (2020) *DFID's Approach to Value for Money – Guidance for external partners*. London: DFID. (www.alnap.org/dfid%E2%80%99s-approach-to-value-for-money-%E2%80%93-guidance-for-external-partners).
- Drew, R. (2021) *COVID-19 Global Evaluation Coalition, evaluating the coherence of the international response to the COVID-19 pandemic scoping study for the COVID-19 Global Evaluation Coalition*. Paris: OECD. (www.alnap.org/covid-19-global-evaluation-coalition-evaluating-the-coherence-of-the-international-response-to-the).
- Dunsch, F., Lombardini, S. and Heirman, J. (2021) 'Lean impact evaluations: Experimental evidence in adaptive humanitarian interventions'. [Blog]. 12 July. Rome: WFP. (<https://www.alnap.org/%E2%80%9Clean%E2%80%9D-impact-evaluations-experimental-evidence-in-adaptive-humanitarian-interventions>).
- Elice, P. (2021) *Impact evaluations in forced displacement contexts: A guide for practitioners*. (www.alnap.org/impact-evaluations-in-forced-displacement-contexts-a-guide-for-practitioners).
- FAO. (2019) *Evaluation of the FAO Strategy for Partnerships with the Private Sector*. Rome: FAO. (<https://www.alnap.org/evaluation-of-the-fao-strategy-for-partnerships-with-the-private-sector>).
- Featherstone, A., Mowjee, T., Fleming, D., Tong, K., Gros, C. and Evans-Gutierrez, L. (2019) *Evaluation of the coverage and quality of the UNICEF humanitarian response in complex humanitarian emergencies*. New York: UNICEF. (<https://www.alnap.org/evaluation-of-the-coverage-and-quality-of-the-unicef-humanitarian-response-in-complex>).
- Featherstone, A., Mowjee, T., Lattimer, C. and Poole, L. (2019) *OCHA evaluation of country-based pooled funds: Global synthesis report*. New York: OCHA. (www.alnap.org/ocha-evaluation-of-country-based-pooled-funds-global-synthesis-report).
- Field, J. (2016) 'A culture clash? Exploring "appropriateness" and "relevance" in the response to Typhoon Haiyan/Yolanda, in *Essays on Humanitarian Effectiveness*. London: Save the Children. (<https://www.alnap.org/essays-on-humanitarian-effectiveness>).

Fisher, D., Hagon, K., Lattimer, C., O'Callaghan, S., Swithern, S. and Walmsley, L. (2018) World Disasters Report: Leaving no one behind. Geneva: IFRC. (www.alnap.org/world-disasters-report-2018-leaving-no-one-behind).

GLOW Consultants and Humanitarian Advisory Group. (2019) Final evaluation report: Responding to Pakistan's internally displaced (RAPID) fund (RAPID FUND-II). Washington D.C.: USAID. (www.alnap.org/final-evaluation-report-responding-to-pakistan%E2%80%99s-internally-displaced-rapid-fund-rapid-fund-ii).

HAG. (2021) Accountability to affected people: Stuck in the weeds. Melbourne: HAG. (www.alnap.org/accountability-to-affected-people-stuck-in-the-weeds).

Hagens C., Morel, D., Causton A. and Way, C. (2012) Guidance on monitoring and evaluation. Baltimore: CRS. (www.alnap.org/help-library/guidance-on-monitoring-and-evaluation).

Hallam, A. (1998) Evaluating humanitarian assistance programmes in complex emergencies. London: ODI. (<https://www.alnap.org/help-library/guidance-for-evaluating-humanitarian-assistance-in-complex-emergencies-0>).

Heider, C. (2017) Rethinking evaluation: Reflections from Caroline Heider. Washington D.C.: World Bank Group. (www.alnap.org/rethinking-evaluation-reflections-from-caroline-heider).

Hernandez, K., Ramalingham, B., and Wild, L. (2019) Towards evidence-informed adaptive management: A roadmap for development and humanitarian organisations. London: ODI. (<https://www.alnap.org/towards-evidence-informed-adaptive-management-a-roadmap-for-development-and-humanitarian>).

IASC. (2013) The centrality of protection in humanitarian action. Geneva: IASC. (<http://alnap.org/the-centrality-of-protection-in-humanitarian-action-statement-by-the-inter-agency-standing-committee>).

IASC. (2019) Inter-Agency Humanitarian Evaluation of the drought response in Ethiopia 2015 – 2018. Geneva: IASC. (<https://www.alnap.org/inter-agency-humanitarian-evaluation-of-the-drought-response-in-ethiopia-2015-2018>).

IASC. (2020) Inter-Agency Humanitarian Evaluation of the response to Cyclone Idai in Mozambique. New York: IASC. (<https://www.alnap.org/inter-agency-humanitarian-evaluation-of-the-response-to-cyclone-idai-in-mozambique>).

IASC. (2021) Inter-Agency Humanitarian Evaluation of the COVID-19 humanitarian response. New York: IASC. (<https://www.alnap.org/inter-agency-humanitarian-evaluation-of-the-covid-19-humanitarian-response>).

IASC. (2022) Inter-Agency Humanitarian Evaluation (IAHE) of the Yemen crisis. New York: IASC. (<https://www.alnap.org/inter-agency-humanitarian-evaluation-iahe-of-the-yemen-crisis>).

IASC. (n.d.) 'Inter-Agency Humanitarian Evaluations'. [Webpage]. New York: IASC. (<https://interagencystandingcommittee.org/inter-agency-humanitarian-evaluations>).

ICRC. (2016) World Humanitarian Summit: ICRC's comment on the UN Secretary General's report. Geneva: ICRC. (www.alnap.org/world-humanitarian-summit-icrcs-comment-on-the-un-secretary-generals-report).

- ICRC. (2022) ICRC Evaluation Strategy 2022-2024. Geneva: ICRC. (www.alnap.org/icrc-evaluation-strategy-2022-2024).
- IFRC and ICRC. (1994) The code of conduct for the International Red Cross and Red Crescent movement and non-governmental organisations (NGOs) in disaster relief. Geneva: IFRC. (www.alnap.org/help-library/code-of-conduct-for-the-international-red-cross-and-red-crescent-movement-and-ngos-in).
- IFRC. (2011) IFRC framework for evaluation. Geneva: IFRC. (<https://www.alnap.org/ifrc-framework-for-evaluation>).
- IFRC. (2019) Everyone counts. Geneva: IFRC. (<https://www.alnap.org/everyone-counts-may-2019-issue>).
- IOM. (2020) IOM Monitoring and Evaluation Guidelines. IOM Guide. Geneva: IOM (<https://www.alnap.org/iom-monitoring-and-evaluation-guidelines>).
- Kennedy-Chouane, M. G. (2013) 'Developing OECD DAC guidance on evaluating in settings of violent conflict and fragility'. *Journal of Peacebuilding & Development*, 8(2): 110-115. (www.alnap.org/developing-oecd-dac-guidance-on-evaluating-in-settings-of-violent-conflict-and-fragility).
- Kennedy-Chouane, M. G. (2020) 'Evaluation criteria: What has changed, why, and why does it matter?' [Webinar]. January 16. (www.alnap.org/evaluation-criteria-what-has-changed-why-and-why-does-it-matter).
- Lough, O., Barbelet, V. and Njeri, S. (2022) Inclusion and exclusion in humanitarian action: Findings from a three year study. London: HPG/ODI. (<https://www.alnap.org/inclusion-and-exclusion-in-humanitarian-action-findings-from-a-three-year-study>).
- Low, I., Robinson, L. and Hesaie, D. (2019) Final evaluation: Building resilient communities in Fiji project. (<https://www.alnap.org/final-evaluation-building-resilient-communities-in-fiji-project>).
- Lundgren, H. (2017) DAC evaluation criteria: A discussion and issues paper for the 21st meeting of the DAC Network on Development Evaluation. Paris: OECD. (www.alnap.org/dac-evaluation-criteria-a-discussion-and-issues-paper-for-the-21st-meeting-of-the-dac-network-on).
- Jacimovic, R. and Bostoen, K. (2017) Monitoring and evaluation framework for WASH market-based humanitarian programming guidance document. Oxford: Oxfam. (<https://www.alnap.org/monitoring-and-evaluation-framework-for-wash-market-based-humanitarian-programming-guidance-document>).
- Meier, C., Steets, J., Derzsi-Horvath, A., Minh Bui, M. and Reichhold, U. (2018). Evaluation of diversity, inclusion, and AAP in ICRC operations: Evaluation report. Berlin : GPPi. (<https://www.alnap.org/evaluation-of-diversity-inclusion-and-aap-in-icrc-operations-evaluation-report>).
- Metcalfe-Hough, V., Fenton, W., Saez, P. and Spencer, A. (2022) The Grand Bargain in 2021: An independent review. London: HPG/ODI. (<https://www.alnap.org/the-grand-bargain-in-2021-an-independent-review>).
- Minea, L. (1994) 'The International Relief System: A Critical Review' in OECD DAC. (1999) Guidance for evaluating humanitarian assistance in complex emergencies. Paris: OECD. (www.alnap.org/guidance-for-evaluating-humanitarian-assistance-in-complex-emergencies).

- MOPAN. (2018) MOPAN 3.0 methodology manual. Paris: MOPAN. (<https://www.alnap.org/mopan-30-methodology-manual>).
- Mountfield, B. (2015) Sphere for monitoring and evaluation. Geneva: The Sphere Project (<https://www.alnap.org/sphere-for-monitoring-and-evaluation>).
- MSF. (2017). Evaluation manual: a handbook for initiating, managing and conducting evaluations in MSF. Vienna: MSF. (www.alnap.org/help-library/evaluation-manual-a-handbook-for-initiating-managing-and-conducting-evaluations-in-msf).
- Noltze, M., Euler, M. and Verspohl, I. (2018) Sustainability in German development cooperation: Meta-evaluation. Bonn: DEval. (<http://www.alnap.org/sustainability-in-german-development-cooperation-meta-evaluation>).
- Obrecht, A., Warner, A., and Dillon, N. (2017) Working paper: Evaluating humanitarian innovation. HIF/ALNAP Working Paper. London: ALNAP/ODI. (<https://www.alnap.org/evaluating-humanitarian-innovation-hif-alnap-working-paper>).
- Picciotto, R. (2013) 'The logic of development effectiveness: Is it time for the broader evaluation community to take notice?' Evaluation, 19(2): 155–170. (www.alnap.org/the-logic-of-development-effectiveness-is-it-time-for-the-broader-evaluation-community-to-take).
- OCHA. (2013) Operational peer review: Guidance terms of reference report template. New York: OCHA. (<https://aap-inclusion-psea.alnap.org/help-library/operational-peer-review-template>).
- OCHA. (2018) Inter-Agency Humanitarian Evaluations: Process guidelines. New York: OCHA. (www.alnap.org/inter-agency-humanitarian-evaluations-process-guidelines).
- OECD. (1991) Principles for evaluation of development assistance. Paris: OECD. (<https://www.alnap.org/help-library/principles-for-evaluation-of-development-assistance-0>).
- OECD. (1999) Guidance for evaluating humanitarian assistance in complex emergencies. Paris: OECD. (<https://www.alnap.org/guidance-for-evaluating-humanitarian-assistance-in-complex-emergencies>).
- OECD. (2007) Encouraging effective evaluation of conflict prevention and peacebuilding activities: Towards DAC guidance. Paris: OECD. (www.alnap.org/encouraging-effective-evaluation-of-conflict-prevention-and-peacebuilding-activities-towards-dac).
- OECD. (2010) Quality standards for development evaluation. Paris: OECD. (<https://www.alnap.org/quality-standards-for-development-evaluation>).
- OECD. (2012) Evaluating peacebuilding activities in settings of conflict and fragility: Improving learning for results. Paris: OECD. (<https://www.alnap.org/evaluating-peacebuilding-activities-in-settings-of-conflict-and-fragility-improving-learning-for>).
- OECD. (2018a) Summary of the workshop on OECD DAC evaluation criteria: Progressing the dialogue. Paris: OECD. (www.alnap.org/summary-of-the-workshop-on-oecd-dac-evaluation-criteria-progressing-the-dialogue).
- OECD. (2018b) OECD DAC evaluation criteria: Summary of consultation responses. Paris: OECD. (www.alnap.org/oecd-dac-evaluation-criteria-summary-of-consultation-responses).

- OECD. (2019) Better criteria for better evaluation revised evaluation criteria definitions and principles for use. Paris: OECD. (<https://www.alnap.org/better-criteria-for-better-evaluation-revised-evaluation-criteria-definitions-and-principles-for-use>).
- OECD. (2021) Applying evaluation criteria thoughtfully. Paris: OECD. (<https://www.alnap.org/applying-evaluation-criteria-thoughtfully>).
- Ofir, Z. (2017) 'The DAC criteria, Part 11. From criteria to design principles'. [Blog]. 14 November. (www.alnap.org/the-dac-criteria-part-11-from-criteria-to-design-principles).
- Palenberg, M. (2011) Evaluation working papers: Tools and methods for evaluating the efficiency of development interventions. Bonn: German Federal Ministry for Economic Cooperation and Development. (www.alnap.org/evaluation-working-papers-tools-and-methods-for-evaluating-the-efficiency-of-development).
- Patton, M. Q. (2020) 'Evaluation criteria for evaluating transformation: Implications for the coronavirus pandemic and the global climate emergency'. American Journal of Evaluation, 42(1): 53-89. (www.alnap.org/evaluation-criteria-for-evaluating-transformation-implications-for-the-coronavirus-pandemic-and-the).
- Peersman, G. (2014) Evaluative criteria: Methodological briefs impact evaluation no. 3. New York: UNICEF. (www.alnap.org/evaluative-criteria-methodological-briefs-impact-evaluation-no-3).
- Proudlock, K., Ramalingam, B. and Sandison, P. (2009) 'Improving humanitarian impact assessment: bridging theory and practice', in ALNAP's 8th Review of Humanitarian Action: Performance, Impact and Innovation. London: ALNAP/ODI. (<https://www.alnap.org/improving-humanitarian-impact-assessment-bridging-theory-and-practice-alnaps-8th-review-of>).
- Puri, J., Aladysheva, A., Iversen, V., Ghorpade, Y. and Brück, T. (2015) What methods may be used in impact evaluations of humanitarian assistance? Bonn: IZA. (<https://www.alnap.org/what-methods-may-be-used-in-impact-evaluations-of-humanitarian-assistance>).
- Puri, J., Aladysheva, A., Iversen, V., Ghorpade, Y. and Brück, T. (2017) 'Can rigorous impact evaluations improve humanitarian assistance?'. Journal of Development Effectiveness, 9(4): 519-542. (<https://www.alnap.org/can-rigorous-impact-evaluations-improve-humanitarian-assistance>).
- Ramalingam, B., Wild, L. and Buffardi, A. (2019) Making adaptive rigour work: Principles and practices for strengthening monitoring, evaluation and learning for adaptive management. London: ODI. (<https://www.alnap.org/making-adaptive-rigour-work-principles-and-practices-for-strengthening-monitoring-evaluation-and>).
- Rossignoli C. M., Giani, A., Di Iacovo, F., Moruzzo, R. and Scarpellini, P. (2017) 'Enhancing participatory evaluation in a humanitarian aid project'. Evaluation, 23(2): 134-151. (<https://www.alnap.org/enhancing-participatory-evaluation-in-a-humanitarian-aid-project>).
- Samad, Z. A. and Moschini, B. (2016) Humanitarian assistance in Lebanon overview, challenges, and recommendations. Beirut: Arab NGO Network for Development. (<https://www.alnap.org/humanitarian-assistance-in-lebanon-overview-challenges-and-recommendations>).
- Save the Children. (2012) Evaluation handbook. London: Save the Children International. (<https://www.alnap.org/evaluation-handbook>).

- Slim, H. (2020) You don't have to be neutral to be a good humanitarian. Geneva: The New Humanitarian. (<https://www.alnap.org/you-don%E2%80%99t-have-to-be-neutral-to-be-a-good-humanitarian>).
- Sphere. (2018) The Sphere handbook: Humanitarian charter and minimum standards in humanitarian response, fourth edition. Geneva: Sphere. (<https://www.alnap.org/the-sphere-handbook-humanitarian-charter-and-minimum-standards-in-humanitarian-response-fourth>).
- Stoddard, A. (2005) Review of UNICEF's evaluation of humanitarian action. New York: UNICEF Evaluation Office. (www.alnap.org/help-library/review-of-unicefs-evaluation-of-humanitarian-action).
- Tanner, L. and Moro, L. (2016) Missed Out: The role of local actors in the humanitarian response in the South Sudan conflict. (<https://www.alnap.org/missed-out-the-role-of-local-actors-in-the-humanitarian-response-in-the-south-sudan-conflict>).
- UNEG. (2016a) Norms and standards for evaluation. New York: UNEG. (<https://www.alnap.org/uneq-norms-and-standards-for-evaluation>).
- UNEG. (2016b) Reflecting humanitarian principles in evaluation. New York: UNEG. (www.alnap.org/help-library/reflecting-humanitarian-principles-in-evaluation).
- UNFPA. (2019a) Evaluation Handbook: How to design and conduct a country program evaluation at UNFPA. New York: UNFPA. (www.alnap.org/evaluation-handbook-how-to-design-and-conduct-a-country-program-evaluation-at-unfpa).
- UNFPA. (2019b) Evaluation of the UNFPA capacity in humanitarian action (2012-2019). New York: UNFPA. (<https://www.alnap.org/evaluation-of-the-unfpa-capacity-in-humanitarian-action-2012-2019>).
- UNICEF. (2018) Humanitarian Cash Transfers Monitoring and Evaluation Guidance. New York: UNICEF. (www.alnap.org/humanitarian-cash-transfers-monitoring-and-evaluation-guidance).
- UNICEF. (2022) Impact evaluation for resilience learning in South Sudan. New York : UNICEF. (www.alnap.org/impact-evaluation-for-resilience-learning-in-south-sudan).
- UNOIOS. (2014) Inspection and evaluation manual. New York: UNOIOS. (www.alnap.org/help-library/inspection-and-evaluation-manual).
- USAID. (2021) 'Draft technical guidance for monitoring evaluation and reporting for emergency activities'. (<https://www.alnap.org/bureau-for-humanitarian-assistance-technical-guidance-for-monitoring-evaluation-and-reporting-for-and-change-year-to-2022>).
- WFP. (2019) WFP impact evaluation strategy 2019 - 2026. Rome: WFP. (<https://www.alnap.org/wfp-impact-evaluation-strategy-2019-2026>).
- WFP. (2020) Evaluation of the WFP People Strategy (2014-2017): Evaluation Report: Volume 1. Rome: WFP. (www.alnap.org/help-library/evaluation-of-the-wfp-people-strategy-2014-2017).
- WFP. (2021a) Impact evaluations to optimize humanitarian interventions: Brief. Rome: WFP. (<https://www.alnap.org/impact-evaluations-to-optimize-humanitarian-interventions-brief>).
- WFP. (2021b) Technical note: Evaluation questions and criteria. Rome: WFP. (www.alnap.org/technical-note-evaluation-questions-and-criteria).

WFP. (2022) WFP Evaluation policy 2022. Rome: WFP. (www.alnap.org/wfp-evaluation-policy-2022).

World Vision. (2020) Final evaluation of the cyclone response program: Final report. Bulawayo: World Vision. (<https://www.alnap.org/final-evaluation-of-the-cyclone-response-program-final-report>).

ANNEX 1: LIST OF EVALUATIONS

Evaluation name		Year	Region or country
Inter-Agency Humanitarian Evaluation of the Drought Response in Ethiopia 2015 - 2018	https://www.wfp.org/publications/inter-agency-humanitarian-evaluation-drought-response-ethiopia-2015-2018	2019	Ethiopia
Norwegian Refugee Council Global Cash Evaluation 2019	https://www.alnap.org/help-library/nrc-global-cash-evaluation	2019	Ethiopia Somalia Iraq
Indonesia Earthquakes and Tsunami Response Final Evaluation	https://reliefweb.int/report/indonesia/indonesia-earthquakes-and-tsunami-response-mdrid013-final-evaluation-executive	2019	Indonesia
Inter-Agency Humanitarian Evaluation of the Yemen Crisis	https://interagencystandingcommittee.org/system/files/2022-07/Yemen%20IAHE%20Final%20Report%2C%2013%20July%202022%20%28English%29.pdf	2022	Yemen
Inter-Agency Humanitarian Evaluation on Gender Equality and the Empowerment of Women and Girls	https://interagencystandingcommittee.org/system/files/2021-01/The%20Inter-Agency%20Humanitarian%20Evaluation%20%28IAHE%29%20on%20Gender%20Equality%20and%20the%20Empowerment%20of%20Women%20and%20Girls%20%28GE-EWG%29-Report.pdf	2020	Global
Inter-Agency Humanitarian Evaluation of the Response to Cyclone Idai in Mozambique	https://www.wfp.org/publications/inter-agency-humanitarian-evaluation-response-cyclone-idai-mozambique	2020	Mozambique
OCHA Evaluation of Country Based Pooled Funds	https://www.unocha.org/sites/unocha/files/2019%20OCHA%20Evaluation%20of%20CBPFs%20-%20Synthesis%20Report.pdf#:~:text=OCHA%20Evaluation%20of%20Country-Based%20Pooled%20Funds%20-%20Global.An%20evaluation%20team%20from%20KonTerra%20conducted%20the%20evaluation.	2019	Somalia South Sudan Occupied Palestinian Territory Iraq Afghanistan DRC Ethiopia Yemen Syria
WFP and The Cash Learning Partnership (CaLP) Evaluation of the 2017 Somalia Humanitarian Cash-Based Response	https://www.wfp.org/publications/somalia-humanitarian-cash-based-response-2017-joint-evaluation	2018	Somalia
CARE International in Pakistan Humanitarian Project in South Waziristan Tribal District End of Project Evaluation Report (2019).	https://careevaluations.org/evaluation/humanitarian-project-in-south-waziristan-tribal-district-swtd-end-of-project-evaluation-report/	2019	Pakistan
ACTED Final Evaluation Report – Disaster Response in Yemen 2019	https://www.alnap.org/help-library/final-external-evaluation-report-of-concerns-disaster-response-in-yemen	2019	Yemen
Norwegian Church Aid - Sustainable WASH Assistance to the 2010 Flood-Affected Communities in Naushahro Feroze, Sindh 2017	https://www.norad.no/en/toolspublications/publications/ngo-evaluations/2017/final-evaluation--sustainable-wash-assistance-to-the-2010-flood-affected-communities-in-naushahro-feroze-sindh/	2017	Pakistan

Evaluation of UNHCR's Level-3 Emergency Response to Cyclone Idai	https://www.unhcr.org/afr/608fd5cf4.pdf	2021	Mozambique Malawi Zimbabwe
Evaluation of the European Union's Humanitarian Interventions in India and Nepal, 2013-2017	https://www.alnap.org/system/files/content/resource/files/main/kr-04-18-059-2a-n.pdf	2018	India Nepal
Formative Evaluation of UNICEF Work to Link Humanitarian and Development Programming	https://www.unicef.org/evaluation/formative-evaluation-unicef-work-link-humanitarian-and-development-programming-0	2021	Ethiopia Indonesia
Review of the UNICEF L2 Response In Venezuela	http://www.uneval.org/evaluation/reports/detail/18184	2020	Venezuela
Sustainable WASH Assistance to the 2010 Flood-Affected Communities in Naushahro Feroze, Sindh	https://www.norad.no/en/toolspublications/publications/ngo-evaluations/2017/final-evaluation--sustainable-wash-assistance-to-the-2010-flood-affected-communities-in-naushahro-feroze-sindh/	2017	Pakistan
World Vision Cyclone Idai DEC Evaluation Final Report	https://www.alnap.org/help-library/world-vision-cyclone-idai-dec-evaluation-final-report	2020	Zimbabwe
Evaluation of the 2017 Somalia Humanitarian Cash-Based Response	https://www.wfp.org/publications/somalia-humanitarian-cash-based-response-2017-joint-evaluation	2018	Somalia
Independent Review – Iraq Humanitarian and Stabilisation Package	https://www.dfat.gov.au/sites/default/files/iraq-humanitarian-and-stabilisation-package-review-2017-2020.pdf	2020	Iraq
Disasters Emergency Committee - East Africa Response Review: South Sudan	https://issuu.com/decuk/docs/dec_east_africa_appeal_south_sudan	2017	South Sudan
Real-Time Evaluation of World Vision's Response to the Refugee Influx in Lunda Norte, Angola	https://www.alnap.org/help-library/real-time-evaluation-of-world-vision%E2%80%99s-response-to-the-refugee-influx-in-lunda-norte	2018	Angola
NRC Global Cash Evaluation	https://www.alnap.org/help-library/nrc-global-cash-evaluation	2019	Ethiopia Somalia Iraq
Real-Time Assessment (RTA) of UNICEF's Response to COVID-19 in Latin America and the Caribbean (LAC)	https://www.alnap.org/system/files/content/resource/files/main/Regional%20RTA%20Report%20LACRO_DEF%20210411.pdf	2021	Argentina Dominican Republic El Salvador Venezuela (Bolivarian Republic of)
External Ex-post Evaluation of the "Multisectoral Protection Response for Vulnerable Populations in Ecuador affected by the Humanitarian Crisis"	https://careevaluations.org/evaluation/external-ex-post-evaluation-of-the-multi-sectoral-protection-response-for-vulnerable-populations-in-ecuador-affected-by-the-humanitarian-crisis-project-executed-by-care-alas-de-colibri-foundation/	2020	Ecuador
Humanitarian Project in South Waziristan Tribal District (SWTD): End of Project Evaluation Report	https://careevaluations.org/evaluation/humanitarian-project-in-south-waziristan-tribal-district-swtd-end-of-project-evaluation-report/	2019	Pakistan
Final External Evaluation Report: Disaster Response in Yemen	https://www.alnap.org/help-library/final-external-evaluation-report-of-concerns-disaster-response-in-yemen	2019	Yemen
Response to the Rohingya Humanitarian Crisis Phase II Evaluation	https://static1.squarespace.com/static/5ab0691e5417fc8a1ee9a417/t/6110c3e-3c5e4d5325312fc75/1628488683798/AHP+Bangladesh+Phase+II+Evaluation+Final+Report.pdf	2021	Bangladesh
Response to the 2018 Papua New Guinea Highlands Earthquake Evaluation	https://static1.squarespace.com/static/5ab0691e5417fc8a1ee9a417/t/5da3adc6e91096688f5c2a50/1571008011101/PNG+2018+Highlands+Earthquake+Evaluation+Report_FINAL.pdf	2019	PNG

External review of DFAT's Support to TB Response in PNG - 2011 to 2018	https://www.alnap.org/help-library/external-review-of-dfats-support-to-tb-response-in-png-2011-to-2018	2019	PNG
Midterm Review of Malaysia COVID-19 Response	https://www.ifrc.org/media/49544	2021	Malaysia
Missed Out: The Role of Local Actors in the Humanitarian Response in the South Sudan Conflict (2016)	https://oxfamlibrary.openrepository.com/handle/10546/606290	2016	South Sudan
Timor-Leste Drought Operation	https://www.ifrc.org/media/13652	2017	Timor-Leste
Real-Time Evaluation Indonesia: Earthquakes and Tsunami (Lombok, Sulawesi) 2018	https://www.alnap.org/help-library/real-time-evaluation-indonesia-earth-quakes-and-tsunami-lombok-sulawesi-2018	2019	Indonesia
Responding to Pakistan's IDP Fund (Responding to Pakistan's Internally Displaced (RAPID) Fund-II) (Concern Worldwide; 2019)	https://admin.concern.net/sites/default/files/documents/2020-09/Pakistan%20RAPID%20Fund%20II%20Final%20Evaluation%20Sept%202019.pdf	2019	Pakistan
Evaluation of the UNHCR Regional Refugee Response to the Venezuela Situation	https://www.alnap.org/help-library/evaluation-of-the-unhcr-regional-refugee-response-to-the-venezuela-situation	2020	Brazil, Colombia, Ecuador and Peru
Real-Time Review of DEC's Response to Cyclone Idai	https://www.alnap.org/help-library/real-time-review-of-decs-response-to-cyclone-idai	2019	Mozambique, Malawi, Zimbabwe
Evaluation of the WFP Regional Response to the Syrian Crisis (2015-2018)	https://www.alnap.org/help-library/evaluation-of-the-wfp-regional-response-to-the-syrian-crisis-2015-2018	2018	Syria, Jordan, Lebanon, Turkey, Egypt, Iraq
Independent Evaluation of Plan International DEC-Funded response to the Nepal Earthquakes (2015)	https://www.alnap.org/help-library/independent-evaluation-of-plan-international-dec-funded-response-to-the-nepal	2018	Nepal
Independent Evaluation of UNHCR's L3 Emergency Response in Bangladesh	https://www.alnap.org/help-library/independent-evaluation-of-unhcrs-l3-emergency-response-in-bangladesh	2018	Bangladesh
Independent Evaluation of UNHCR's Emergency Response to the Rohingya Refugees Influx in Bangladesh 2017-2018	https://www.ids.ac.uk/publications/independent-evaluation-of-unhcrs-emergency-response-to-the-rohingya-refugees-influx-in-bangladesh-august-2017-september-2018/	2018	Bangladesh
Evaluation of UNHCR Prevention of, and Response to, SGBV in Brazil Focusing on the Population of Concern from Venezuela (2017-2018)	https://www.alnap.org/help-library/evaluation-of-unhcr-prevention-of-and-response-to-sgbv-in-brazil-focusing-on-the	2019	Brazil Venezuela (Bolivarian Republic of)
Evaluation of the European Union's Humanitarian Interventions in India and Nepal, 2013-2017	https://www.alnap.org/system/files/content/resource/files/main/kr-04-18-059-2a-n.pdf	2018	India, Nepal
Evaluation of IFRC West Africa Ebola Viral Disease Appeal Response Sierra Leone and Liberia (2018)	https://www.alnap.org/help-library/evaluation-of-ifrc-west-africa-ebola-viral-disease-appeal-response-sierra-leone-and	2018	Sierra Leone, Liberia
Independent Evaluation of the UNHCR South Sudanese Refugee Response in White Nile State, Sudan (2013 - 2018)	https://www.alnap.org/help-library/independent-evaluation-of-the-unhcr-south-sudanese-refugee-response-in-white-nile-state	2018	Sudan
Final Evaluation Report	https://admin.concern.net/sites/default/files/documents/2020-09/Pakistan%20RAPID%20Fund%20II%20Final%20Evaluation%20Sept%202019.pdf	2019	Pakistan

Humanitarian Assistance in the Pacific: An Evaluation of the Effectiveness of Australia's Response to Cyclone Pam	https://humanitarianadvisorygroup.org/insight/humanitarian-assistance-in-the-pacific-an-evaluation-of-the-effectiveness-of-australias-response-to-cyclone-pam/	2017	Pacific
Final Report for the Final Evaluation of OFDA Response Program (July to October 2021)	https://reliefweb.int/report/syrian-arab-republic/final-report-final-evaluation-ofda-response-program-july-october-2021	2021	Syria
Evaluation of the Humanitarian Innovation and Evidence Programme (HIEP): Summative Phase 2: Final Report	https://www.gov.uk/research-for-development-outputs/evaluation-of-the-humanitarian-innovation-and-evidence-programme-hiep-summative-phase-2-final-report	2018	Global
Evaluation of DFID's Humanitarian Response to Typhoon Haiyan (Yolanda)	https://www.gov.uk/government/publications/evaluation-of-dfids-humanitarian-response-to-typhoon-haiyan	2015	Philippines
DEC – East Africa Response Review: South Sudan (2017)	https://issuu.com/decuk/docs/dec_east_africa_appeal_south_sudan	2017	South Sudan
UNICEF Review of the L2 Response in Venezuela (2020)	https://www.researchgate.net/publication/355873806_REVIEW_OF_THE_UNICEF_L2_RESPONSE_IN_VENEZUELA_FINAL_REPORT	2020	Venezuela
Final Evaluation of OFDA Response Program July to October 2021 (CARE Turkey and partners in NorthWest Syria)	https://reliefweb.int/report/syrian-arab-republic/final-report-final-evaluation-ofda-response-program-july-october-2021	2021	Syria
Final Evaluation of the Emergency Response and Economic Recovery for Eastern Democratic Republic of the Congo	file:///C:/Users/poult/Downloads/Samaritan's%20Purse%20-%20DRC%20-%20720FDA19GR00267%20-%20Program%20Performance%20Evaluation%20Report%20-%202021.pdf	2022	Democratic republic of Congo
The National Health Cluster in Yemen: Assessing the Coordination of Health Response during Humanitarian Crises	https://jhumanitarianaction.springeropen.com/articles/10.1186/s41018-022-00117-y	2022	Yemen
Evaluation of the European Union's Humanitarian Response to the Refugee Crisis in Turkey	https://www.alnap.org/help-library/evaluation-of-the-european-union%E2%80%99s-humanitarian-response-to-the-refugee-crisis-in	2019	Turkey
Plan International UK's DEC-Funded Response to the Rohingya Refugee Crisis in Bangladesh	https://www.alnap.org/help-library/plan-international-uks-dec-funded-response-to-the-rohingya-refugee-crisis-in-bangladesh	2019	Bangladesh
Responding to Pakistan's Internally Displaced (RAPID) Fund (RAPID Fund-II)	https://www.alnap.org/help-library/evaluation-of-ifrc-west-africa-ebola-viral-disease-appeal-response-sierra-leone-and	2018	Sierra Leone, Liberia
Evaluation of the Coverage and Quality of the UNICEF Humanitarian Response in Complex Humanitarian Emergencies	https://reliefweb.int/report/world/evaluation-coverage-and-quality-unicef-humanitarian-response-complex-humanitarian	2019	Global
Corporate Emergency Evaluation of the WFP Regional Response to the Syrian Crisis	https://docs.wfp.org/api/documents/WFP-0000100097/download/	2018	Syria

Related ALNAP publications

- Summary Brief: Review of the OECD DAC criteria for evaluating humanitarian action

ALNAP

alnap.org/me

ALNAP
ODI
203 Blackfriars Road
London SE1 8NJ
United Kingdom
alnap@alnap.org