Evaluating the Coherence of the International Response to the COVID-19 Pandemic

Scoping Study for the COVID-19 Global Evaluation Coalition

Roger Drew

Executive Summary

S1. The COVID-19 Global Evaluation Coalition commissioned a scoping study focused on providing support to participants planning evaluation of the coherence of the international response to the COVID-19 pandemic, based on the OECD DAC evaluation criterion of coherence. This study was intended to inform the development of thoughtful, insightful and useful evaluation and to prompt critical thinking by suggesting relevant concepts and ways of thinking that can be applied in different contexts. This was a desk-based study consisting of review of almost 200 documents and interviews with 16 key informants.

S2. The terms of reference specified three tasks to be carried out through the study with 2-4 questions to be answered in each task. These are shown in Table S1.

Table S1: Study tasks and questions[1] as identified in the terms of reference

Task	Questions
TASK 1. Scoping the topic of coherence in the context of COVID-19 and identification of strategic evaluation questions.	1. What are the most relevant evaluation questions related to Coherence, and which stakeholders are interested in these questions? 2. Which coherence issues are raised at different levels of analysis: institutional, country-level and global? 3. What aspects of the COVID-19 response and recovery effort (the immediate health response, secondary effects including on education and livelihoods, or building back sustainably and equitably) are the most pertinent when it comes to evaluating the coherence criterion? 4. What are the boundaries between coherence and other criteria, including relevance and effectiveness?
TASK 2. Analysis of feasibility of answering the identified questions	5. Which evaluation approaches and methods will enable agencies to evaluate coherence – and answer the identified questions – in ways that are meaningful, feasible and manageable? 6. Is there sufficient, relevant data being collected and/or available to evaluate coherence of the response to the pandemic (dimensions identified above)? 7. Which data systems and resources need to be developed to maximise the feasibility and success of evaluations?
TASK 3. Identification of appropriate, credible processes and ways of working	8. Which overall ways of working will support Coalition participants to design credible, ethical, timely and appropriate evaluations? 9. Which processes will ensure plans and deliverables are inclusive, credible and maximise opportunities for audiences to use findings to inform decisions?

S3. The study was framed in terms of the OECD DAC (2019) criterion and definition of coherence "*the compatibility of the intervention with other interventions in a country, sector or institution*".

S4. It considers how coherence interacts with other concepts within and beyond the OECD DAC evaluation criteria. The main sub-divisions of coherence, internal and external, were also explored. While many of the studies reviewed were conducted prior to this definition being agreed, the study uses this definition as a starting point for analysis and recommendations. As a

---

[1] During the study it was subsequently agreed that the report will focus on questions 1-5 and 8 and 9 as they reflect the priorities of the COVID-19 Global Evaluation Coalition.

secondary point of reference, the study draws on the Coalition's strategic evaluation question that focuses on coherence: "*to what extent are responses[2] aligning to ensure coherent approaches at global, regional and country levels*?"

S5.	The study reviews available evidence of approaches to evaluating coherence prior to the COVID pandemic, drawing in particular on work in the humanitarian sector and on policy coherence for development. This identified considerable heterogeneity in terms of what elements of coherence were being evaluated ranging from instrumental approaches which view coherence in terms of the benefits it brings to the intervention being evaluated to approaches which evaluate coherence in terms of a particular sector or system, other sectors, the totality of an institution's interventions or broader themes and topics, such as the Sustainable Development Goals (SDGs).

S6.	The study identified different ways of working, such as joint or inter-agency evaluations, and various subjects of evaluation, for example particular institutions, particular themes or particular countries. In many cases, an evaluation combines different subjects, for example evaluating an intervention by a particular agency on a particular topic in a particular country.

S7.	The study then reviews experience of evaluating coherence in responses to COVID-19. In most cases, this experience is drawn from evaluations in the design or early implementation phase. The same elements of coherence were identified as for evaluations of coherence prior to COVID-19. Similar ways of working were also identified with perhaps more prominent joint or inter-agency evaluations, particularly among multilateral agencies. As with evaluations of coherence before COVID-19, evaluations in particular countries tended to be of particular agencies or themes and not the totality of that country's COVID-19 response.

S8.	The study then reviews the COVID-19 Global Evaluation Coalition strategic evaluation question on coherence[3] and how this might be used as the basis for developing questions for a specific evaluation. Throughout the report, examples of specific questions included in particular evaluations are presented with further information about design, implementation and findings. The report presents a process for developing coherence-related questions that reflects the need for evidence and to consider the specific context of an individual evaluation.

S9.	The study then considers different ways of evaluating and learning about the coherence of COVID-19 responses. These include particular approaches to evaluation, such as real-time evaluation, specific tools and methods and particular ways of working, such as joint evaluations and different approaches to evaluation synthesis. As noted in other reviews, the vast majority of evaluation reports reviewed relied solely on descriptive or analytical narrative based on collecting data from interviews and review of documents and management information. Isolated examples of use of other methods and tools were identified including indices, analytical frameworks, policy coherence matrices, evaluation rubrics (including RAG rating systems), social network analysis; and various forms of toolkit.

S10. The report concludes with a section of equity and inclusion, some discussion and conclusions and a number of recommendations:

- The newly-adopted OECD DAC evaluation criterion on coherence provides a good basis for evaluations wishing to assess the coherence of responses to COVID-19. The criterion is helpful as it emphasises the importance of evaluating an intervention's fit. The delineation of internal and external coherence is particularly useful for evaluations of well-defined entities. However,

---

[2] Responses from national governments, humanitarian agencies and development partners.
[3] And suggested examples of questions for thematic and country-level questions.

the distinction may be less useful where the entity being evaluated is complex or amorphous, e.g. a United Nations agency or where an evaluation is not of a specific entity, e.g. a thematic evaluation.

- Evaluations wishing to include the coherence criterion should first identify the subject of the evaluation and what elements of coherence will be evaluated in the specific context under consideration. This will involve understanding the entity or entities being evaluated and identifying other relevant actors. It will also involve being clear as to "*coherence with what*" is being evaluated, for example, as illustrated in Figure 3. Is the evaluation interested in fit with other interventions in the same sector or another sector? Or is the main focus on how the intervention fits with what other parts of the entity are doing? Or is the main focus on fit with broader themes and topics, such as the SDGs, the humanitarian-peace-development triple nexus, human rights, equity and inclusion.

- While generic questions on coherence may be a helpful starting point and it may also be possible to learn from questions others have used, each evaluation will need to carefully develop evaluation questions that are relevant to the intervention(s) being evaluated and their context.

- Where possible, evaluations should seek to go beyond simple analytical narratives as a way of describing and communicating findings concerning coherence. The use of tools, such as red amber green (RAG) ratings to present findings from rubric-based evaluative approaches may be useful to aid communication but they have to date been used by relatively few evaluations relating to coherence.

- Evaluators should consider equity and inclusion when designing and implementing evaluations focused on the coherence of responses to the pandemic. This may involve reflecting on the relevance of the leave no one behind agenda within responses to the pandemic, if and how marginalised groups have benefited from the coherence of responses and whether their needs and perspectives are being taken into consideration as plans develop.

- The review proposes ways of working to support the development of meaningful, feasible and manageable evaluations of coherence. Where possible, joint evaluations may have particular advantages in answering questions related to coherence. Where these are not possible, different synthesis approaches may be useful in summarising evaluative evidence from evaluations conducted by different actors.

- Practical steps can be taken to support the development of evaluation portfolios that can be synthesised. These include developing and applying a shared analytical framework to inform the design of evaluations, working in ways that support complementarity rather than duplication of evaluations, development of a comprehensive evidence base, and development ofwork plans that include sufficient time for participatory reflection and collaboration across different agencies

Table of Contents

Glossary

|	
ADB	Asian Development Bank
ALNAP	Active Learning Network for Accountability and Performance
AMR	Antimicrobial Resistance
CARE	Cooperative for Assistance and Relief Everywhere
CDI	Commitment to Development Index
CGD	Center for Global Development
COVID	Coronavirus Disease
CRID	Center of Resources for Innovation and Development
DFID	Department for International Development
ECDC	European Centre for Disease Prevention and Control
ECDPM	European Centre for Development Policy Management
EPYPSA	Estudios, Proyectos y Planificación Soociedad Anónima
EU	European Union
EvalNet	The DAC Network on Development Evaluation
FAO	Food and Agriculture Organization of the United Nations
FGM	Female Genital Mutilation
GHRP	Global COVID-19 Humanitarian Response Plan
IADB	Inter-American Development Bank
IAHE	Inter-Agency Humanitarian Evaluation
ICAI	Independent Commission for Aid Impact
IFRC	International Federation of Red Cross and Red Crescent Societies
ILO	International Labour Organization
IOM	International Organization for Migration
IPA	Innovations for Poverty Action
K4D	Knowledge, Evidence and Learning for Development
MENARID	Middle East North Africa Regional Integrated Development
MPTF	Multi-Partner Trust Fund
OCHA	United Nations Office for the Coordination of Humanitarian Affairs
ODI	Overseas Development Institute
OECD DAC	Organisation for Economic Cooperation and Development's Development Assistance Committee
OIE	World Organisation for Animal Health
PCD	Policy Coherence for Development
RAG	Red Amber Green
SDG	Sustainable Development Goal
TOR	Terms of Reference
UK	United Kingdom
UN	United Nations
UNDAC	United Nations Disaster Assessment and Coordination
UNDP	United Nations Development Programme
UNFPA	United Nations Population Fund
UNHCR	The UN Refugee Agency
UNICEF	United Nations Children's Fund
WFP	World Food Programme
WHO	World Health Organization

Introduction

1. The EvalNet Secretariat commissioned this rapid scoping and initial evaluability assessment that uses the OECD-DAC definition of "*coherence*", to provide support to COVID-19 Global Evaluation Coalition participants to plan evaluations of the coherence of the international response to the COVID-19 pandemic. This involves providing strategic and practical advice to support participants to develop evaluation plans that are feasible, credible and useful. Full terms of reference for this study are available in Annex 1 (p67).

2. This report first briefly explains the methods used (paragraphs 6-9) before seeking to provide some definition and analytical framing of coherence for the discussion that follows (paragraphs 6-11). It then presents experience of evaluating coherence prior to COVID-19 (paragraphs 12-30) before then considering how coherence has been or will be evaluated in evaluations of COVID-19 responses (paragraphs 31-49). The report then seeks to identify what questions an evaluation might ask about coherence and how these might be developed (paragraphs 50-75) before assessing what approaches, tools, methods and ways of working might be available to evaluate coherence (paragraphs 76-86). The report concludes with some reflections on the importance of considering equity and inclusion, a discussion about the key findings, conclusions and a number of recommendations.

The COVID-19 Global Evaluation Coalition

3. The COVID-19 Global Evaluation Coalition is an independent collaboration bringing together development evaluation units of countries, United Nations organisations and other multilateral institutions. The Coalition seeks to provide credible evidence to inform international co-operation supporting non-clinical responses to and recovery from the COVID-19 pandemic in developing countries, helping to ensure that lessons are learned and that the global development community delivers on its promises. The Coalition does this by supporting and communicating both individual participants' evaluations, and joint work involving multiple participants. This collaborative approach seeks to maximize synergies and learning, while reducing duplication in evaluating different elements of the COVID-19 pandemic response. The OECD DAC Network on Development Evaluation (EvalNet) serves as Secretariat for the Coalition.

4. To provide a basis for assessment and to accommodate the exceptional circumstances facing implementing partners and evaluation units, the Coalition is taking a phased and modular approach to its work. This involves both individual evaluation work carried out by participants around a common evaluation framework (see Figure 1) and joint work involving multiple participants. This approach aims to support coherence of activities, minimise duplication, and maximise opportunities for shared learning and collaborative work. A series of evaluations and related analytical work will be led by participants of the Coalition over time to meet different needs. This scoping study builds on Coalition-led work to develop strategic evaluation questions and mapping work to analyse global evaluation plans across the Coalition. It also complements a scoping study that has been commissioned by the Secretariat to support participants of the Coalition to develop feasible, credible and useful evaluation plans that focus on Building Back Greener.
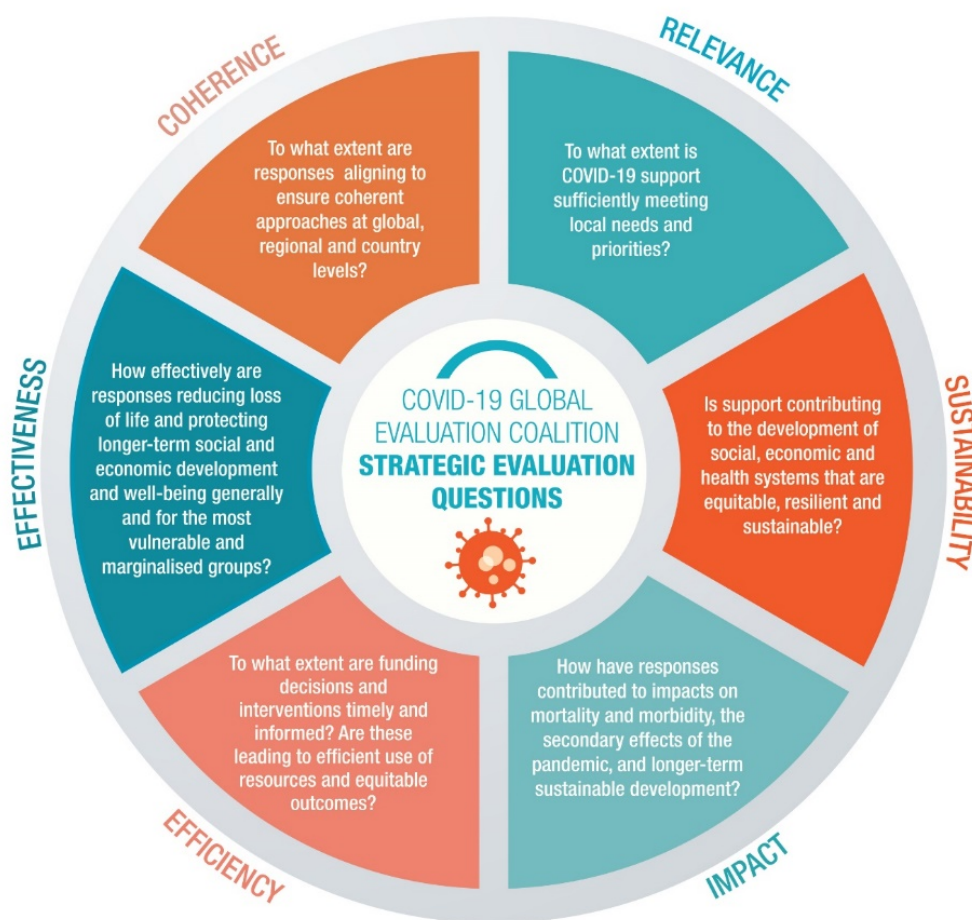
Figure 1: Strategic Evaluation Questions Framework by the COVID-19 Global Evaluation Coalition

5. There was an expectation of linking the study to the work OECD had done/was doing to develop strategic evaluation questions. However, this work was happening at the same time, meaning that these questions were iterated over the lifetime of the study. This study was drafted in parallel with other streams of Coalition-led work, notably work to finalise the strategic evaluation questions and the development of the Building Back Greener scoping study. Given the emerging nature and development of the pandemic, and the need for the Coalition to work in ways that are both strategic and demand driven it has been necessary for streams of work to be completed in parallel to each other.

Defining and Framing Coherence

6. For this study, understanding of the term coherence was provided by the revised OECD DAC criteria which added coherence as a criterion (OECD DAC Network on Development Evaluation, 2019).[4] Coherence is defined as "*the compatibility of the intervention with other interventions in a country, sector or institution*".

7. The note that accompanies this definition provides further guidance to inform evaluations

---

[4] However, it should be noted that many of the evaluations reviewed, particularly those which occurred prior to the occurrence of the COVID-19 pandemic pre-dated the adoption of this criterion.

'The extent to which other interventions (particularly policies) support or undermine the intervention, and vice versa. 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.' (OECD, 2021, p45)

8. Systems thinking is an over-arching conceptual and theoretical approach which informed the definition of coherence, adaptation of all the OECD/DAC evaluation criteria and drafting of the OECD Guidance 'Better Criteria for Better Evaluation' (OECD, 2021). Its application within evaluation is characterized by the identification of three concepts: dynamic interrelationships between different actors, different perspectives, and boundaries (Williams and Imam, 2007)

9. Systems thinking has been applied within a range of analytical disciplines and within some evaluations, including a small number of evaluations that have focused on coherence. Systems thinking, as a conceptual and theoretical approach, was not applied within the majority of evaluations reviewed for this study although it has informed some evaluations in in the humanitarian field (e.g. ALNAP, 2018; Burrett, 2019). A recent discussion paper, based on review of multiple humanitarian evaluations concluded that most looked at project or single-agency level and that there was "*a marked absence of studies that looked at systemic or response-wide issues*" (Darcy and Dillon, 2020).

10. Table 1 briefly presents some concepts that are related to coherence based on how different evaluations that are reviewed here have been conducted. Some of these concepts are other OECD DAC criteria and this material emphasises the inter-connectedness of these criteria. Others are not OECD DAC evaluation criteria but they have been used as evaluation criteria in particular contexts, such as humanitarian settings (OECD DAC, 1999) and by other organisations, such as ALNAP (ALNAP, 2006a; ALNAP, 2016) and others (for example European Commission, 2006).

Table 1: Brief review of concepts that are related to coherence

Within other OECD DAC criteria	Beyond the other OECD DAC criteria
While consideration of fit between national government interventions and interventions of other actors may be considered to be coherence, consideration of alignment to national policies falls under the OECD DAC evaluation criterion of *relevance*.	*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 (ALNAP, 2006)
While many aspects of partnership and coordination fall under coherence, elements of partnership and coordination which are "*instrumental*", i.e. which are intended for the benefit of the intervention, e.g. to achieve its objectives, fall under the OECD DAC evaluation criterion of *effectiveness*.	*Coordination* is defined by the European Union's Heads of Evaluation Task Force (2003) as 'activities of two or more development partners that are intended to mobilise aid resources or to harmonise their policies, programmes, procedures and practices so as to maximise the development effectiveness of aid resources'
While connectedness between the humanitarian and development nexus might be	*Complementarity,* *defined by the European Union's Heads of Evaluation Task force '*is

considered to be related to coherence, it could also be considered under *sustainability*.	intended to ensure that Community development policy shall be complementary to the policies pursued by the Member States'.
The OECD/DAC guidance Applying Evaluation Criteria Thoughtfully notes that there are links between coherence and **efficiency** 'incoherent interventions may be duplicative, thus wasting resources (OECD, 2021, p45). Consequently, evaluations may wish to consider the extent to which efforts to ensure the coherence of activities support and/or limit the efficiency of the intervention.	
Impact: evaluations may explore and understand the extent to which efforts to ensure coherence of different interventions and strategies have increased and/or limited different impacts for different groups.	

11. The main sub-division of the coherence criterion within the OECD/DAC criteria is between internal and external coherence. Internal coherence relates to fit within the same institution or government including internal policies and any commitments to international laws and agreements. External coherence relates to fit with other actors' interventions in a given context. These other actors would include other development and humanitarian agencies but would also include other types of actors, such as military and security actors, where relevant, and local-level actors. This sub-division, of internal and external, is based on the assumption that an organisation or entity can be clearly and consistently identified and defined. This may be difficult for some complex entities. For example, in an evaluation of a UN agency, the entity may be defined as the UN as a whole, as the particular agency or potentially as the agency's country office depending on the perspective of whoever is commissioning the evaluation. In systemwide evaluations, e.g. of the humanitarian system, internal may be defined in terms of being within the system or as internal to one or more organisations operating within that system. A second issue is that some evaluations may not be of a single entity, e.g. they may be of a theme or issue. In such evaluations, internal coherence may need to be considered in terms of multiple agencies or entities.

Experience of evaluating coherence to date

12. Much of the experience of evaluating coherence comes from the humanitarian sector (Hallam, 1998; OECD DAC, 1999; ALNAP, 2006a; Ministry of Foreign Affairs of Denmark, 2006; ALNAP, 2018; Darcy and Dillon, 2020) (see Box 1) including from crises in general (UNDP, 2020a; Bastøe et al, 2020; World Bank Independent Evaluation Group, 2020; WFP, 2020c; UNDP 2021) (see Box 2).

Box 1: Much of the experience of evaluating coherence comes from the humanitarian sector

In 1998, Alistair Hallam produced a good practice review on evaluating humanitarian assistance programmes in complex emergencies. One of the recommendations of this was use of evaluation sub-criteria, such as connectedness and coherence. It documented that one of the first sub-evaluations to use these criteria had been the Joint Evaluation of Emergency Assistance to Rwanda in 1996. Coherence was specifically suggested as a criterion for evaluating humanitarian assistance in complex emergencies by OECD in 1999. In 2006, ALNAP produced a guide for humanitarian agencies on using the OECD DAC criteria, including coherence, to evaluate humanitarian action. This gave good practice examples, including the evaluation of UNHCR's emergency preparedness and response in Kosovo. It also identified key questions to ask in the case of coherence including:

- Why was coherence lacking or present?
- What were the particular political factors that led to its coherence or its lack?
- Should there be coherence at all?

In 2006, Denmark included coherence as an additional criterion for evaluating humanitarian action. In its assessment of the state of the humanitarian system, ALNAP includes a section assessing coherence. These assessments cover a number of implicit and explicit questions on coherence including:

- To what degree are humanitarian efforts coherent with core principles and international humanitarian law?
- To what degree are humanitarian actors effective in encouraging support for international humanitarian law and international refugee law?
- What factors affect coherence? (level of ambition, an increasingly hostile geopolitical environment, increased links between humanitarian and development/stabilisation activities, and lack of understanding/commitment on the part of humanitarian staff)

It is therefore clear that coherence has been long-considered a criterion for humanitarian evaluations. ALNAP and OECD included coherence as a criterion for humanitarian evaluations before the new OECD/DAC criteria were adopted, which are applicable to development and humanitarian evaluations. Despite this, Darcy and Dillon's 2020 discussion paper raised concerns that "*the great majority* [of humanitarian evaluations] *deal with context-specific crisis responses by individual agencies; relatively few are concerned with system-wide performance or organisational performance across a range of different contexts*". Part of the problem might be (as stated in the ALNAP 2006 guide] that there is an understanding that "*coherence may be less relevant for evaluating single-agency or single-project interventions*" when potentially such interventions may be most at risk of fitting poorly with what others are doing (or with what other parts of the same agency are doing) and might conversely benefit most from evaluating coherence.

Box 2: There is experience of evaluating coherence from evaluations of humanitarian responses to a range of crises

One of the lessons UNDP considered it had learned from providing governance support to countries in crisis for different reasons (e.g. conflict in Yemen, natural disaster in the Philippines) was the importance of a well-designed crisis response plan for establishing strong partnerships and for encouraging coherent humanitarian and development interventions. This lesson was included with others in UNDP's 2021 document focused on learning from past crises for recovering from COVID-19. While none of the other lessons explicitly mention coherence, many are relevant to it including that, in the sector of livelihoods restoration and job creation, frameworks for cooperation and coordination across UN agencies and other partners are critical, crisis-response interventions have a greater impact when coupled with a broad package of development support and clear transition from crisis response to recovery is essential.

In their 2020 blog, Per Bastøe, Wendy Brusse and Jörg Faust argued that responses to COVID needed to be based on lessons learned from evaluations of responses to previous epidemics and disasters. While they do not mention coherence explicitly, some of the four lessons they identify relate to coherence implicitly including the importance of donor coordination to avoid donor fragmentation and project proliferation and the need to align and transition between humanitarian assistance and longer-term co-operation for resilient country-owned health systems.

The 2020 WFP strategic evaluation of their capacity to respond to emergencies identified that sometimes there were tensions between elements of coherence, for example between "*striving for coherence with government priorities on the one hand and adherence to humanitarian principles on the other*".

---

13. Some experience of evaluating coherence comes from specific humanitarian crises including:

- Previous health crises (Office of Evaluation and Oversight IADB, 2020; Gold and Hutton, 2020), such as Ebola (Thormar, 2013; United Nations et al, 2015; World Bank, 2015; Momoh et al, 2016; Ali and Hutton, 2016; WFP, 2017; UNICEF, 2017; ECDC, 2017; Lamoure and Juillard, 2020), Avian Flu (World Bank Independent Evaluation Group, 2020) and cholera (UNICEF, 2018) (see Box 3).

Box 3: There is experience of evaluating coherence in evaluations of responses to health crises, particularly Ebola

In considering what could be learned from evaluating past public health crises, IADB did not explicitly refer to coherence. However, one of the lessons learned was the importance of collaboration and knowledge sharing, namely to learn from other emergencies and to harness knowledge and expertise from across the organisation. Similarly, although the 2020 blog by Gold and Hutton, for the World Bank, did not explicitly refer to coherence, some of the lessons learned relate to coherence implicitly including, under mounting a rapid response, that partnerships contribute to mitigating risks related to rapid project preparation and cooperation and coalition building among countries can strengthen response performance and address longer-term needs.

Much of the experience of evaluating coherence in previous health crises comes from Ebola in West Africa. For example, in 2015, the United Nations and others looked at the response to Ebola in Liberia, Sierra Leone and Guinea. One of their recommendations was the need for regional organisations, such as the African Union, the Economic Community of West African States and the Mano River Union, to envisage joint activities and to work together with each other and with national governments to "*ensure coherent linkages and coordination mechanisms between country and regional level interventions to ensure synergies*". While the World Bank's review of Nigeria's response to Ebola does not mention coherence explicitly, it did conclude that one of the key lessons was the importance of building on and strengthening existing systems rather than introducing new, parallel structures.

When Momoh and others evaluated the Disasters Emergencies Committee response to Ebola, they had specific questions and a section on coherence. In both phases (emergency phase 1 and recovery phase 2), coherence was conceptualised largely in terms of collaboration with faith-based leaders and community heads. In phase 2, coherence also included consideration of relationships with government ministries including health, agriculture and social welfare. Specifically, the questions on coherence for this evaluation were:
- How did phase II build on the successes and learnings generated from phase I, was there a logical progression from phase I to phase II. [i.e. were the phases of the project coherent with each other]
- Did projects in each phase complement one another?
- Have partners been able to link with one another during the project and has there been added value if this has occurred? [this could be seen as internal coherence within the Development Emergencies Committee or external coherence for individual agencies]

However, it is not particularly clear if or how the material presented in the report on coherence answers these specific questions.

When, in 2016, Ali and Hutton evaluated the DFID Ebola Emergency Response Fund in Sierra Leone, they did not explicitly talk about coherence. However, the evaluation did ask how well the DFID Ebola Emergency Response Fund worked with partners, government, stakeholders and donors under the evaluation criterion/area of efficiency. Although initially the evaluation included, as part of the assessment of impact, the question to what extent did the DFID Ebola Emergency Response Fund contribute to a solid and coordinated Ebola virus disease response, this was later merged with another question, under the criteria/area of quality and relevance of design, to what extent the DFID Ebola Emergency Response Fund's adopted approach help respond to Ebola virus disease control.

In contrast, WFP's evaluation of its response to the Ebola virus disease crisis in West Africa in 2014-15 specifically considered coherence and appropriateness among its evaluation criteria. Questions asked related to coherence included:
- Was WFP's response coherent with national priorities and effectively and efficiently coordinated with the governments of Ebola affected countries?
- To what extent was WFP's response coordinated with the United Nations Mission for Ebola Emergency Response and other UN agencies, enabling synergies and multiplying opportunities at strategic and operations levels and taking account of the shifting frameworks for coordination?
- Was WFP's response coherent and aligned with the priorities of other partners, enabling synergies at operations levels?
- To what extent was a transition strategy developed and integrated in implementation, namely in terms of partnerships and stakeholders' involvement and their capacities strengthened through WFP's response?
- To what extent was WFP's response (and activities) aligned to WFP's corporate policies? To what extent were these policies relevant to operational needs and objectives?
- To what extent was WFP's response delivered in a timely, efficient and successful manner by consolidating and coordinating already implemented interventions, and by addressing/advocating to address critical gaps (including coverage, partnerships, and access)?

In UNICEF's 2017 evaluation of its response to the Ebola outbreak in West Africa 2014-15, coherence was one of the evaluation criteria considered. Two of the five key evaluation questions focused on this and were:
- How well coordinated internally was UNICEF's response to Ebola?
- How well coordinated externally was UNICEF's response to Ebola?

Each of these questions had a section of the report dedicated to it. Section III considered ten internal UNICEF functions (such as supply and logistics, finance and administration) but all but one of these (knowledge management) were assessed in terms of enabling effectiveness. Section IV on external coordination is much briefer and is specifically focused on whether elements of strategic and operational coordination "*helped or hindered the effectiveness of UNICEF's response*".

Box 3: There is experience of evaluating coherence in evaluations of responses to health crises, particularly Ebola (continued)

In ECDC's 2017 evaluation of ECDC Ebola deployment in Guinea, coherence was included as one of the evaluation's criteria. The main evaluation question in this area was to what extent does the activity counteract with other activities, internal or external to the Centre, with similar objectives. Other possible research questions identified in this area included:
- To what extent were the deployment activities aligned with the EU and international response objectives in Guinea?
- To what extent were the deployment activities well-coordinated with the EU institutions?

Responses to these questions were succinctly presented in tabular form, i.e.

Coherence	
To what extent were the deployment activities aligned with the EU and international response objectives in Guinea?	As stated above, this evaluation has demonstrated that the ECDC support to the outbreak response activities in the field was well aligned with the needs. As the teams were embedded within and deployed under WHO, their activities were well aligned with the international response.
To what extent were the deployment activities well-coordinated with the EU institutions?	However field coordination and cooperation with EU services and WHO in Guinea throughout this deployment recorded the most negative perceptions from the deployed experts due to: • confusion regarding ECDC/WHO roles and responsibilities, and • lack of consistent link between the ECDC mobilised experts and the EU services in the field.

In 2020, Lamoure and Juillard pulled together a lessons paper for ALNAP based on responding to Ebola epidemics. However, this did not explicitly consider coherence. Rather, its lessons were organised in four areas – (A) healthcare, water, sanitation and hygiene and body management; (B) context, communication and community engagement; (C) Ebola's effects on healthcare, mental health protection, education and livelihoods; and (D) coordination and funding. While perhaps area D is most immediately relevant to coherence, some of the topics in areas B & C also touch on coherence.

Other health crises from which learning is possible include avian influenza and cholera. In 2020, the World Bank's Independent Evaluation Group published learning from evaluation of World Bank experience with avian influenza between 2006 and 2013. This cited concerns that the economic costs of a previous outbreak (of Severe Acute Respiratory Syndrome) had been disproportionately high because of the lack of a coherent, coordinated global response. The paper distinguishes strategic and technical lessons. Some of the strategic lessons are relevant to coherence, including the value of a global framework for avian influenza projects, working across sectors and drawing on expertise from international technical agencies, such as WHO, FAO, OIE and UNICEF. In 2018, UNICEF published an evaluation of its level 3 response to the cholera epidemic in Yemen. One of the areas of focus for this evaluation was on what role the organization played in coordinating, leading or facilitating the response of the wider system, through its cluster leadership and otherwise. The report had a short section on coherence which focused on different components/sectors of the response – health; water, sanitation and hygiene; communication for development; and nutrition including how well these were planned and harmonised together. Some of the 16 recommendations related to coherence including:
- Recommendation 4 which was focused on addressing lack of internal coherence. The recommendation noted that there had been a lack of coherence both in the advisory input on cholera from different UNICEF sections and between the different components of the UNICEF programme.
- Recommendation 5 which was focused on addressing lack of external coherence, particularly between UNICEF and WHO.
- Recommendation 6 which was focused on coordination processes. The recommendation documented that coordination of the 2017 response in Yemen was confused, with multiple mechanisms overlapping and running in parallel. In particular, the respective roles of the clusters (health/water, sanitation and hygiene) vis-à-vis the emergency operations centres were poorly defined. Another essential component of preparedness is the clarification and simplification of the cholera-related coordination processes and the respective roles of the Cholera Task Force, the emergency operations centres, the health/ water, sanitation and hygiene clusters, the Office for the Coordination of Humanitarian Affairs and the Humanitarian Country Team/Inter-Cluster Coordination Mechanism.

- Natural disasters, such as earthquakes (Grünewald et al, 2010; Bhattacharjee and Lossio, 2011; Patrick 2011; Alam and Balthazar, 2011; EPYPSA, 2011; The Humanitarian Coalition, 2012; Baker et al, 2015; Murtaza et al, 2016; Dara, 2016; Bhattacharjee, 2017; Key Aid Consulting, 2018) (see Box 4 for Haiti earthquake and Box 5 for Nepal earthquake), tsunamis (Bhattacharjee, 2005; ALNAP, 2006b; Telford et al, 2006) (see Box 6) and hurricanes/typhoons (Hanley et al, 2014; ICAI, 2014; Itad, 2015; Dy and Stephens, 2016; CRID, 2017; IOM, 2017; White, 2019) (see Box 7).

Box 4: There is experience of evaluating coherence in evaluations of responses to natural disasters: Example of Haiti 2010 earthquake

In 2010, François Grünewald and others conducted an inter-agency real-time evaluation in Haiti three months after the earthquake. This evaluation did not specifically consider coherence as an evaluation criterion but did include a number of evaluation questions related to coherence particularly under the response area of coordination and connectedness. Evaluation questions were organised by four response areas – response covering the needs; strategic and operational planning and resource mobilization; coordination and connectedness; and context and needs. Each response area had a number of overarching questions. Three of the response areas had either one or two overarching questions but there were six related to coordination and connectedness. There were a further 33 specific questions across the response areas. The questions related to coordination and connectedness were:

Overarching questions:
- Has an inclusive and well-managed coordination system been established early on, including with the national actors, the military and all other relevant stakeholders?
- Were activities planned in support to pre-existing response plans, structures and capacities?
- Was the coordination system supported by an efficient communication and information management system (e.g., enhancing information flow within the field, between field and headquarters)?
- What systems have been put into place to monitor, report and evaluate the efficiency and effectiveness of the overall response? How adequate are these for measuring progress against objectives?
- How adequately have cross-cutting issues been dealt with in all aspects of the response and in all clusters/ sectors?
- Was an inclusive common strategy for security and access developed?

Specific questions:
- To what extent does the coordination system support relief and recovery alike?
- In what ways, if any, has the cluster approach led to a more strategic response in terms of predictable leadership, partnership, cohesiveness and accountability?
- How effective has inter-cluster coordination been (with specific focus on cross cutting issues, Protection and Early Recovery)?
- How effective has the set-up of the support hub in Santo Domingo, with shadow cluster-related functions in addition to the cluster activation, been?
- How effectively has the humanitarian community coordinated the response with the Government of Haiti and the international military forces?
- Has an effective integrated accountability framework been put in place? How well functioning and robust is it?
- In what ways, if any, has the government's leadership capacity been strengthen as it has the primary responsibility to respond to its people's needs?
- In what ways, if any, have national and local capacities been capitalized on and strengthened (e.g., in needs assessments?)
- How effectively have partnerships with civil society organizations and the affected communities themselves been built-up in order to maximize local ownership, and thereby enhance effectiveness, accountability and sustainability?

However, it is not clear how these many questions were (or could have) been answered in the evaluation. The report's findings are structured around two main themes, the quality of the response and its structure. Some elements of coherence are covered, for example, the valuable role played by cluster coordination, challenges were faced in terms of staff turnover and the massive influx of international non-governmental organisations, good coordination between UNDAC and OCHA, and good coordination between donors.

In the 2011 evaluation of the OCHA response to the Haiti earthquake, Bhattacharjee and Lossio do consider specific OECD/DAC evaluation criteria including coordination, coherence and connectedness. Evaluation questions were organised around several key areas of enquiry including coordination. Questions related to coordination included:

- How effective was OCHA in supporting the development of adaptive, inclusive and coherent coordination architecture in Haiti? How effectively did OCHA support the Humanitarian Coordinator/Deputy Humanitarian Coordinator and cluster coordinators in their roles in Haiti?
- Were Coordination tools appropriate and effectively administered?
- How effective was the inter-cluster coordination in Haiti after the earthquake?
- How did OCHA facilitate the government's participation in and ownership of cluster processes?
- To what extent national and regional organisations participated in the response tools namely, UNDAC and the International Search and Rescue Advisory Group?
- How effective was OCHA in facilitating coordination with the Department of Peacekeeping Operations and to what extent it effectively interacted with the United Nations Stabilisation Mission in Haiti on behalf of humanitarian community to build up a mutually supportive relationship?
- Examine the role played by OCHA in leveraging the capacity of military forces while ensuring independence, neutrality and impartiality of all humanitarian response?
- How effective was OCHA in facilitating clusters in transition phase from moving from relief to recovery/reconstruction?
- Role played by OCHA in developing tools and processes for system-wide learning on key issues. What were the effects of OCHA's tools, guidance and training on OCHA's performance and on the functioning of coordination system in Haiti?
- Did the Humanitarian Coordinator and cluster members find OCHA's support "predictable and systematic' in the Haiti response?
- To what extent OCHA's support enabled the Humanitarian Coordinator to report on performance of the entire humanitarian system in terms of results?

Box 4: There is experience of evaluating coherence in evaluations of responses to natural disasters: Example of Haiti 2010 earthquake (continued)

Material of relevance to coherence is presented in the summary under the heading of coordination, in section 3 in relation to OCHA's core function of coordination and in section 7 in a specific assessment of the criteria of coordination, connectedness and coherence. For example, topics covered in the summary include the importance of establishing coordination structures beyond the capital city, different understandings of the "principle of last resort" in terms of using military and civil defence assets, and tension between OCHA's humanitarian mandate and the longer-term role of the United Nations Stabilisation Mission in Haiti. Section 7 covers the criterion of coordination by cross-referencing section 3. The section which covers the criterion of connectedness covers the respective roles and strengths of OCHA and the United Nations Stabilisation Mission in Haiti, tensions over use of military assets and linking short-term emergency response with long-term development programmes. The section on coherence focuses on consistency in policies and practices, particularly in relation to gender, sexual exploitation, and on tensions between humanitarian and development budgets.

Because this evaluation report uses multiple analytical framings (questions, key areas of enquiry, evaluation criteria) the report is long with substantial areas of repetition and duplication. The distinction between connectedness and coherence is not particularly clear. For example, why are tensions between humanitarian and development budgets covered under coherence and not under connectedness?

In 2011, Jonathan Patrick identified a number of emerging evaluation lessons from the Haiti Earthquake Response. This was organised around evaluation criteria, including coherence. Three key lessons were identified in relation to coherence, namely the need to support humanitarian leadership from the start; the need to engage with the military in advance to establish protocols regarding division of responsibility, channels of communication and broad coordination mechanisms; and the need to better explore how the humanitarian community can best learn from and implement previous lessons.

In 2011, Alam and Balthazar carried out a real-time evaluation of ActionAid's International Haiti Emergency Response Programme. This did not explicitly consider coherence. That same year, EPYPSA carried out an evaluation for IFRC looking at meeting of shelter needs. The report does not explicitly identify evaluation criteria or questions used but some of the findings imply consideration of some aspects of coherence. For example, the findings section starts with general coordination system difficulties. The only explicit mention of coherence comes in a discussion of concerns from the Haitian government that the shelter response design was not sufficiently adapted to the Haitian context. But, the evaluators quote one government respondent as saying "*we didn't show a coherent shelter and housing plan…*"

The Humanitarian Coalition's 2012 final evaluation report of the Haiti 2010 earthquake response does not explicitly mention coherence. However, one of the report's main findings relates to coordination and collaboration with others and this identified the challenge of comprehensive targeting and coverage, as well as standardisation of the assistance package and accountability systems so that all beneficiaries received comparable assistance regardless of the assisting agencies; limited coordination with government; mixed experiences in relation to cluster coordination; mixed experiences of local government regarding involvement of non-governmental organisations; and generally positive interaction with other civil society groups such as associations and committees.

Box 5: There is experience of evaluating coherence in evaluations of responses to natural disasters: Example of Nepal 2015 earthquake

In 2015, Jock Baker and others carried out for IFRC a real-time evaluation of the Nepal Earthquake Response Operation. This clustered findings under three groups of evaluation criteria. The second grouped coordination with efficiency and effectiveness while the third focused on connectedness alone. The section on coordination covered cooperation with the International Committee of the Red Cross and external coordination (focused on the relative roles of IFRC and the Nepal Red Cross Society). The section on connectedness examines the relationship between capacity building, preparedness, emergency intervention and recovery. In 2016, Niaz Murtaza and others carried out for the ACT Alliance a final evaluation of the Nepal Earthquake Response. This did not explicitly consider coherence but structured the evaluation under four criteria, one of which included coordination. Questions considered under that criterion included:

- How effective was the ACT Secretariat in facilitating and coordinating the response efforts – within requesting members and other ACT members working in Nepal for earthquake responses?
- How did the ACT Nepal programme optimize the value of ACT Alliance's joint appeal system to create greater impact? Were appropriate synergies, institutional platforms and existing national strategy used to leverage ACT response?
- How are organizations addressing the issue of coordination and what leadership are they demonstrating with regards to the challenges?
- Assess the effectiveness of the collaboration. Coordination and coordination mechanisms among ACT members/partners of ACT Alliance as well as with other stakeholders.
- Were the needs and priorities of the affected population, ACT donors and policy standards of ACT Alliance met?
- What were the strengths, weaknesses and opportunities of ACT visibility in a multi-actor ACT response?
- How was the coordination with UN clusters and other external mechanisms? Was it possible to utilize the ACT forum structure to better participate and influence those platforms?

The findings section answered each of these questions in turn although the third and fourth question were combined for this purpose and the fifth question was shortened to only consider ACT donors and policy standards of ACT Alliance.

In Dara's 2016 final report of the evaluation of UNICEF's response and recovery efforts to the Gorkha earthquake in Nepal, coherence was not included as a specific evaluation criterion but coordination was and so was connectedness (with sustainability). The questions for these criteria were:

Connectedness and sustainability
- To what extent have the linkages between relief, recovery and development been addressed in UNICEF's response?
- To what extent did the existing preparedness measures facilitate UNICEF's response?
- To what extent have UNICEF's activities contributed to strengthening Nepal's institutional capacity?

Coordination
- Were UNICEF's resources and staff sufficient to ensure that it could adequately perform its role as cluster lead during emergency, recovery and development? To what extent did UNICEF cluster leads comply with the responsibilities defined in the IASC's ToR of cluster leads at country level?
- To what extent did the internal coordinating tools facilitate the emergency response?

Findings sections were prefaced with these questions but then structured around key themes. For connectedness and sustainability, these were contingency planning/preparedness measures; planning for transition; transition-oriented activities incorporated into programmes. For coordination, these were internal coordination, compliance with cluster lead agency responsibilities, coordination arrangements in UNICEF-led clusters, sub-national coordination and communication with communities.

In 2017, Abhijit Bhattacharjee carried out an evaluation for Islamic Relief Worldwide of their Nepal Earthquake Response. This included assessment against a number of evaluation criteria with coherence and connectedness grouped together. The questions identified for these criteria were:
- Was Islamic Relief Worldwide 's response coherent with relevant Islamic Relief Worldwide policies, international principles and standards?
- To what extent have interventions been coordinated with national/local government and international humanitarian system?
- Are there internal coordination /communication challenges that affected the project? How have these been addressed?
- Were the interventions carried out taking into account longer-term and interconnected problems, and capacity of communities/local authorities?

The summary does address the first of these points but then digresses into issues of capacity building of local masons, migration and reliance on remittances without being clear how these relate to the questions on coherence and connectedness. The body of the report is not much clearer.

In 2018, Key Aid Consulting completed a final evaluation for the Red Cross of their Nepal Earthquake Recovery Programme. This did not mention coherence explicitly. Rather the report was structured around the criteria of relevance, effectiveness, value for money and sustainability. There was a section on coordination with external actors (under value for money) which considered coordination within RCM and non-RCM coordination, e.g. with Oxfam.

Box 6: There is experience of evaluating coherence in evaluations of responses to natural disasters: Example of Indian Ocean tsunami 2004

In 2005, Abhijit Bhattacharjee carried out a real-time evaluation for IFRC of the tsunami response in Asia and East Africa. This does not have an explicit focus on coherence as an evaluation criterion but one of its chapters focuses on transition from relief to recovery phase and in that includes a section on coordination. This covered a number of issues including particularly the movement coordination framework and the need to coordinate multiple participating national societies.

In 2006, John Telford and his colleagues produced a synthesis report for the Tsunami Evaluation Coalition based on the joint evaluation of the international response to the Indian Ocean tsunami. While this report did not explicitly analyse coherence as an evaluation criterion, it was a central part of some of the recommendations, for example that all actors should strive to increase their disaster response capacities and to improve the linkages and coherence between themselves and other actors in the international disaster response system, including those from the affected countries themselves.

In 2006, ALNAP convened a meeting in Rome to consider lessons learned from the joint evaluation of response to the Indian Ocean tsunami in order to set the agenda for joint evaluations of humanitarian assistance. One of its conclusions was that joint evaluation may be more suitable for addressing issues of coordination and coherence than evaluations conducted by single agencies.

Box 7: There is experience of evaluating coherence in evaluations of responses to natural disasters: Example of hurricanes and typhoons

In 2014, Teresa Hanley and others produced a report of the IASC Inter-agency Humanitarian Evaluation of the Typhoon Haiyan response. This evaluation did not systematically examine coherence as a criterion. Rather, the evaluation was focused largely on whether articulated results were achieved with some discussion of relevance and timeliness. There were also sections on (i) how well the international response engaged with and strengthened national and local systems, structures and actors for disaster response and (ii) coordination.

Also that year, ICAI conducted a rapid review of the DFID's humanitarian response to typhoon Haiyan in the Philippines. Although the report did not specifically mention consideration of coherence, the summary did highlight under the area of mobilisation that DFID had worked well the Ministry of Defence and other UK Government departments. It also included a specific area on transition, for example to longer-term needs and livelihoods.

In 2015, Itad also evaluated DFID's humanitarian response to typhoon Haiyan (Yolanda). This did not explicitly focus on coherence as an evaluation criterion. Findings were presented in three areas – effectiveness of the DFID response; DFID's contribution to the effectiveness of the humanitarian system and improving DFID and partner approaches to accountability to affected populations. Nevertheless, some of the findings and lessons learned did relate to coherence, for example lesson six was that DFID's influence on the transformative agenda Level 3 response was less than it could have been, in part due to lack of coherence and linkages between field-level monitoring activities and strategic areas DFID would like to influence, such as violence against women and girls, value for money and accountability.

In 2016, Philip Dy and Tori Stephens produced a discussion paper for the Harvard Kennedy School's Program on Crisis Leadership based on the response to typhoon Haiyan focusing on strengthening coordination among Philippine government, civil society and international actors. The discussion paper was based on a case analysis method relying on secondary data and key informant interviews. While the paper does not discuss coherence explicitly, the topic of coordination is relevant. It explained the coordination challenges encountered in three ways:
* While the Philippines has institutionalized approaches for disaster management in laws and policies, government at all levels has varying capacity for implementation
* Both government and the international community failed to adequately partner with civil society, resulting in missed collaboration opportunities
* The Philippine government's operational logic clashed with that of the international system in four domains – locus of control; accountability and pressure to deliver; timelines for relief and recovery; coordination processes. These clashes inhibited effective coordination

In 2017, CRID produced a summative evaluation report focused on CARE Philippines' programme in response to Typhoon Haiyan. This evaluation did not explicitly consider coherence as one of the evaluation criteria it assessed but it did review the programme's partnership strategy. The questions here were:
* How well did the Program contribute to the development of partnerships?
* What can be learned about how to increase private sector participation from the Program?

This material was very focused on programme effectiveness, that is what are the best partnerships to enable CARE to reach its target results.

Also that year, IOM issued a situation report focused on the Caribbean and the response to hurricanes Irma, Maria and Jose. This did not explicitly discuss coherence. In 2019, Gavin White produced a final evaluation report of the IFRC hurricane Irma response operation in Antigua and Barbuda and St Kitts and Nevis. This also did not consider coherence explicitly. The evaluation report was divided by country and focused on support provided, health programming, capacity building etc. Under operational management, there was consideration of coordination with national societies, movement coordination and IFRC internal coordination. Conclusions were framed in terms of other evaluation criteria.

- Financial crises (ADB, 2020a; ILO, 2020a; World Bank Independent Evaluation Group, 2020) (see Box 8).

<div style="border:1px solid; background:#ffffcc; padding:10px;">

Box 8: There is some experience of evaluating coherence in evaluations of responses to financial crises

In 2020, ADB sought to document lessons learned from previous support to micro, small and medium-sized enterprises. However, this does not consider coherence explicitly. There was one relevant finding from Viet Nam which identified that a number of agencies were involved in financial reform and there was need for continuous coordination between them. Also in 2020, ILO sought to draw lessons from previous economic and financial crises concerning effects on workers and enterprises. Although coherence was not explicitly identified as an evaluation criterion, a number of the findings etc. related to coherence. For example, lesson six was that international policy coherence had been a driver to mainstreaming decent work at the international, UN and country levels.

In 2020, the World Bank Independent Evaluation Group compiled lessons based on previous evaluations in relation to crisis response and resilience to systemic shocks. This did not explicitly consider coherence but one of the lessons learned did relate to the importance of multipronged policy responses to pandemics.

</div>

- Conflicts, such as in Syria (Sida et al, 2016; Darcy, 2016; WFP, 2018a-b; UNFPA Evaluation Office, 2019) (see Box 9).

<div style="border:1px solid; background:#ffffcc; padding:10px;">

Box 9: There is some experience of evaluating coherence in evaluations of responses to conflicts

In 2016, Lewis Sida and others carried out an evaluation of the OCHA response to the Syria crisis. Coherence was not used as an evaluation criterion. The evaluation was organised around four key areas of enquiry – leadership of the Syria response; enabling the system; strategy, planning and operations; and management. These areas were mapped to OCHA strategic goals which included coordination and OECD DAC evaluation criteria (but not those specifically for humanitarian contexts, such as coherence). None of the identified main sub-questions touched on coherence although some issues of coherence were identified in the findings.

Also, that year, James Darcy produced an evaluation synthesis and gap analysis of 24 publicly-available evaluative studies concerning the international response to the Syria crisis. This report was structured around seven thematic clusters. The second of these – strategy and planning, coordination and leadership (also called preparedness, strategy, coordination and leadership)– specifically considered coherence of United Nations-led strategy which it termed strategic alignment. In addition, other elements related to coherence were considered in other thematic clusters. For example, working in partnership with government was considered in the first thematic cluster of context-related findings. Findings related to coherence are mainly covered under the second thematic cluster. The guiding questions under this thematic area were:
- To what extent are weaknesses in strategy and planning identified in the findings as lying behind under-performance?
- How well has the overall response been led by the United Nations?
- What picture of inter-agency coordination emerges from the material? And what picture of coordination with governments?

In 2018, WFP produced two evaluation reports related to the Syrian crisis. The first was an evaluation of the WFP regional response to the Syrian crisis between 2015 and 2018. This was structured around seven evaluation criteria including complementarity and coherence. These criteria were used implicitly, to structure the evaluation's findings into four sections around four questions, and explicitly to present conclusions in a table. The evaluation had four main questions, each with two to six sub-questions. The four main questions were matched to evaluation criteria. The questions matched to coherence and complementarity were:
- To what extent was WFP response well aligned with national/regional responses to the crisis?
- How effectively has WFP engaged with collective decision-making within the United Nations system to promote a principled and coherent approach to the humanitarian response?
- To what extent has WFP seized opportunities for joint implementation/collective operational action within the humanitarian response?
- To what extent has the WFP choice of a regional "Syria + 5" model for its humanitarian response supported synergies across countries/programmes?

Second, WFP conducted a decentralised evaluation of its general food assistance to Syrian refugees in Jordan from 2015 to mid-2018. The evaluation had four key evaluation questions including how relevant, appropriate and coherent is the general food assistance (worded elsewhere as "*is the design of the general food assistance activity relevant to the context and contributing to a larger social safety net environment*"). Findings were presented according to these questions. Sub-sections in response to this question included engagement and coordination within the UN system.

</div>

Box 9: There is some experience of evaluating coherence in evaluations of responses to conflicts (continued)

In 2019, UNFPA published a report of an evaluation of their response to the crisis from 2011 to 2018. Findings were structured around ten evaluation questions/criteria including coordination (and leadership within the humanitarian response architecture), coherence (with UNFPA strategic frameworks and with the strategic and normative frameworks of the wider humanitarian system), connectedness (of humanitarian action with longer-term development strategies and processes [the humanitarian-development nexus]) and partnerships. There were more detailed sub-questions under these main questions/criteria namely:

Coordination
- To what extent has the formal leadership of the gender-based violence area of responsibility (at international, hub and country levels) and informal leadership of reproductive health working groups and youth working groups (at hub and country levels) by UNFPA contributed to an improved sexual and reproductive health and rights, gender-based violence and youth-inclusive response?

Coherence
- To what extent is the UNFPA response aligned with: (i) the priorities of the wider humanitarian system (as set out in successive humanitarian response plans and regional refugee and resilience plans); (ii) UNFPA strategic frameworks; (iii) gender equality principles of the United Nations Evaluation Group; (iv) national-level host government prioritisation; and (v) strategic interventions of other United Nations agencies?

Connectedness
- To what extent does the UNFPA response promote the humanitarian-development nexus?

Partnerships
- To what extent does UNFPA leverage strategic partnerships within its response?

14. In addition, the European Union, for many years, published 3C reviews of coordination, coherence and complementarity (see Box 10 and, for example, European Union, 2007[5]). However, they stopped doing these reviews in 2008.

<div style="border:1px solid #000; background-color:#ffffcc; padding:10px;">

Box 10: The 3Cs initiative

This initiative involved a series of joint evaluations with a view to explore and assess the role played by the Maastricht Treaty precepts of coordination, complementarity and coherence in the European Union's development co-operation policies and operations and to determine how far these have been applied in practice and with what impact.

The 3Cs were defined by the European Union's Heads of Evaluation Task Force in 2003.
- *Coordination* was defined as 'activities of two or more development partners that are intended to mobilise aid resources or to harmonise their policies, programmes, procedures and practices so as to maximise the development effectiveness of aid resources'. With regard to co-ordination several levels (international, regional, national, sub-national, sectoral) can be distinguished, as well as differences in content (policies/principles/priorities, procedures, practices) as in intensity (consultation, co-operation, collaboration). Co-ordination is seen as necessary, because a lack of co-ordination could lead to: a donor driven agenda, excessive demands on scarce management capacities, inconsistencies of approach, etc.

- *Complementarity* is intended to ensure that Community development policy 'shall be complementary to the policies pursued by the Member States'. This indicates that development co-operation is a shared competence between the Community and the Member States which can be jointly exercised. It is confirmed that the Community has a specific, but not exclusive competence in the field of development co-operation. In this sense complementarity differs from the concept of 'subsidiarity', which refers to a distribution of competence and decision-making at the most appropriate level. In the case of complementarity both the Commission and the Member States can have competences and tasks at the same level. The notion of complementarity poses the question of its direction, in other words, is it up to the Community to complement the activities of Member States, or the other way around? Another issue is the equal partnership between the Commission and Member States, and reciprocal participation in the elaboration of their respective policies.

- *Coherence* was considered to be the most debated of the 3Cs and was defined as: 'the non-occurrence of effects of policy that are contrary to the intended results or aims of policy.' Much was considered to depend on the perspective of the viewer. For example:
  - A narrow definition would be that objectives of policy in a particular field may not be undermined or obstructed by actions or activities in this same field
  - A wide definition would be that objectives of policy in a particular field may not be undermined or obstructed by actions or activities of government in that field or in other policy fields

  With regard to policy coherence this means that it can focus on one terrain or field of policy only, or try to make links with other fields, domains or policies. This can lead to a range of consequences:

  - Incoherence in European development policy itself
  - Incoherence between different sets or parts of foreign policy and development co-operation policy
  - Incoherence between development co-operation policies and policies in other fields, which can in theory be all parts of European policy making

  An important aspect is the distinction between intended and unintended incoherence in policy-making. This stresses that there is no hierarchy in policies and that given a certain set of goals and weighing them against a set of goals in another policy field, incoherence can also be deliberate.

A total of eight evaluation reports were published from 2004 to 2008. Each of these covered a specific subject or field but each also asked a number of common process-oriented questions, namely:

- What steps did relevant actors take, individually or jointly, to improve coherence, complementarity and/or coordination?
- Which, if any, enabling mechanisms and/or frameworks were used or put into place? And for what purpose?
- What results were achieved, intentionally or unintentionally? Why, and why not?
- What constraints or opportunities did the actors encounter while implementing their actions/mechanisms? How did they deal with these?

</div>

---

[5] This cites the website on http://www.three-cs.net/index.html

15. Within the development sector, much of the focus on evaluation of coherence has been on policy coherence for development (PCD) (for example OECD, 2003; Lockhart 2005; Picciotto, 2005; OECD 2008; Barry et al, 2010; Keijzer and Oppewal, 2012; Fellesson and Román, 2016; Núñez-Borja et al, 2018; Stave et al, 2018; European Commission 2019; OECD 2019; Koff et al, 2020; European Commission, undated; Select Committee on International Development, undated) (see Box 11).

Box 11: Evaluations of policy coherence for development: Examples of background material

In 2003, <u>OECD</u> produced a policy brief on policy coherence as vital for global development. It defined policy coherence for development by stating:

> *"Policy coherence for development means taking account of the needs and interests of developing countries in the evolution of the global economy. It involves the systematic promotion of mutually reinforcing policy actions across government departments and agencies creating synergies towards achieving the agreed objectives, such as promoting knowledge-based economies in poor countries through the appropriate use of information and communication technology"*

It explained why it was important but difficult to achieve and how development countries benefit. It also identified key areas for making progress, namely agriculture, trade policy, investment, knowledge and technology transfer, migration and global policy action, and discussed the role of the OECD. It did not explicitly identify questions that could be asked to assess policy coherence for development.

In 2005, <u>Clare Lockhart</u> produced a paper for ODI for the Senior Level Forum on Development Effectiveness which explored the shift from aid effectiveness to development effectiveness and the importance of strategy and policy coherence in fragile states. It discussed what policy coherence is, issues relating to strategy and policy coherence in a fragile state context and emerging approaches to policy coherence. However, it did not explicitly identify questions that could be asked to assess policy coherence for development. That same year, <u>Robert Picciotto</u> published an article in Evaluation on the evaluation of policy coherence for development. This identified four elements of policy coherence for development – internal coherence, intra-country coherence, inter-country coherence and donor-recipient coherence. It explained the rationale for policy coherence and then discussed this in relation to the European Union, the United Nations, International Financial Institutions, non-governmental organisations and OECD. It discussed policy coherence for development monitoring by UNDP and the World Bank, and particularly focuses on the Commitment to Development Index produced by the Center for Global Development. The paper also discusses the role of evaluability assessments and then a number of dimensions of policy coherence for development, such as accountability and social learning. It ends with sections on:

- *Evaluation options* arguing for qualitative approaches and triangulation techniques pioneered outside the development evaluation profession. Specific options considered include:
  - Systematic independent multi-donor evaluations of international collaborative multi-country development programs currently in place to deliver global public goods, share knowledge across countries, or set business and/or professional standards.
  - Vertical multi-country reviews of individual policies (aid, trade, migration, etc.) on a regional or global basis.
  - Systematic assessments of the PCD aspects of national policies on a horizontal basis.
  - A sample of representative developing countries within which the development impact of specific reforms in OECD policies (e.g. increase in quality or quantity of aid, gradual removal of cotton subsidies, reduced immigration restrictions, etc.) would be tracked and recommendations made both to OECD and to the countries concerned in order to enhance the synergy of policy reforms and to improve the design of aid programs.
  - Independent assessments of the impact of regulatory regimes and standards (whether voluntary or compulsory) on developing countries.
  - An annual progress report on PCD that would build on the CDI index but enhance its legitimacy and evaluative content through formal participation by governments, the private sector and civil society.

- *Evaluation methods* including program evaluation theory, meta-evaluation methods with theory-based evaluation techniques, case study and policy research tools, 'new public management' tools, objectives-based approach, evaluation of partnerships and process evaluations.

- *Evaluation governance* to ensure independence, objectivity and value added arguing for a combination of self- and independent evaluation. Principles identified included:
  - An arm's length relationship with line managers and policy-makers including giving the evaluation team considerable autonomy
  - Capacity to influence policy formulation and decision-making
  - Compliance with principles of accountability, learning and transparency
  - Involving major stakeholders in evaluation design
  - Being clear of respective roles and responsibilities
  - Providing adequate skills and resources for the evaluation and its dissemination
  - Involvement of developing countries in the process

However, the paper did not explicitly identify questions that could be asked to assess policy coherence for development

Box 11: Evaluations of policy coherence for development: Examples of background material (continued)

In 2008, OCED published a policy brief on the lessons learned for policy coherence on development. In addition to identifying key lessons relating to policy coherence for development, it argued that having an assessment methodology "*would help to identify the major synergies, conflicts or trade-offs across several domains that contribute to development (economic, environmental and social). It would help compare the positive and negative impacts on the different dimensions and to tease out potential conflicts in order to achieve more coherent policies towards development*". However, it did not make explicit suggestions for such an assessment methodology.

In 2010, Frank Barry and others published a paper which identified five challenges for policy coherence for development based in experience of EU and Irish policies. These challenges were the opposing interests of domestic and development constituencies, conflicts between development objectives themselves, disagreements between experts on what 'good' development policy looks like, difficulties in identifying the true development interest of developing countries and the growing heterogeneity between and within developing countries. In 2012, Neils Keijzer and Jorrit Oppewal produced a paper for the European Centre for Development Policy Management which reviewed methodological approaches for evaluating coherence in the field of international cooperation. This study, based on a review of 22 studies, focused on four research questions:
- In what ways and to what extent can 'coherence' be defined and operationalised for evaluation purposes?
- To what extent can the relation between coherence and effectiveness/efficiency be evaluated (i.e. is coherence additional or complementary?
- What methods have been used in past studies and evaluations that look into coherence inside or between policies, at what levels (micro, meso, macro), and what are their respective strengths and weaknesses?
- Based on the answers to the first three questions, what practical and methodological dilemmas can be observed with regard to improving the evaluation of policy coherence in the specific field of policies on international cooperation?

The paper concluded that there was, at that time, no widely-accepted definition of coherence. It also observed that the studies most commonly used were interviews, document analysis and descriptive statistics. It noted that, "*analysing the respective strengths and weaknesses of different methods was challenging due to two reasons. First of all, few if any studies offered much reflection on the benefits and limitations of their methodological approaches. Secondly, methods are not intrinsically useful as the usefulness depends on the way in which the method was applied*".

In 2016, Måns Felleson and Lisa Román produced a report on results and responsibility for the Swedish Policy for Global Development. The report considered policy coherence in terms of motivation, coordination, cognition, voluntary responsibility and whether there is a common understanding. It considers how policy coherence for development fits with a results-based management approach and considers the examples of migration and development and higher education and research. In the concluding discussion, the report considers whether aid is a catalyst for coherence.

In 2018, Carmen Núñez-Borja and others conducted an external evaluation of the European Union's policy coherence for development covering the period from 2009 to 2016. This was based on a conceptual framework which recognised policy coherence for development as an evolving concept while recognising that the European Union had a particular approach to this. To address this the evaluation team reconstructed an intervention logic. The evaluation sought to answer eight evaluation questions mapped to various evaluation criteria including coherence and these were used to frame presentation of the findings. All questions are shown here (as they all relate to policy coherence for development). The criteria to which the questions are considered to refer are shown in brackets.
- To what extent has the EU PCD approach and its operational framework responded to evolving needs? (relevance)
- To what extent has the EU PCD approach been aligned with wider EU policy and evolving international obligations of the EU? To what extent have PCD inputs and activities been adequate to implement the EU PCD approach? (coherence)
- To what extent has the EU PCD approach (PCD specific mechanisms) led to raised awareness on PCD, which in turn has indirectly influenced policy-making? (efficiency)
- To what extent has the EU PCD approach influenced existing or planned policies/initiatives likely to affect developing countries so that they take into account development objectives? (effectiveness/efficiency)
- To what extent has the EU PCD approach created additional value beyond what could be achieved by EU Member States acting independently? (EU added value)
- To what extent have changes in the design and implementation of EU policies and initiatives brought about by incorporating a PCD approach influenced outcomes and impacts in developing countries? (impact)
- To what extent is the EU PCD approach sustainable? (sustainability)

16. OECD governments have conducted reviews and evaluations particularly focused on policy coherence for development. These include Austria (Austrian Development Agency, 2021); the EU (European Commission, 2018); Finland (Zetter et al, 2019); Ireland (Barry et al, 2010); the Netherlands (Ministry of Foreign Affairs Netherlands, 2019); Norway (Stave et al, 2018); Sweden (Fellesson and Román, 2016; Government Offices of Sweden, 2020); and the UK (Select Committee on International Development, undated) (see Box 12). Policy Coherence for Development is one of the topics that can be covered in OECD peer reviews of bilateral donors' development cooperation (for examples - see OECD 2017; OECD, 2020d; OECD, 2020e) (see Box 13).

Box 12: Evaluations of policy coherence for development: Examples from OECD governments

In January 2021, the Austrian Government started a project to conduct an evaluation and evidence synthesis of its whole-of - government approach. The material identified is fairly brief and might be considered a summary terms of reference. It does link the concepts of whole-of-government and policy coherence noting that the latter term is perhaps broader. The material identified does not include specific questions that the evaluation will answer.

In 2018, the European Commission published an evaluation of its generalised scheme of preferences. This included six specific questions including to what extent the current generalised scheme of preferences is coherent with EU's relevant policies. The report is structured around particular methods, e.g. economic impact analysis, social and human rights impact analysis and environmental impact analysis. It also contains thematic (textile and machinery sectors) and country case studies (Bangladesh, Ethiopia, Bolivia and Pakistan). Attempts are made to systematically answer the evaluation questions, including the question on coherence, in the conclusions section.

In 2019, Roger Zetter and others conducted an evaluation on forced displacement and Finnish development policy. This evaluation seeks to assess "*how coherently [Finland's] development policy and its targets relating to forced displacement have been implemented and how the coherence could be enhanced*". There were three main evaluation questions and the third of these - "*to what extent and how do the approaches to forced displacement and humanitarian-development nexus rooted in the Development Policy Programme's help establish policy coherence between Finnish policies?*" – focuses on coherence. Rather than having sub-questions on coherence, this evaluation identifies three judgement criteria and these are systematically assessed (in an annex). They are:

- ***Mechanisms to promote policy coherence within the Ministry for Foreign Affairs are in place and operate effectively***
- ***There is coherence between relevant Ministry for Foreign Affairs policies on forced displacement and humanitarian-development nexus and those of other Government Ministries/Departments (e.g. Ministry of the Interior, Ministry of Defence, Prime Minister's Office) and the Ministry for Foreign Affairs' partners – bilateral and multilateral development co-operation partners (UN, EU and civil society organisations)***
- ***The level of policy coherence achieved is adequate to support the approaches to forced displacement and humanitarian-development nexus***

Issues of coherence are covered in the context analysis, the findings and the conclusions.

In 2019, the Ministry of Foreign Affairs of the Netherlands produced an evaluation of the Dutch Government's policy on responsible business conduct. The evaluation covers four main areas – sector agreements, framework for private-sector instruments, procurement by the government and international initiatives. Coherence is identified as a specific sub-theme in the last two. There is also consideration of Dutch responsible business conduct policy in four countries – Bangladesh, Colombia, Ethiopia and India. The evaluation was based on one main question and seven sub-questions. Two of the sub-questions specifically relate to coherence, namely:

- To what extent is the policy internally coherent (e.g. between activities under the heading of international responsible business conduct policy) and well balanced in terms of policy attention?
- To what extent is the policy externally coherent with and influenced by other policies of the Dutch government (not only the Ministry of Foreign Affairs), for example on trade and competition?

Box 13: Policy coherence for development is assessed in OECD peer reviews: Examples

The 2017 OECD peer review of the Netherlands noted that "*A World to Gain*" and an eight-point action plan reflect the Netherlands' commitment to policy coherence for development. The review concluded that coordination among ministries ensures coherent policies and that the Netherlands' progress on policy coherence was impressive. However, it also observed that better monitoring and a clear timeline were needed.

In 2020, the peer review of Ireland noted that the country was making serious efforts to address challenges to policy coherence for sustainable development. However, there was need to make the commitment to policy coherence for sustainable development more explicit. It was also suggested that Ireland could establish a specific mechanism for monitoring and assessing the transboundary effects of domestic policies and could also seize opportunities for a structured cross-government approach.

In 2020, the peer review of Austria did not include a category of policy coherence for sustainable development. However, under structure and systems, the review pointed out that the fragmented nature of Austria's development cooperation system presented challenges for coordination and a more coherent, whole-of-government approach was suggested. Under delivery and partnerships, the review commented that including all Austrian contributions in country strategies would help ensure a more coherent and coordinated approach. Under fragility, crises and humanitarian aid, the review proposed that evaluating Austria's humanitarian system would be a good first step in updating its humanitarian strategy and making humanitarian assistance coherent with development cooperation and peace building. Recommendations included that the Ministry of Foreign Affairs should be mandated to lead a coherent and coordinated approach to its total development cooperation effort, Austria should empower and resource a focal point or institution to lead on policy coherence for development, including responsibility for developing an action plan to address key areas of incoherence and Austria should update its humanitarian policy to better reflect a pragmatic and coherent approach to engaging in protracted crises. .
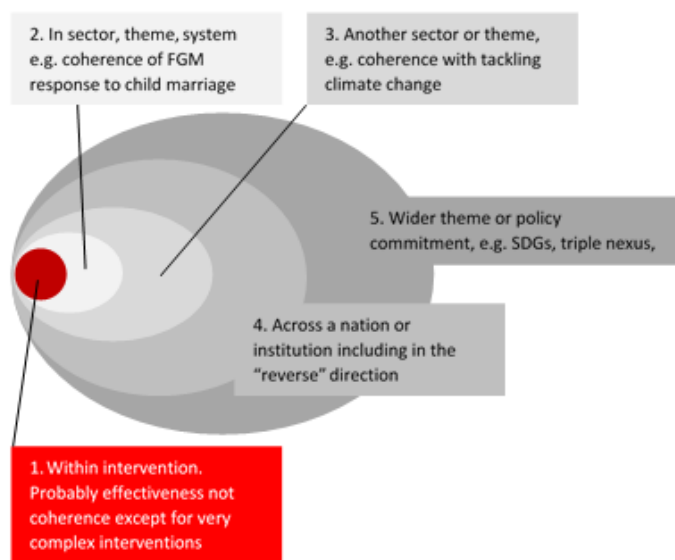
Which coherence issues are raised in these evaluations?

17. Based on this material, there is a high degree of heterogeneity in relation to what kind of coherence and which dimensions of it are being assessed where coherence is addressed. Many evaluations do not consider the issue of coherence. This is not simply a case that evaluations took place before 2019 when OECD adopted coherence as one of its criteria for development evaluations. Although coherence has been long-established as a criterion for humanitarian evaluations (see Box 1), a recent discussion paper based on a review of multiple humanitarian evaluations concluded that only around one third of evaluations considered the criterion of coherence while more than 80% of those evaluations considered the criterion of effectiveness (Darcy and Dillon, 2020). Five different forms of coherence or fit have been identified in this review. These are illustrated in Figure 2 and are briefly explained here.

18. There is coherence within the intervention and questions on this type of "*coherence*" might focus on how coordination mechanisms and/or partnerships are enabling progress of the intervention, e.g. towards achieving its objectives or producing expected results. Such questions reflect an instrumental approach and, as explained in Table 1, such an approach fits better under the OECD DAC criterion of effectiveness rather than under coherence. In such cases, the issue of interest is not how well the intervention fits with other things but rather how partnerships and coordination mechanisms can maximise an intervention's effectiveness. This type of coherence is labelled number 1 in Figure 2 and is denoted by the red circle and the red text box. An example is provided in Box 7 of the CARE Philippines' programme in response to typhoon Haiyan. Assessment of the programme's partnership strategy was focused on how best these partnerships enabled CARE to reach its target results. So, this assessment was more about programme effectiveness than an assessment of how what CARE was doing fitted with what others were doing which would have been an assessment of external coherence.

19. There may be interest in how an intervention on a particular topic or theme fits with interventions on similar topics or themes within a given sector or area. An example of this is the evaluation looking at how an intervention on female genital mutilation fits with interventions in other linked areas, such as early marriage and gender-based violence (UNICEF and UNFPA, 2020). Another example is the ALNAP Lessons Paper on Ebola which looked at similar issues concerning the coherence of Ebola response with other ongoing healthcare provision (Lamoure and Juillard, 2020).

Figure 2: Which coherence issues are raised in these evaluations?

2. In sector, theme, system e.g. coherence of FGM response to child marriage

3. Another sector or theme, e.g. coherence with tackling climate change

5. Wider theme or policy commitment, e.g. SDGs, triple nexus,

4. Across a nation or institution including in the "reverse" direction

1. Within intervention. Probably effectiveness not coherence except for very complex interventions

20. A particular type of this approach might be termed a systemwide evaluation where elements of coherence are considered across a system, for example, the humanitarian system in a given

country (e.g. Burnett, 2019). This approach is not looking at coherence with particular parts of a sector or system but coherence across the entirety of that system.

21. There may be interest in evaluating how an intervention fits with interventions in another sector. For example, this might include looking at how a health programme fits with policies and/or other interventions on the environment or on the economy.

22. There may be interest in evaluating not just how an intervention fits with one intervention in another sector but rather taking a more holistic approach and looking at how a particular intervention fits with the totality of what a nation or institution is doing, e.g. on development. This is much of the focus of work on policy coherence for sustainable development and taking a whole-of-government approach to development, i.e. considering the effects of non-aid policies (including foreign policy) and interventions on development (see Boxes 11 and 12). Currently, there is also a great deal of interest in the "*reverse*" of this, that is how does a country's aid policy fit with its own domestic policies and needs, e.g. what are the benefits to a donor country of its official development assistance? This is a form of coherence and, for the purpose of this paper, it is referred to as "*reverse coherence*" as it is the reverse of what is usually meant by coherence particularly within development circles. No evaluations looking at this form of coherence were identified in this review.

23. There is an even broader view of coherence where the focus is not only on the whole of what one institution or country is doing but rather on a wider theme or commitment, such as the Sustainable Development Goals (Office of Internal Oversight Inspection and Evaluation Division, 2019; Ishida, 2020 and see Box 14). How humanitarian interventions fit with responses focused on development and peace, might also be an example of this type of coherence (Caparini and Reagan, 2019). Some agencies may wish to assess coherence with a broader set of values, e.g. feminist values.

Box 14: Using the coherence criterion to understand better interdependencies between different SDG targets and to consider this explicitly when designing and implementing projects focused on training teachers

In 2020, Yoko Ishida published a paper which looked at how the coherence criterion could be used to contribute to achieving the SDG4.c target for teachers. The research sought to answer the following three questions:

- Which of the other SDG targets positively contributes to the achievement of target 4.c?
- How are they taken into consideration when evaluating international cooperation projects for teacher professional development?
- What might be improved by introducing the concept of coherence in evaluating international cooperation projects for teacher professional development?

The research involved taking the International Council for Science's seven-point framework for classifying goals and targets to identify links between particular SDG targets and then conducting a meta-evaluation of all 21 ex-post evaluations of the teacher professional development projects by the Japan International Cooperation Agency from 2008 to 2018. These evaluations were all conducted prior to the inclusion of coherence as one of the OECD DAC evaluation criteria. The research identified at least 27 SDG targets that have positive interactions with SDG target 4.c for teachers. The meta evaluation element showed that few, if any, of the evaluations systematically examined coherence with SDG targets. The research concludes that "*The concept of coherence should be utilized to ensure that key issues, which might have synergistic or trade-off interactions with the project achievements, are to be assessed in evaluation.*"

---

24. How do the coherence issues identified here fit with or differ from the OECD DAC definition of coherence as an evaluation criterion? The issues are identified based on empirical data and observations drawn from the evaluations and other documents reviewed. Most of these evaluations and documents precede the adoption of the OECD DAC's adoption of the coherence criterion for development evaluations. However, a coherence criterion was in place prior to this for humanitarian evaluations, although the scope of this was narrower, as there were also other criteria at that time which are now included under coherence, such as connectedness and

complementarity. A considerable amount of the information included in this section was not originally explicitly identified or labelled as relating to coherence but it has been included because it fits with the way coherence is now defined and understood by OECD. In addition, there may be some material which was identified as relating to coherence but now falls outside the OECD DAC definition.

25. As a result, the coherence issues identified here are consistent with and build on the OECD DAC definition of coherence as an evaluation criterion. This is not because that definition was used by the evaluations in question but because that definition has been used as the filter through which material on coherence and related topics have been identified in this study.

26. There is one type of activity which might be identified as, or appear to be related to, coherence but which this review excludes from coherence as understood from the OECD DAC definition (see paragraph 18). That is where an agency focuses on coordinating or partnering with others not to ensure compatibility of their intervention with other interventions but to maximise the achievement of their own objectives or targets. As explained in Table 1 (p3), such coordination or partnership is seen as instrumental in nature, relating to the criterion of effectiveness and not to the criterion of coherence.

27. This section picks up on the OECD DAC definition of coherence which considers the compatibility of an intervention with other interventions in a country, sector or institution. Table 2 seeks to identify how the coherence issues identified here fit with those parts of the definition. Issues 2 and 3 in this section relate to how an intervention might be coherent with interventions in a sector and distinguish between coherence with interventions in the same sector and with interventions in another sector. Issue 4 in this section relates to how an intervention might fit with the totality of what a nation (country) or institution is doing. Issue 5 in this section relates to how an intervention might fit with a broader theme or commitment, such as the SDGs. This perhaps does not fit well with these elements of the OECD definition, i.e. compatibility with other interventions in a country, sector or institution.

Table 2: How do the coherence issues identified here fit with the OECD DAC definition of coherence

	Issue identified in this section	Compatibility of the intervention with other interventions in a…			Type of coherence	
		Country	Sector	Institution	Internal	External
1	Partnerships or coordination to maximise an intervention's results	This does not meet the OECD DAC definition of coherence and should be considered under effectiveness				
2	Compatibility of the intervention with other interventions in the same sector		✔		✔	✔
3	Compatibility of the intervention with other interventions in the same sector		✔		✔	✔
4	Compatibility of the intervention with the totality of what a	✔		✔	✔ ✔	

	nation or institution is doing					
5	Compatibility of the intervention with a broader theme or commitment, such as the SDGs				✔	✔

28. Table 2 also shows how the issues of coherence identified here also fit with the split in the definition into internal and external coherence. In most cases, the issue might relate to either internal or external coherence. For example, if consideration is being given to the compatibility of an intervention with other interventions in the same sector, this could relate to other interventions implemented by the same agency (internal) or by another agency or agencies (external). In the case of compatibility with the totality of what a nation or institution is doing, this is fundamentally internal coherence. But, this does depend (see paragraph 11) on precisely how an institution is defined and if that definition is shared and understood. For example, when the Ministry of Foreign Affairs of the Netherlands evaluated the Dutch government's policy on responsible business conduct (see Box 12), it defined internal coherence in terms of activities under the heading of international responsible business conduct policy while it defined external coherence in terms of compatibility with other policies of the Dutch government, that is outside the Ministry of Foreign Affairs. This means that the institution/entity is being defined as the Ministry of Foreign Affairs. If however, the institution/entity is defined as the Dutch government, both these forms of coherence would be internal.

Ways of working and subjects for evaluation[6]

29. Within this material, there are different ways of working and various subjects for evaluation and these are illustrated in Table 3. One specific way of working has been through joint or inter-agency evaluations, for example following the Indian Ocean Tsunami (Telford et al, 2006) and of particular themes (UNICEF and UNFPA, 2020)

30. Subjects for evaluation have included particular donor countries or institutions, particular themes or topics and a particular country or region. However, most of the country evaluations[7] are not really evaluations of a country's response to a particular issue, such as malaria or girls' education. Rather, they mostly view countries as the location in which a particular intervention occurs (see Box 15). Consequently, many of the evaluations combine each of these elements, i.e. an evaluation of an agency's intervention on a particular theme in a particular country, such as UNICEF's response to cholera in Yemen (UNICEF, 2018). This evaluation is cited as one example to illustrate the observed tendency of humanitarian evaluations to focus on *context-specific crisis responses by individual agencies* (Darcy and Dillon, 2000).[8]

---

[7] See final row of Table 3.

Table 3: What were the subjects of evaluations that included consideration of coherence before the COVID-19 pandemic?

Subjects of Evaluation	Examples
Institution or country (donor)	• Ireland (Barry et al, 2010; OECD, 2020d); Norway (Stave et al, 2018); Sweden (Fellesson and Román, 2016; Government Offices of Sweden, 2020); and the UK (Select Committee on International Development, undated) • Multilaterals – ILO (ILO, 2020a); IOM (IOM, 2017; IOM, undated); UNDP (UNDP, 2020a-d; UNDP 2021); UNFPA (UNFPA Evaluation Office, 2019; UNICEF and UNFPA, 2020; UNFPA Evaluation Office, 2020b); UNICEF (Dara, 2016; UNICEF, 2017; UNICEF, 2018; UNICEF and UNFPA, 2020); WFP (WFP, 2017; WFP, 2018a-b; WFP, 2019; WFP, 2020c-d); and the World Bank (World Bank IEG, 2020; Gold and Hutton, 2020; World Bank Independent Evaluation Group, 2020) • Red Cross (Bhattacharjee, 2005; EPYPSA, 2011; Thormar, 2013; Key Aid Consulting, 2018; White, 2019) and CSOs, e.g. ActionAid (Alam and Balthazar, 2011); DEC (Momoh et al, 2016); and Islamic Relief (Bhattacharjee, 2017)
Theme or sector	For example support to enterprise (ADB, 2020a); social protection (UNDP, 2020b) (incl cash transfers (COVID-19 Global Evaluation Coalition, 2020c)); livelihoods (UNDP, 2020c); health (UNDP, 2020d) (incl AMR (Government Offices of Sweden, 2020), cholera (UNICEF, 2018)); formal private sector (World Bank IEG, 2020); food security (OECD and ECDPM, 2013; COVID-19 Global Evaluation Coalition, 2020a); gender equality (incl in education) (COVID-19 Global Evaluation Coalition, 2020b,d); natural resource management (Hodge-Mitchell et al, 2014); forced displacement (Zetter et al, 2019); FGM (UNICEF and UNFPA, 2020) and shelter (EPYPSA, 2011)
Country or region	Afghanistan (Anderson, 2016; Zetter et al, 2019); the Caribbean (IOM, 2017; White, 2019); Guinea (ECDC, 2017); Haiti (Grünewald et al, 2010; Patrick, 2011; Alam and Balthazar, 2011; Bhattacharjee and Lossio, 2011; EPYPSA, 2011; The Humanitarian Coalition, 2012), Iran (Hodge-Mitchell et al, 2014); Nepal (Baker et al, 2015; Murtaza et al, 2016; Dara, 2016; Bhattacharjee, 2017; Key Aid Consulting, 2018); Nigeria (World Bank, 2015; WFP, 2019); Philippines (Hanley et al, 2014; ICAI, 2014; Itad, 2015; Dy and Stephens, 2016; CRID, 2017); Syria (Sida et al, 2016; WFP, 2018a-b; Zetter et al, 2019; UNFPA Evaluation Office, 2019); Uganda (Thormar, 2013); and Yemen (UNICEF, 2018)

Box 15: Most of the evaluations which specified a particular low or middle-income country conceptualised that country as the location in which the evaluation occurred rather than as a key part of the evaluation's subject

None of the evaluations reviewed were evaluations of an entire national programme on a particular topic but rather they tended to be evaluations of a particular agency's intervention on a theme in that country. This approach tends to treat the country (and its institutions) not as an active participant in the intervention or a key part of the subject of the evaluation but as a rather passive location in which activities happen. This issue is reflected in some of the feedback received which was concerned that while an institution, policy, actor, government, theme or response could be the object of an evaluation (evaluand) a country or region could not be. Rather, these were just locations in which interventions and evaluations take place. This certainly reflects the practice observed here but does not negate the need for more evaluations of entire national programmes on particular topics.

Two evaluations were documented in Afghanistan. One related to Denmark's engagement and integrated approach in the country while the other used the country as a case study of forced displacement and Finnish development policy. The two evaluations in the Caribbean relate respectively to IOM and IFRC engagement following hurricanes. The evaluation in Guinea related to ECDC's Ebola deployment there. The evaluations in Haiti all relate to the 2010 earthquake and concern responses by ActionAid, IFRC, OCHA, the Inter-Agency Standing Committee and the Humanitarian Coalition. Some of the reports, including one for OECD, do attempt to take an overview of the entire response. The evaluation in Iran related to UNDP support to MENARID Iran for integrated natural resource management. The evaluations in Nepal all relate to the 2015 earthquake and concern responses by IFRC, the ACT Alliance, UNICEF and the British Red Cross. The evaluations in Nigeria relate to Ebola and the work of the World Bank and the emergency response in North-East Nigeria by WFP. The evaluations in the Philippines relate to responses to typhoon Haiyan by the Inter-Agency Standing Committee, DFID and CARE Philippines. One report did look specifically at coordination among the Philippine government, civil society and international actors. Evaluations of responses to the crisis in Syria were conducted for OCHA, WFP and as a case study of forced displacement and Finnish development policy. The evaluation in Uganda was of the Ebola response by IFRC. The evaluation in Yemen was of UNICEF's response to cholera.
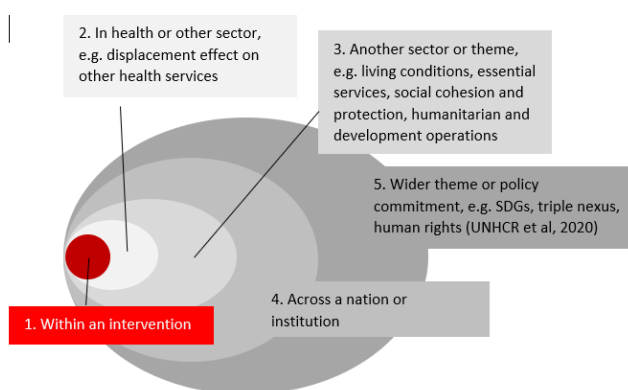
COVID-19 evaluations

31. Initially, many agencies produced guidance on how to conduct evaluations during the COVID-19 crisis (OECD DAC Network on Development Evaluation and the Independent Evaluation Office of UNDP, 2020; WFP, 2020a; Office of Internal Oversight Services Inspection and Evaluation Division, 2020a; UNICEF, 2020a; UNFPA Evaluation Office, 2020a; FAO, 2020; ILO, 2020b; UN Women, 2020; WHO, 2020; UNDP Independent Evaluation Office, undated; UNODC, undated). Much of this focuses on how evaluation practice needed and needs to be modified in the light of the pandemic.

32. The COVID-19 Global Evaluation Coalition (2021c) has conducted an initial analysis of the COVID-19 evaluations based on the plans and documents shared by 46 participants, of which 25 were multilateral and 21 bilateral agencies. As of August 2021, details were received of 679 evaluations. Of these, 67 evaluations (~10%) are planned or are under-consideration solely focused on COVID-19 with a further 123 (~18%) covering questions related to COVID-19. Overall, while 34 (5%) were being conducted jointly, that is by more than one agency, this was the case for 19 (28%) of the 67 evaluations focused solely on COVID-19.[9]

33. In some cases, it was difficult to identify supporting documents for these evaluations. This may be because they were not publically available and/or at an early stage of planning (pre-TOR). Most evaluations with documents available for analysis are ongoing and a small number have been completed.

Which coherence issues are raised in these evaluations?

34. As with evaluations before the onset of COVID-19, there is a high degree of heterogeneity in relation to what kind of coherence is being assessed. Many of the evaluations reviewed do not have a focus on coherence. The same five forms of coherence or fit have been identified in this part of the review. These are illustrated in Figure 3 and are briefly explained here.

35. There may be a focus on coherence within the intervention, i.e. how do coordination mechanisms or partnerships make the intervention more effective? As explained in Table 1 (p3), this is considered an instrumental approach and fits better under the OECD DAC criterion of effectiveness rather than under coherence. This type of coherence is not discussed further here and it is shown in red in Figure 3.

Figure 3: Which coherence issues are raised in these evaluations?

2. In health or other sector, e.g. displacement effect on other health services

3. Another sector or theme, e.g. living conditions, essential services, social cohesion and protection, humanitarian and development operations

5. Wider theme or policy commitment, e.g. SDGs, triple nexus, human rights (UNHCR et al, 2020)

4. Across a nation or institution

1. Within an intervention

---

[9] Caution is needed in interpreting these figures. These are self-reported evaluation plans. Some of these may not be realised and there is wide variety in terms of size and scope. The landscaping exercise is only as complete as the reporting exercise. Agencies that have not yet reported may be planning substantive evaluations but these would not yet be captured in this analysis.

36. There may be interest in how an intervention fits with interventions on similar topics within a given sector or area. Where the intervention is a health intervention, this approach may look at coherence across other elements of health. A good example of this approach to coherence is found in the lesson-learning brief (Avdeenko and Heesemann, 2020) which considers the effects of COVID and responses to it on other aspects of health including immunisation services, maternal care and mental health. Where a COVID-related intervention is not specifically related to health, this form of coherence might relate to another sector. As with evaluations pre-COVID, there may be interest in evaluating coherence across an entire system, e.g. the humanitarian system.

37. There may be interest in evaluating how an intervention fits with interventions in another sector. For example, this might include looking at how a health programme fits with policies and/or other interventions on living conditions, essential services, social cohesion and protection and humanitarian and development operations.

38. There may be interest in evaluating not just how an intervention fits with one intervention in another sector but rather taking a more holistic approach and looking at how a particular intervention fits with the totality of what a nation or institution is doing, e.g. on development. Sometimes such evaluations refer to taking a whole-of-government approach (see paragraph 22). To date, relatively few COVID reviews or evaluations have taken this approach even where they have been focused on a particular organization (see Box 16).

Box 16: Learning from Belgium (ENABEL's) evaluation of their response to the pandemic

The Belgium International Development Agency (Enabel) internal evaluation department commissioned a real-time evaluation (covering the period from March to August 2020) of Enabel's response to the COVID-19 pandemic. The evaluation worked in real time to understand the decision-making mechanisms during the crisis. It aimed to identify what worked well and what needs improvement, what facilitated the response or, on the contrary, what constraints and challenges were encountered and how these were overcome

The evaluation includes a number of findings about the coherence of ENABEL's responses:
For example, it found Enabel's working approach of «dual anchoring» within the country's institutions ensures immediate availability of field staff at central and decentralised levels, and capacity to monitor changing needs of partner countries, guide the response and strengthen socio-economic resilience over time

One example of Enabel's evaluation is highlighted below:

'In DRC and Niger, integration of Enabel's health intervention technical teams within the Ministry of Health facilitates rapid internal decision-making to readjust activities in line with the national response plan. In these two countries the good dynamics existing between Belgian governmental cooperation and its national, regional and local partners, and in particular the special relationship between Enabel and the Ministry of Public Health, have clearly supported the response. The presence of active teams in the field at decentralised level who were rapidly available for the response was a factor in its success. For example in Niger, in the framework of the new intervention co-financed by the EU and Lux-Dev (Team Europe Niger) project teams in the Health Districts (Dosso, Maradi and Zinder), could easily work with the Directors of the targeted hospitals and the managers of the resuscitation services to identify their needs/inventories and to quickly find the most sustainable solutions (for example, the choice of

> tropicalized equipment in small quantities where there is HR to use them, rather than acquiring many respirators in the absence of enough HR to operate them).'p10
>
> The evaluation concludes that 'a strong "OneTeam Belgium" had a positive impact on the coherence and coordination of the response and enhanced visibility of Belgium' p11

39. An example of an agency that has reviewed its approach to international development in the light of COVID-19 is the Scottish Government. This includes substantive consideration of coherence with other areas of policy, including trade, health etc. (see Box 20).

40. Some of the evaluations reviewed do seek to look at how interventions on COVID fit with broader themes or commitments, including the SDGs, the triple nexus and human rights (UNHCR et al, 2020) (see Box 17).

> **Joint Evaluation of the Protection of Rights of Refugees during the COVID-19 pandemic**
>
> The COVID-19 pandemic appears to have challenged the protection of the fundamental rights of refugees in a way that is profound and with possible lasting impacts. Understanding how widespread this is, how effective international cooperation and the combined response of key actors has been, and what we can learn from the steps taken will be crucial to the implementation of current operations and the design of future strategies and plans.
>
> The joint evaluation will be carried under the auspices of the **COVID-19 Global Evaluation Coalition**. The purpose of the evaluation is to examine the effectiveness of international cooperation, and the combined response of host states, agencies and non-state actors, in ensuring the protection of the rights of refugees during the COVID-19 pandemic: to identify emerging good practice, innovation and adaptation to protection responses. Two years since the Global Refugee Forum, a high-level officials meeting (HLOM) is scheduled for December 2021 to review progress towards the objectives of the Global Compact on Refugees and chart the way forward. Progress data and information is being gathered throughout the year to inform this meeting. We are considering different ways to present the findings and discuss the recommendations from this evaluation as a part of this process.
>
> The objectives of the evaluation are as follows:
>
> - To ascertain the **coherence and coverage of refugee rights promotion** and incorporation into international cooperation in the context of national COVID-19 responses;
> - To determine the **effectiveness of the combined contribution of states, agencies and non-state actors** efforts towards enabling refugees to realize their rights in the context of COVID-19.
> - To identify **good practices and lessons** that can be shared for preparedness and application in future emergencies, including a focus on innovation and scalable adaptive solutions
>
> The evaluation is managed by the Evaluation Units of UNHCR, Ministry of Foreign Affairs of Finland, Governments of Colombia and Uganda, and the humanitarian system network ALNAP.

Points about how the issues identified here compare with the OECD DAC definition of coherence are covered in paragraphs 24 to 28.

Ways of working and subjects for evaluation[6]

41. Within this material, there are different ways of working and various subjects for evaluation and these are illustrated in Table 4 and Boxes 20-22. Joint or inter-agency evaluations, particularly of multilateral agencies, are more prominent. Examples include the Inter-Agency Humanitarian Evaluation (IAHE) of the Global COVID-19 Humanitarian Response Plan (GHRP); protection of rights of refugees during COVID-19 (UNHCR et al, 2020) (see Box 17); Response and Recovery Multi-Partner Trust Fund (MPTF) (Freeman et al, 2021) (see Box 18); and possibly a UN systemwide evaluation.[10]

---

[10] Although it is not clear if anything beyond the MPTF evaluation is planned.

Box 18: Evaluation of the UN COVID-19 Response and Recovery Multi-Partner Trust Fund (MPTF)

The inception report for this evaluation presents early lessons and assesses the fund's evaluability. Coherence is mentioned as part of the primary objective of the inception report which was "*to support learning and accountability of the UN COVID-19 MPTF by drawing lessons that are significant in the context of the Resident Coordinator system to improve coherent programming from early lessons; and, to conduct an evaluability assessment of the Fund in order to examine approaches to conduct a final evaluation.*"

The lesson learning component of the inception phase was structured around three areas of investigation. These were presented as positive statements rather than questions to avoid inviting yes/no answers and the risk of shifting the emphasis of the exercise from lesson learning to accountability. These statements were as follows:

- Area of Investigation 1: Progress in the reforms enabled the United Nations Development System to mount a coherent response to the pandemic, including the rapid launch and operation of the Fund and Socio-Economic Response Plan.
- Area of Investigation 2: The Fund and the Socio-Economic Response Plans were designed to take advantage of progress made in United Nations Development System reform and to support and strengthen the reform process
- Area of Investigation 3: Progress on United Nations Development System reforms and the design, governance and management of the Fund and Socio-Economic Response Plans combined to facilitate an integrated, cross-mandate United Nations Development System response to the socio-economic impacts of COVID-19, relevant to the needs of programme countries

The first of these refers to coherence specifically. Despite the reluctance to frame the three areas of investigation as overarching questions, the report does identify a total of 20 questions for early learning distributed across the three areas of investigation. None of these explicitly refer to coherence. The six questions for the first area of investigation are listed here:

- Did the progress and evolution of the MPTF mechanism in response to the Funding Compact assist in rapid and efficient set-up and operation (including disbursement) of the Fund and development of the Socio-Economic Response Plans? How?
- Did the Fund, through the use of the MPTF mechanism and the work of the MPTF Office and the Secretariat on improving transparency, accountability and results reporting elicit a positive response from donors? What may have inhibited resource mobilization and what could be done differently?
- Do the financing mechanism, governance arrangements, consultative processes and approval criteria of the Fund provide value added at a global and country level in comparison with other funding mechanism: including from the perspective of the United Nations Country Team and host governments?
- What constraints may have limited the efficiency of the MPTF mechanism in mobilizing resources, allocating funds to programmes and disbursing funds to participating United Nations Country Team entities? What should be done differently?
- Was progress on strengthening the independence and capacity of the Resident Coordinator and the Resident Coordinator Office advanced enough to enable them to respond quickly to the need for an integrated approach to Socio-Economic Response Plans and Fund supported programme proposals? Was the United Nations Country Team ready for the required level of collaboration and coordination?
- Did the COVID-19 development emergency help to accelerate innovations in use of digital resources to accelerate collaboration, consultation and transparency among key stakeholders? If so, how?

Some of these questions do implicitly appear to relate to coherence, for example, the third question about the value added of the Fund, the fifth question about the readiness of the United Nations Country Team for collaboration and coordination and the sixth question about the use of digital resources to accelerate collaboration, consultation and transparency. In addition, there are other early questions which may relate to the criterion of coherence and some examples are given here:

- Question 8: What linkages and synergies are evident between the fund and other, related MPTFs including, for example the SDG fund, UN Partnership to Promote the Rights of Persons with Disabilities fund and the Spotlight Initiative fund? How could they be strengthened moving toward the end of the COVID-19 Fund.
- Question 10: How have the Socio-Economic Response Plans contributed to or learned from the plans/actions of national governments, bilateral agencies and international financial institutions?
- Question 11: Did the governance arrangements, criteria for proposal calls, funding levels approved and support provided to Resident Coordinators and United Nations Country Teams contribute to the development of Socio-Economic Response Plans based on joint, integrated action by United Nations Country Team entities, relevant to national needs and priorities?
- Question 14: Are the United Nations Country Team entities (resident and non-resident UN agencies) able to participate in joint proposal development and programming?
- Question 15: What processes were used by Resident Coordinators (with support of Resident Coordinator Offices) to coordinate the United Nations Country Team response to proposal calls? Which approaches were most effective?
- Question 16: Are the Socio-Economic Response Plans aligned with national priorities and how can they influence the UN programme cycle going forward?
- Question 17: How have the Socio-Economic Response Plan and the proposal and implementation process for Fund programme support contributed to a more integrated United Nations Country Team approach allowing for work across mandates which draws on the combined capacities of the United Nations Country Team? How is this relevant to the national needs including those of the most vulnerable?
- Question 18: How have Socio-Economic Response Plan and Fund processes at country level contributed to a stronger offer of capacities and services across the United Nations Country Team, especially relating to policy support responsive to the national context? Including for principles of inclusion: gender equality, human rights, disability and leave no-one behind?

Box 18: Evaluation of the UN COVID-19 Response and Recovery Multi-Partner Trust Fund (MPTF) (continued)

Annex 1 of the report presents a lessons learning matrix and this contains a further layer of sub-questions/potential lines of enquiry. However, the preamble to the matrix explains that *"not all lines of enquiry will be explored during the exercise as some will provide richer and stronger evidence for Early Lessons than others."* Reference is made to addressing all sixteen assessment questions although it appears that these have now been expanded to 20. There are a total of 59 sub-questions. Two of these do specifically mention coherence and both of these fall under question 18. They are:

- Do key stakeholders at country level, including the Resident Coordinator and Resident Coordinator Office staff, staff of United Nations Country Team entities, national government staff, and representatives of excluded groups feel that United Nations Country Team entities provided a more coherent offer of their capabilities, especially around policy support for the socio-economic response to COVID-19?
- If so, did the Fund and the Socio-Economic Response Plan encourage/facilitate this more coherent policy engagement?

In addition, many of the sub-questions could implicitly relate to coherence. Examples include sub-questions related to:

- Question 3: In comparison to other pooled funds, most specifically the SDG fund, are their characteristics of the Fund that provide value-added? What are they and what value added do they bring? Are there specific aspects of how the Fund incorporated principles of gender, human rights, disability and leave no-one behind that represent comparative advantages or value-added for the Fund?
- Question 5: Did United Nations Country Teams respond positively and collaboratively to the call for proposals and the proposal development process as managed by the Resident Coordinator? If not, what factors contributed to the lack of response?
- Question 6: At a headquarters level, did the COVID-19 emergency and the attendant constraints it imposed on travel and other forms of collaboration accelerate the move to digital means of consultation, collaboration, decision making? Did it help to accelerate transitions to more inclusive and transparent processes, including in Fund approval? Did the COVID-19 emergency help to accelerate consultations between the Fund and Resident Coordinators (and United Nations Country Teams) at country level using remote methods? Do Resident Coordinators, Resident Coordinator Offices, and United Nations Country Team entities feel that they had more meaningful and strategic input into the direction of the Fund than is normally the case for pooled funding mechanisms? If so, why and how?
- Question 8: At a global level, how do the Advisory Committee of the Fund, the MPTF Office and the Fund secretariat work to promote linkages, especially with the SDG fund? At country level, are Resident Coordinators and United Nations Country Team entities working to ensure alignment across the different funds? As the Fund progresses towards its end date, are there elements of its design, operation and governance which could be adopted by the Joint SDG Fund?
- Question 10: Are there specific references in the Socio-Economic Response Plans to plans or programmes of national governments, bilateral agencies and international financial institutions? What indications are there that Socio-Economic Response Plans have been influenced or influenced national government, bilateral agency or international financial institution plans?
- Question 11: What processes were used by Resident Coordinators to convene and animate/motivate United Nations Country Team entities to collaborate in the development of the Socio-Economic Response Plans? Which were most effective? Did specific instruments of the Fund, including the allocation of funds for proposal development and Socio-Economic Response Plan planning, assist the Resident Coordinator in engaging United Nations Country Team members in the development of the Socio-Economic Response Plans? Did the resulting Socio-Economic Response Plans encompass/envision integrated and coordination action by the United Nations Country Team entities in one or more of the five pillars? If so, what contributed to the commitment for coordinated action? If not, what impeded it?
- Question 14: Are United Nations Country Team entities (resident and non-resident UN agencies) responsive to calls for collaboration, coordination, and joint programming in order to access the Fund resources given the small allocations available when compared to identified financial requirements in each pillar? If they are, what is working to incentivise and enable their responsiveness? Is participation in successful proposals for programme funding dominated by large, operational, and resident United Nations Country Team entities? If so, why? Alternatively, is participation by smaller and non-resident UN entities broader in the case of the Fund when compared to similar financing mechanisms including the SDG Fund? If so, why? What processes and criteria support this? Where smaller United Nations Country Team entities and non-resident UN agencies have participated in development of the Socio-Economic Response Plan or in funded programmes, what steps were involved in ensuring their participation? How did they overcome the disadvantage of either a small in-country footprint or non-resident UN agency status?
- Question 15: What steps were taken by the Resident Coordinator and what support provide by the Fund to facilitate participation by smaller United Nations Country Team entities and non-resident UN agencies? What factors promoted or inhibited their participation? How did Resident Coordinators convene and coordinate United Nations Country Team members in responding to appeals for proposals by the Fund? What approaches were most effective? What did not work? In the experience of the Resident Coordinator and United Nations Country Team entities were the Socio-Economic Response Plans a useful framework for coordinating an integrated approach to policy support and advocacy by the United Nations Country Team? If not, what was lacking?
- Question 16: Did the socio-economic impact assessments underlying the Socio-Economic Response Plans recognize and, if possible, incorporate social and economic diagnostic material from the Common Country Assessments? When seen from the perspective of United Nations Country Team entities and national governments, does the content of the Socio-Economic Response Plans align to the extent it can to the content of the relevant United Nations Development Assistance Framework/United Nations Sustainable Development Cooperation Framework? What are the issues and challenges faced by United Nations Country Team entities as they collaborate to advocate for investments and policies consistent with the Socio-Economic Response Plan and the United Nations Development Assistance Framework/United Nations Sustainable Development Cooperation Framework?

42. Subjects for evaluation are broadly similar to the types of evaluation of coherence before the COVID-19 pandemic (see paragraph 29). As with evaluations of coherence before COVID-19, there were no evaluations of the coherence of an entire national response to COVID-19. All were caveated in some way, e.g. by theme or agency. Some high-income countries, such as Sweden, have begun to evaluate their domestic response to COVID-19 but reports to date have focused on a particular theme, elderly care (Coronakommissionen, 2020) (Box 19).

> **Box 19: The Corona Commission in Sweden**
>
> In order to evaluate the management of the virus outbreak and the effects of the outbreak, the Swedish Government established a Corona Commission in June 2020. According to its website, the Commission has been instructed to evaluate the measures taken by the government, the relevant administrative authorities, the regions and the municipalities to limit the spread of the virus that causes COVID-19 disease and its effects. The Commission is also expected to make an international comparison with relevant countries of the various measures taken and the effects of the measures. To date, the Commission has produced a report focused on elderly care and is scheduled to produce a more wide-ranging interim report by the end of October 2021.
>
> The English summary of the report on elderly care does not mention coherence explicitly and the report's focus is the domestic situation in Sweden. Nevertheless, it does contain material related to coherence within that context, for example, the fragmented nature of elderly care in Sweden.

Table 4: What were the subjects of evaluations of COVID-19 responses that included consideration of coherence?

Subjects of Evaluation	Examples
Institution or country (donor)	• International Federation of Red Cross and Red Crescent Societies (IFRC, 2021); Acaps (Acaps, 2020) • International Financial Institution, such as the Asian Development Bank (Asquith and Bloom, 2020) (see Box 16) • Country, e.g. Belgium (Cota, 2021) (see Box 16); Scotland (Scottish Government, 2020 and 2021) • European Union (European Commission, 2020b) • UNHCR (UNHCR, 2021); UNICEF (UNICEF, 2020b-d and 2021a-b; Chazaly and Goldman, 2021); WFP (WFP, 2020b); WHO (The Independent Panel for Pandemic Preparedness and Response, 2021)
Theme or sector	e.g. food security (UN, 2020a; COVID-19 Global Evaluation Coalition, 2020a); vaccination (Lewin and Glenton, 2020; Glenton and Lewin, 2020); migration (IOM, 2020; IFRC, 2021); gender and inclusion (IFRC, 2021); communication and community engagement and accountability (IFRC, 2021); remote working (IFRC, 2021); digital

	solutions (European Investment Bank, 2020; Lewin and Glenton, 2020); cash transfers (IPA 2020a-b); refugees/IDPs (UNHCR et al, 2020; Kuhnt and Schüttler, 2020); informal sector (K4D, 2020); participation of women (K4D, 2020)
Country or region	e.g. Africa (European Investment Bank, 2020); Benin (Cota, 2021); Cambodia (UNICEF, 2020b); Colombia (IPA, 2020a-b); DRC (Cota, 2021); Malawi (Scottish Government, 2020); Malaysia (UNICEF, 2020d); MENA (Chazaly and Goldman, 2021); Mongolia (UNICEF, 2021a); Niger (Cota, 2021); Pakistan (Scottish Government, 2020); Rwanda (Scottish Government, 2020); South Asia (UNICEF, 2021b); Thailand (UNICEF, 2020c); Zambia (Scottish Government, 2020)

Box 20: Institutions or (donor) countries as the subjects of evaluations of COVID-19 responses that included consideration of coherence

In 2021, IFRC issued terms of reference for an IFRC-wide COVID-19 evaluation. The purpose of this evaluation was framed in terms of three evaluation criteria including coherence which was defined as adding value while avoiding duplication of effort. However, while there are proposed evaluation sub-questions for the other two criteria (effectiveness and relevance), no specific evaluation questions are included for coherence. One distinctive feature of the terms of reference is that it maps out other relevant research and learning activities which might inform the evaluation.

Acaps have produced an analytical framework for COVID-19 which seeks to ensure a level of coherence to the COVID-19 analysis. It focuses on the interplay between the COVID-19 pandemic, policies and interventions put in place as a reaction to the pandemic and behaviours and actions that are adopted both as a result of the pandemic itself and as a result of the policies and interventions adopted. This interplay takes place across all spheres of human life which are grouped into four pillars – health; living conditions or essential services; social cohesion and protection; and humanitarian and development operations.

In 2020, the Scottish Government announced plans to review its international development programme in the light of COVID-19. A summary report was published in March 2021. Issues of coherence emerged in a number of findings of the review. For example, in line with the Scottish Government's commitment to policy coherence, the new programme principles will apply not only to the international development programme but also to other areas of policy including climate, health, trade, education and equalities. Specifically, the principles:
- Will apply to the Climate Justice Fund and climate adaptation and water programmes
- Will apply to the NHS Scotland Global Citizenship Programme
- Are aligned to key reforms in education including Learning for Sustainability and the newly-established programme to enhance race equality and anti-racism education
- Are consistent with Vision for Trade particularly around the trade challenge of ensuring coherence with international development outcomes
- Are aligned with broader approaches to equity

In 2020, the European Commission drafted terms of reference for a fast-track assessment of the EU's initial response to COVID-19 crisis in partner countries and regions. This assessment has five specific objectives and is expected to cover 17 countries or regions. The second specific objective explicitly mentions coherence, namely assessing "*whether the Team Europe approach has supported a more coherent and efficient policy and operational response by the EU institutions, EU Member States, the European Investment Bank and the European Bank for Reconstruction and Development to COVID-19 in partner countries/regions. Particular attention should be paid to how the Team Europe approach has been coordinated at country/region level and whether this coordination has improved overall European coordination and joint work*". In addition, the fifth specific objective is implicitly about coherence as it requires assessing "*whether the EU response has been consistent with the Joint Communication on the EU's global response, the 2017 Joint Communication on Resilience and the EU's rights-based approach, and to provide evidence on how EU external action support in the area of Gender Equality and Women's Empowerment, in line with the 2018 guidelines 'Evaluation with gender as a cross-cutting dimension'*".

Identified evaluation questions are organised around five evaluation criteria including EU added value and coherence (as one criterion). The overarching question for this criterion is "*has the Team Europe initial response added benefits to what would have resulted from action taken by the EU institutions and EU Member States on their own?*" Sub-questions include:
- Has the Team Europe approach contributed to reinforcing existing, establishing new and/or implementing effective co-ordination mechanisms with other EU donors and international organisations at country and regional levels? This should relate particularly to the areas of analysis, planning, prioritisation, data collection, research, monitoring and evaluation.
- To what extent was the EU support adding value to national and/or regional responses or plans addressing the COVID-19 crisis? To what extent has the EU managed to position itself as a key player in the global fight against COVID-19?
- Has EU external action support ensured overall coherence and complementarity between its interventions (regardless of their delivery methods, funding channels and instruments)?
- What was the effect of the EU response under the Team Europe approach on on-going cooperation, at country or regional levels? Are there any direct benefits or drawbacks? In the case of identified drawbacks, are there mitigation measures in place or plans to address them in short or medium term?

In January 2021, UNHCR outlined briefly how it would approach evaluating responses to COVID-19. This identified three COVID-19 related evaluations which it would be leading or supporting. In addition to the joint evaluation on the protection of refugee rights during COVID-19 and the inter-agency humanitarian evaluation of the GHRP, UNHCR was planning a meta-evaluation of UNHCR's own adaptation and response. While the information about the meta-evaluation is limited, it does include an explicit focus on coherence. Part of the statement of content and scope is that "*the meta-evaluation will include a focus on the coherence and effectiveness of UNHCR's collaboration and coordination with partners (including other UN agencies), governments and civil society organizations in responding to COVID-19. The analysis and recommendations will be framed as being future-focused and generate evidence that could be used to enhance the organization's understanding of the effects of COVID-19 and government's actions to mitigate COVID-19, and the agency's effectiveness, coherence and relevance of the adaptations and responses to COVID-19.*"

Box 20: Institutions or (donor) countries as the subjects of evaluations of COVID-19 responses that included consideration of coherence (continued)

UNICEF has conducted a number of reviews of areas of its work related to COVID-19. These include:
- A review of risk communication and community engagement initiative for COVID-19 prevention behaviours in Cambodia. This did not explicitly refer to coherence.
- An After Action Review of Thailand Country Office response to COVID-19 crisis. This analysed findings according to four evaluation criteria including coherence divided into internal and external elements. Internal coherence was defined as within UNICEF's Thailand Country Office while external coherence was defined as with government and other actors. Of the review's three recommendations, the third related to better managing the expectations of UNICEF headquarters and regional office. The review was structured around four typical overarching questions – what did we intend (or plan) to do; what actually happened; what went well and why; and what can be improved (and why) and what should we change in coming period (and in future responses)? Some detailed questions were developed based on these relating to internal and external coherence. These were:
  - Internal coherence – to what extent were UNICEF's COVID-19 interventions consistent between the various sections of the office and were there any synergies established between interventions by various sections of the office?
  - External coherence – to what extent were UNICEF's COVID-19 interventions consistent with government, UN and other actors' policies, priorities and interventions?
  The findings section on coherence is structured around internal and external coherence. In line with the definition and the evaluation questions, internal coherence is mainly explored in terms of internal to the UNICEF Thailand country office. However, issues related to interaction with UNICEF regional office and headquarters are also raised here implying perhaps a broader definition of internal, i.e. inside UNICEF as a whole.
- Real-time assessments of COVID-19 for UNICEF Malaysia and Mongolia. These did not explicitly refer to coherence although there was some consideration of value addition along with consideration of success. Other areas assessed varied by country but included adaptability, equity, reaching beneficiaries, reaching the most disadvantaged, leaving no child behind and timeliness.
- A real-time assessment of COVID-19 for UNICEF in South Asia. This did mention the lack of coherence that might come from remote working and so this issue was explored in this assessment which covered Afghanistan, Bangladesh, Bhutan, India, Maldives, Nepal, Pakistan and Sri Lanka. It included similar areas to some of the other real-time assessments, such as adaptability, but it also included a short section on the perceived quality of partnerships with UNICEF by government and other partners. The report contained a section on what UNICEF should do more of, less of or should do differently. Under the last heading, there was a request for more cross-sectoral work and to partner more with others, such as WHO and WFP.
- A real-time assessment of COVID-19 for UNICEF in the Middle East and North Africa. This did not mention coherence explicitly. The assessment included case studies from Egypt, Iran, Jordan, Oman, Tunisia and Yemen. The findings were mostly highly technical, e.g. on infection prevention and control, education services etc. but there was also material on implementation and opportunities. Within the opportunities section, there was some material implicitly related to coherence, such as strengthening the partnership with WHO and Ministries of Health opportunities for increased intersectoral coordination.

WFP has produced terms of reference for an evaluation of their response to the COVID-19 pandemic from 2020 to 2021. These are very extensive and cover OECD DAC evaluation criteria including coherence, based on the OECD DAC definition. Specific evaluation questions related to coherence include:
- How well has WFP fulfilled its role as a partner in the collective humanitarian response, at country, regional and at global level?
- To what extent has WFP maintained/broadened its global and national partnerships during the crisis, and what mutual benefits did this bring?
- To what extent and how well has WFP supported national responses to Covid-19?
- To what extent has WFP delivered its intended role in the global response (GHRP) e.g. in common services/upstream supply services?

WHO has established an Independent Panel for Pandemic Preparedness and Response. Their terms of reference do not explicitly mention coherence. The panel is expected to carry out an impartial, independent and comprehensive evaluation of the WHO-coordinated international health response to COVID-19 as one important step and measure to implement the request in the World Health Assembly resolution. In their second report on progress in January 2021, the Panel identified that they would pay particular attention to the coherence and prioritization of recommendations and evidence provided to countries. Some of their findings implicitly relate to coherence, for example, the need for whole-of-government and whole-of-society responses and their consideration of global and regional leadership.

Box 21: Themes or sectors as the subjects of evaluations of COVID-19 responses that included consideration of coherence

Evaluations of COVID-19 responses have considered a wide range of themes and sectors as subjects for evaluation.

In 2020, the United Nations produced a policy brief on the impact of COVID-19 on food security and nutrition. However, this did not explicitly consider coherence. The COVID-19 Global Evaluation Coalition generated lessons from evaluation related to food security. While this too did not explicitly mention coherence, it did raise a number of issues that implicitly relate to coherence including the need for:

- "Nexus" thinking, that is coordinating humanitarian, development, peace and stabilisation programming.
- A multi-sector approach that combines food availability, food access and food use with gender equality and intersectional approaches; nutrition programming; hygiene awareness; animal health services; and environmental and climate adaptation considerations, including disaster risk reduction approaches.
- Partnerships with the private sector, academia, governments and banks.

Claire Glenton and Simon Lewin prepared two briefs for the COVID-19 Global Evaluation Coalition on communicating with the public about vaccines and effects of digital interventions for promoting vaccination uptake. These did not consider coherence explicitly.

In relation to migration, in 2020, IOM issued a call for applications to carry out a real-time evaluation of the IOM East and Horn of Africa COVID-19 response. Part of the purpose of this was to improve coordination and coherence among stakeholders involved in the operations. The evaluation was structured around a number of evaluation criteria including coherence, coordination and connectedness. Specific evaluation questions were identified for these questions including:

Coherence/coordination
- To what extent are current partnerships increasing or compromising synergy?
- To what extent are partners strategic mandates being leveraged effectively?
- What is IOM's added value to the COVID-19 response activities?
Connectedness
- To what extent are interventions linking with longer-term recovery initiatives?
- Does the intervention design have clear linkage between emergency, transition/recovery to development?

The IFRC-wide COVID review (see Box 20) noted that it would be informed by thematic-specific research in a number of areas including migration; protection, gender and inclusion; risk communication and community engagement and accountability; and remote working.

In 2020, the European Investment Bank produced a report focused on Arica's digital solutions to tackle COVID-19. This did not explicitly consider issues of coherence.

IPA have published studies of cash transfers in COVID responses in Colombia. They do not explicitly consider coherence although they do document the benefits of cash transfers across different sectors, such as food security and health.

In addition to the joint evaluation of the protection of the fundamental rights of refugees during the COVID-19 pandemic (see Box 17), Jana Kuhnt and Kirsten Schüttler produced a report for the German Federal Ministry for Economic Cooperation and Development on how to ease the impact of COVID-19 on displaced populations in low- and middle-income countries. This did not explicitly consider coherence although one of the key messages did emphasise the importance of integrated approaches.

K4D have produced a number of HelpDesk reports and emerging issues papers relating to COVID-19 including on the informal sector in Sub-Saharan Africa and the participation of women and women's rights organisations in decision-making. However, neither of these specifically address coherence.

Distinguishing internal and external coherence

43. While in theory the distinction between internal and external coherence may be clear – synergies and interlinkages between the intervention and other interventions carried out by the same institution or government as opposed to consistency of the intervention with other actors' interventions – the distinction may be less clear or helpful in practice, for example where the evaluation is not of an institution or government or there may be differing understandings of where the boundaries of that institution or government fall.

44. For example, in an After Action Review for UNICEF in Thailand (UNICEF 2020c and see Box 20), internal coherence was defined as within UNICEF's Thailand Country Office while external coherence was defined as with government and other actors. Questions were then defined focused in these areas. Potentially, had the evaluation stuck strictly to these definitions and questions, it would have had to conclude that consideration of coherence with UNICEF headquarters and regional office was either out of scope or could potentially be considered under external coherence if UNICEF headquarters and regional office were considered to be "*other actors*". In the end, the evaluation did consider coherence with headquarters and regional office as part of internal coherence effectively moving the perceived boundary of internal/external to include all of UNICEF as internal, not just the country office (see Figure 4).

45. Another example is provided from the evaluation of the Dutch Government's policy on responsible business conduct. This used the concepts of internal and external coherence and defined them as – internal relating to within the Ministry of Foreign Affairs and external as relating to other parts of the Dutch Government (see Figure 4). Is this definition in line with the OECD DAC definition? It depends. If the evaluation is taking the Ministry of Foreign Affairs as the institution being evaluated, then the definition fits. However, if the Dutch government is taken as the reference point, then both forms of coherence would be internal. It is unclear if coherence with other actors was pertinent to this evaluation. If so, this could have been overlooked by this definition.

46. Similar issues arise in many other settings. For example, when the Disaster Emergencies Committee organises a response to a crisis (see Box 3), coherence between the organisations within that response might be conceptualised as external coherence from the point of view of an individual organisation but as internal coherence from the point of view of the Disaster Emergencies Committee as a whole. Similarly, when ECDC responds to Ebola in Guinea (see Box 3), coherence with other EU institutions may be seen as external coherence from the perspective of ECDC but internal coherence from the point of view of the EU as a whole.

Figure 4: Where are the boundaries between internal and external coherence? Examples of UNICEF After Action Review in Thailand
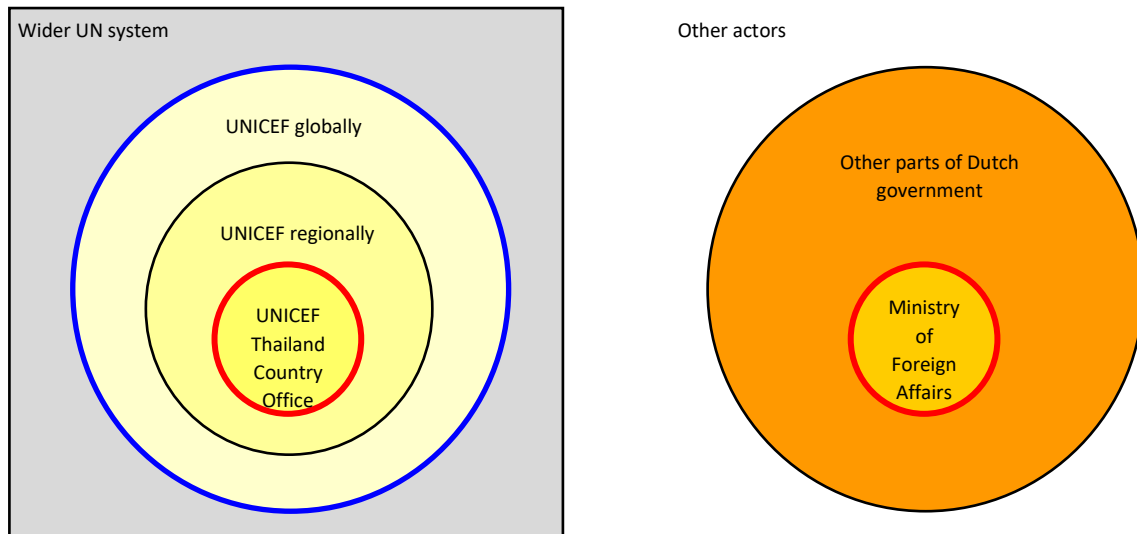
Diagrammatic representation of boundaries between internal and external coherence in After Action Review for UNICEF Thailand showing boundary as per the definition (red line) and boundary as per reporting practice (blue line)

Diagrammatic representation of boundaries between internal and external coherence in evaluation of Dutch Government's policy on responsible business conduct showing boundary as defined (red line)

47. Does it matter how internal and external coherence are defined or if they are defined at all? Probably not. The critical issue is to identify what aspects of coherence matter and then evaluate them. Thinking about internal and external coherence may be a useful way of systematically deciding what aspects of coherence matter to ensure they are included. There is a risk that, if internal and external coherence are defined in a way that does not do this, important elements of coherence may be overlooked by the evaluation. For example, if the After Action Review for UNICEF Thailand had applied their definition rigidly, important issues of coherence between the UNICEF country office and the regional office and headquarters could have been overlooked. Similarly, if matters of coherence with actors outside the Dutch government are important in terms of responsible business conduct, these risk being overlooked as a result of the way internal and external coherence have been defined.

48. It is also extremely likely that, particularly in large and complex evaluations, different stakeholders will have different perspectives as to where an entity's boundaries lie and therefore what constitutes internal and external coherence. If the distinction between internal and external coherence is important, it may be necessary to surface these differences and make them explicit. An alternative, and perhaps preferable, approach would be to ensure that key areas of coherence are included in the evaluation regardless of whether they are internal or external.

49. It may be that OECD or a particular evaluation wishes to take a more normative position. For example, if OECD wished to emphasise the importance of whole-of-government approaches, it may wish to specify that, in the case of government agencies, internal coherence *should* be

defined as synergies and interlinkages between the intervention and other interventions carried out by the same *government* and not only by the same Ministry. Similarly, if an evaluation of one part of the United Nations' system wanted to emphasise the concept of "*one UN*", it might wish to define internal coherence as within the UN rather than limited to a particular UN agency.

Proposed questions on coherence

The role of questions in evaluation

50. Before seeking to identify evaluation questions on coherence, it may be helpful to consider what evaluation questions are and why they are needed. Of course, many data collection methods used for evaluations involve questions that are asked of respondents, e.g. through a survey or questionnaire or in an interview. However, in most cases, it is not these types of questions that are intended when reference is made to evaluation questions or specifically to key evaluation questions. Rather, these are the high-level questions that the evaluation is expected to answer and which form the basis for evaluation design including selection of methods. In most cases, well-designed evaluations have a small number of agreed key evaluation questions. Some advise a maximum of five to seven such questions (BetterEvaluation, 2016) while others consider that the best approach is to identify one or two key evaluation questions as this will ensure that they are the big, meaningful issues for the evaluation to address (Ferretti, 2021). The implication of this is that where an evaluation includes consideration of multiple evaluation criteria, of which coherence is only one, there may be scope for only one key evaluation question relating to coherence.

51. This approach of selecting a small number of key evaluation questions is really about framing the evaluation and setting its scope. In some cases, this framing will be used as the basis for analysis and reporting, that is the findings section of the report may answer each question in turn. In other cases, these questions may feed into, or be mapped to, another analytical framework, such as one based on the OECD DAC evaluation criteria, and these are then used to structure the report.

52. It is clear from a cursory glance at the evaluations reviewed that this approach has not been followed by all evaluations identified. In some cases, an evaluation has identified a small number of key questions but, even in some of those cases, the number of questions may be high given that identified questions may relate only to coherence and coherence may be only one of many criteria being evaluated. In some cases, there are a huge number of evaluation questions to answer. For example, in the case of the inter-agency real-time evaluation following the Haiti earthquake, there were more than 40 evaluation questions (Grünewald et al, 2010 and Box 4). In such cases, it is impractical to answer all the questions systematically and therefore not possible to use the questions for analytical framing. In those cases, the evaluators will have little choice but to treat the questions as an advisory menu from which they might select particular questions to focus on or which they might use in a more general and less literal sense. In the inception report for the evaluation of the UN COVID-19 Response and Recovery MPTF, this issue is addressed explicitly (Freeman et al, 2021 and Box 18). The inception report identifies three areas of investigation. While these are not framed as questions, they probably serve the same purpose as key evaluation questions. However, in addition, there are a further 20 questions and 58 sub-questions or lines of enquiry. The inception report says explicitly that "*not all lines of enquiry will be explored during the exercise as some will provide richer and stronger evidence for Early Lessons than others*". The report is less clear as to why so many questions and lines of enquiry have been identified if they can't all be explored.

53. Often the process for developing terms of reference implicitly incentivises question proliferation. In her blog, Michaela Raab provides insight into how terms of reference are commonly compiled. *"I remember the first evaluation I commissioned, back in the last quarter of the 20th century. I asked my colleague how to write terms of reference (TOR). She said, "Just take the TOR from some other project and add questions that you find important". I picked up the first evaluation TOR I came across, found all the questions interesting and added lots, which I felt showed that I was smart and interested in the project. Then I shared the TOR in our team and others followed suit, asking plenty more interesting questions."* She then discusses the problems this approach causes and concludes by saying, *"… but please keep the list of evaluation questions short and clear"* (Raab, 2020). While official documents such as terms of reference and inception reports might not be expected to explain the process of question selection so candidly, it seems likely that similar processes may be being followed in some cases. The review shows that the evaluations with the longest menu of questions are joint evaluations. This is likely to be because there was opportunity for more people to comment and add questions.

54. While having a proliferation of questions may seem harmless enough, it does carry a number of risks. First, there is a danger that important questions the evaluation needs to answer might be overlooked if there is not a clear distinction between key evaluation questions, which are being used to frame the evaluation, and a broader group of other questions which constitute a menu of options which may inform the evaluation's analysis but do not need to be answered systematically. It may be wiser, if possible, to drop these latter optional questions as there is a risk of misunderstandings occurring and a danger of a mis-match of expectations. Someone who has contributed a question and sees it reflected in the terms of reference might reasonably expect that question to be answered explicitly by the evaluation and may be disappointed if it is not. Similarly, there are likely to be problems if the evaluator understands the questions as a menu or guidance while the commissioner of the evaluation or their stakeholders are expecting the questions to be answered systematically and literally.

55. In a small number of cases, alternatives to key evaluation questions are used to frame the evaluation and set its scope. In some cases, this is done using OECD DAC evaluation criteria to explain what the evaluation will focus on. For example, in the UNFPA evaluation of their response to the Syria crisis, the questions were pegged to ten evaluation criteria (see Figure 5). Although the criterion title was used as a shorthand for the question, each criterion had a more detailed question underneath it but this was limited to one per criterion.[11] In the case of the coherence criterion this was based on how well the response was aligned with particular priorities, strategic frameworks, principles, prioritisations and interventions (UNFPA Evaluation Office, 2019 and see Box 9 for details). In others, lines of enquiry or areas of investigation are identified, e.g. as in the evaluation of the UN COVID-19 Response and Recovery MPTF, (Freeman et al, 2021 and Box 18).

Figure 5: Using evaluation criteria as shorthand for key evaluation questions

---

[11] Recognising, as in the case of coherence, that multiple sub-questions can be included within one main question.

This is an extract from the table of contents of the UNFPA evaluation of its response to the Syria crisis from 2011 to 2018 showing how evaluation criteria can be used to define the main areas of an evaluation. Each of these areas is underpinned by a more detailed evaluation question.

COVID-19 Global Evaluation Coalition strategic evaluation questions

56. Given this, it is helpful that the COVID-19 Global Evaluation Coalition (2021a) has put together proposed strategic questions for COVID-related evaluations across the six OECD DAC evaluation criteria. For each criterion, there is one over-arching evaluation question and a number of examples of thematic/country-level questions (see Figure 6). These questions will be considered in more detail when seeking to identify potential questions for an evaluation of coherence in relation to COVID-19.

57. The COVID-19 Global Evaluation Coalition (2021a) collectively developed six strategic evaluation questions which are the basis for the design and delivery of participants' individual and collective evaluation activities. These questions were developed based on a series of consultations with participants of the Coalition, sharing and analysis evaluation plans to identify shared areas of interest and priorities. They have been designed to capture learning across all stages of the COVID-19 pandemic: health response; safeguarding lives and livelihoods; humanitarian response; human rights and socio-economic responses; and recovering better.[12] They aim to support the development of a more coherent, robust and comprehensive evidence base that can be synthesised, including within sectoral and thematic reviews. The Coalition secretariat intends to complete a series of syntheses studies based around these questions.

58. The Coalition has also developed a longer list of examples of questions that could be included within thematic or country-level evaluations. These are based on a series of consultations with participants of the Coalition, sharing of evaluation plans and collective work to identify suitable questions. It is envisaged that participants will interpret and apply these questions in ways that are appropriate and useful based on their needs for evidence, specific contexts in which they are working, available resources and priorities.

Figure 6: COVID-19 Global Evaluation Coalition strategic evaluation questions

---

[12] UN Comprehensive Response to COVID-19, 2020 https://www.un.org/sites/un2.un.org/files/un_comprehensive_response_to_covid-19_june_2020.pdf
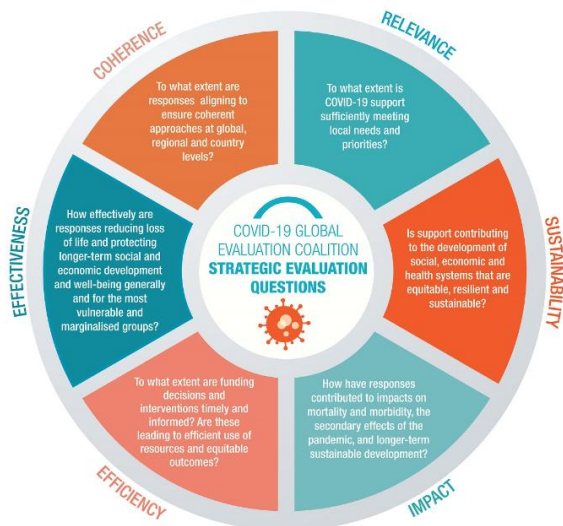
Examples of questions on coherence for thematic or country-level evaluations

If and how is support aligned between national governments, humanitarian agencies and development partners? Are ways of working contributing to or limiting the coherence, coordination and value of responses?

To what extent are there synergies and coherence in COVID-19 related responses across humanitarian-development nexus? What are the drivers and barriers to alignment?

To what extent are domestic actions of development partners coherent with the COVID-19 responses? To what extent have these domestic actions supported equitable access to vaccines, personal protective equipment (PPE) and other resources?

To what extent are plans and actions aligned with international human rights agreements?

Questions related to coherence contained in evaluations identified in this scoping study

59. A large number of evaluation questions related to coherence have been identified through this study and examples of these are included in text boxes throughout the report. While this may initially appear overwhelming, the strategic questions on coherence proposed by the COVID-19 Global Evaluation Coalition provides a useful organising and analytical framework as almost all the questions are variations on that question, namely to what extent are responses aligning to ensure coherent approaches at global, regional or country levels? While the precise context and content varies, the vast majority of the questions identified are variations of this.

60. There are some exceptions. The first is the question articulated in the ALNAP guide for humanitarian agencies in 2006 (see Box 1), namely "*should there be coherence at all*"? The coalition's strategic question, and almost all the coherence questions identified in this study, are premised on the basis that coherence is desirable and the evaluation should assess to what extent this is taking place. While this is probably true in most contexts, there may be some interventions and some evaluations where coherence is not particularly pertinent. However, it may be better to determine this in the design or inception phase of an evaluation and then exclude this criterion from the evaluation design rather than including "*should there be coherence at all*" as a key evaluation question, although this might be appropriate in exceptional circumstances. One issue (as explained in Box 1) is that the ALNAP guide states that "*coherence may be less relevant for evaluating single-agency or single-project interventions*" when potentially such interventions may be most at risk of fitting poorly with what others are doing (or with what other parts of the same agency are doing) and might conversely benefit most from evaluating coherence. As with evaluations which claim that other evaluation criteria, such as impact or sustainability, are not pertinent in their context, any such claim on coherence should be clearly justified and potentially questioned and challenged through any evaluation governance mechanisms in place.

61. Second, while almost all questions identified are variations on the strategic "*to what extent*" question, there are some which explicitly go beyond this and seek to understand why the situation is as it is, that is what the determining factors are. Examples of such explicit questions are found in the ALNAP guide for humanitarian agencies (ALNAP, 2006a and Box 1) and in the inception report for the evaluation of the UN COVID-19 Response and Recovery Multi-Partner Trust Fund (Freeman et al, 2021 and Box 18). However, it seems likely that, in many cases, questions like the strategic question proposed by the COVID-19 Global Evaluation Coalition based on "*to what extent*" imply that there will be some attempt to understand the determining factors for the situation being as it is. This is particularly likely where such questions are being used as

key evaluation questions (BetterEvaluation, 2021) to frame and to indicate the scope of the evaluation. Where it is crucial to specify that understanding such factors is expected as part of the evaluation, this could be added to the "*to what extent*" question. This might mean that the COVID-19 Global Evaluation Coalition's strategic evaluation question could be "*To what extent are responses aligning to ensure coherent approaches at global, regional or country levels? What are the factors determining factors for the extent of coherence observed?*"

62. The material that follows presents some examples of specific evaluation questions that apply the COVID-19 Global Evaluation Coalition strategic evaluation question in particular contexts (see Figure 7). These contexts are drawn from the OECD DAC definition of coherence (green boxes in Figure 7), the COVID-19 Global Evaluation Coalition strategic evaluation question (yellow boxes in Figure 7) and the analytical framing used in this study (blue boxes in Figure 7).

Figure 7: Using the OECD DAC definition of coherence, the COVID-19 Global Evaluation Coalition strategic evaluation question and the analytical framing used in this study to unpack and understand contexts in which the COVID-19 Global Evaluation Coalition strategic evaluation question on coherence may be applied.
(Colour coding – green boxes relate to OECD DAC definition of coherence; yellow boxes relate to levels specified in the COVID-19 Global Evaluation Coalition strategic question on coherence; blue boxes relate to the analytic framing used in this study)
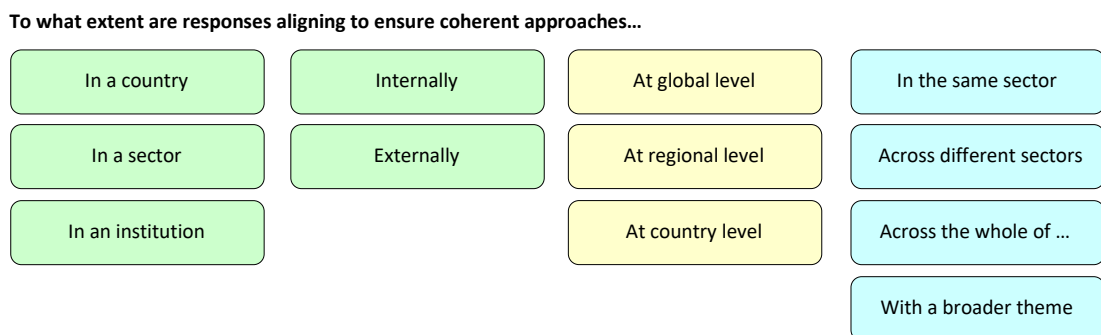
63. There are many examples of questions which focus on the extent to which responses are aligning to ensure coherent approaches in a country or at country level. For example, in the European Commission's fast-track assessment of the EU's initial response to the COVID-19 crisis in partner countries and regions (European Commission, 2020b and Box 20), there are several questions to apply at the country level, including "*to what extent was the EU support adding value to national… responses or plans addressing the COVID-19 crisis?*" Similarly, WFP's evaluation of its response to the COVID-19 pandemic (WFP, 2020b and Box 20) has questions focused on the country level, including "*to what extent and how well has WFP supported national responses to COVID-19?*"
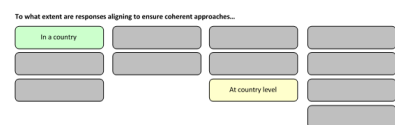
64. There appear to be fewer examples of questions which focus on the extent to which responses are aligning to ensure coherent approaches within a sector or across different sectors. It may be reasonable to assume that where an evaluation relates to a particular sector, that
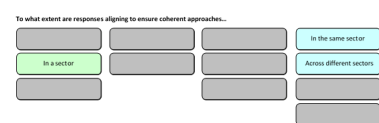
evaluation's questions might implicitly relate to that sector. For example, while the questions identified in the call for applications for a real-time evaluation of the IOM East and Horn of Africa COVID-19 response (IOM, 2020 and Box 21) are fairly generic, it may be reasonable to assume that they will examine the extent to which responses are aligning to ensure coherent approaches in relation to migration. Similarly, based on the evaluation lessons identified in the food security sector (COVID-19 Global Evaluation Coalition, 2020a and Box 21), it may be reasonable to ask the extent to which responses on food security (including food availability, food access and food use) are aligning to ensure coherent approaches with responses in other sectors, such as gender equality, nutrition, hygiene, animal health and environment and climate adaptation.

65. There are many examples of questions which focus on the extent to which responses are aligning within an institution. For example, in WFP's evaluation of its response to Ebola in West Africa (WFP, 2017 and Box 3), one of the questions asked was "*to what extent was WFP's response (and activities) aligned to WFP's corporate policies? To what extent were these policies relevant to operational needs and objectives?*" Similarly, in the evaluation of Islamic Relief Worldwide's response to the Nepal earthquake (Bhattarcharjee, 2017 and Box 5), one of the questions asked was "*was Islamic Relief Worldwide 's response coherent with relevant Islamic Relief Worldwide policies, international principles and standards?*" Also, in the evaluation of the UNFPA response to the Syria crisis (UNFPA, 2019 and Box 9), one of the questions asked was *"to what extent is the UNFPA response aligned with UNFPA strategic frameworks*". There are some points of nuance here. First, while it is common for evaluations to take an institutional lens and look for example at how parts of their own response on a particular topic fit together, there is a sense in which doing this is more about effectiveness than coherence. It is only when the evaluation seeks to assess the extent to which the response aligns with other policies and interventions of the institution, as is the case in the three examples of questions illustrated here, that this might be moving beyond evaluating response effectiveness to evaluating coherence within an institution. When the evaluation examines how the totality of what an institution is doing interacts with what it is doing in a particular intervention and vice versa, this may be considered to be taking a whole-of-institution approach.
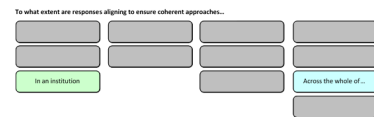
66. Relatively few evaluations explicitly distinguish between the extent to which responses are aligning internally and externally but some do and some examples of this have been discussed earlier (see paragraphs 43 to 49). There are other examples. In UNICEF's evaluation of its response to the Ebola outbreak in West Africa (UNICEF, 2017 and Box 3), the questions on coherence were simply *"how well coordinated internally was UNICEF's response to Ebola?"* and *"how well coordinated externally was UNICEF's response to Ebola?"* In addition, there are many examples of questions which implicitly explore internal or external coherence without explicitly identifying these as such. For example, in the evaluation synthesis and gap analysis of the international response to the Syria crisis (Darcy, 2016 and Box 9), there is a question that is implicitly about external coherence namely "*what picture of inter-agency coordination emerges from the material? And what picture of coordination with governments?*".
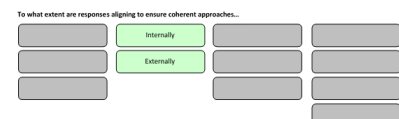
67. Many evaluations include questions which seek to explore the extent to which responses are aligning at global level. One example of these questions, drawn from the terms of reference for WFP's evaluation of its response to the COVID-19 pandemic (WFP, 2020b and Box 20) covers all
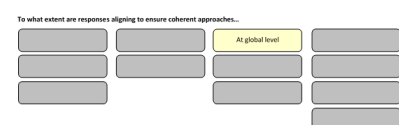
three levels identified in the COVID-19 Global Evaluation Coalition strategic question, namely "*how well has WFP fulfilled its role as a partner in the collective humanitarian response, at country, regional and at global level?*" There are other examples of questions which assess the extent to which responses are aligning at global level including other questions within the same WFP evaluation, such as "*to what extent has WFP delivered its intended role in the global response (GHRP) e.g. in common services/upstream supply services?*" In its assessments of the state of the humanitarian system (e.g. ALNAP, 2018 and Box 1), ALNAP includes the following questions, "*to what degree are humanitarian efforts coherent with core principles and international humanitarian law?*" and "*to what degree are humanitarian actors effective in encouraging support for international humanitarian law and international refugee law?*". While the wording of the question implies a focus on global positioning and visibility, the question "*to what extent has the EU managed to position itself as a key player in the global fight against COVID-19?*" included in the fast-track assessment of the EU's initial response to the COVID-19 crisis (European Commission, 2020b and Box 20) could be seen as asking to what extent the response was aligned (or seen as adding value) at the global level.

68. Similarly, a number of evaluations include questions which seek to explore the extent to which responses are aligning at regional level. For example, in the evaluations of WFP's regional response to the Syrian crisis (WFP 2018, Box 9), one of the questions was "*to what extent was WFP response well-aligned with regional responses to the crisis?*" Similarly, in the real-time evaluation of ADB's response to the COVID-19 pandemic (Asquith and Bloom, 2020 and Box 16), one of the questions was "*how… coherent is ADB being in allocating its resources to support the Asia and Pacific region during the COVID-19 crisis?*" In the case of the evaluation of WFP's response to Ebola in West Africa (WFP. 2017 and Box 3), there were a number of questions which related to the extent to which responses were aligned within the region and the countries of the region including:
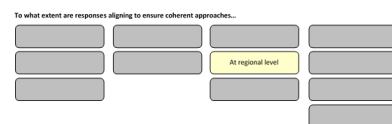
- *Was WFP's response coherent with national priorities and effectively and efficiently coordinated with the governments of Ebola affected countries?*
- *To what extent was WFP's response coordinated with the United Nations Mission for Ebola Emergency Response and other UN agencies, enabling synergies and multiplying opportunities at strategic and operations levels and taking account of the shifting frameworks for coordination?*
- *Was WFP's response coherent and aligned with the priorities of other partners, enabling synergies at operations levels?*

69. Finally, there are some evaluations which seek to explore the extent to which responses are aligning with a broader theme or policy. This ties in to assessment of the extent to which responses are aligning at the global level, e.g. in the case of international humanitarian and international refugee law (see paragraph 68). In addition, some humanitarian evaluations explore the extent to which responses are aligning with the humanitarian and development (and peace) nexus. For example, in the UNFPA evaluation of their response to the Syria Crisis (UNFPA Evaluation Office, 2019 and Box 9) one of the questions was "*to what extent does the UNFPA response promote the humanitarian-development nexus?*" Perhaps surprisingly, there are relatively few evaluations which explore the extent to which responses are aligning with the Sustainable Development Goals and efforts by the institution being evaluated and others to reach them. A recent study (Ishida, 2020 and Box 14) of 21 evaluations of teacher professional development projects
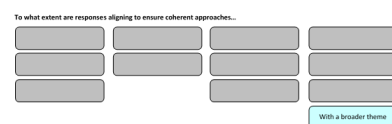
supported by the Japan International Cooperation Agency concluded that few, if any, of the evaluations systematically examined coherence with SDG targets. There are some exceptions. For example, some of the questions in the evaluation of the United Nations COVID-19 Response and Recovery Multi-Partner Trust Fund (Freeman et al, 2021 and Box 18) relate to coherence between this trust fund and others including one related to SDGs. For example, one of the questions in that evaluation is "*what linkages and synergies are evident between the fund and other, related MPTFs including, for example the SDG fund…? How could they be strengthened moving toward the end of the COVID-19 Fund?*"

Developing questions on coherence for inclusion in future evaluations of responses to COVID-19

70. Given the material above (see paragraphs 50 to 55) on the importance of identifying a small number of key evaluation questions to frame and set the scope of an evaluation, it is likely that most evaluations could take the strategic evaluation question proposed by the COVID-19 Global Evaluation Coalition and then adapt and apply it for their particular context. It is difficult and potentially ill-advised to try to identify more detailed questions based on this strategic evaluation question which can then be applied as a blueprint to any evaluation of responses to COVID-19 without fully considering specificities of context. With this in mind, Figure 8 presents a flow chart that could be used to adapt and apply the strategic evaluation question on coherence for particular COVID-19 related evaluations.

71. This is structured around the analytical framing used for this study (see, for example, Figures 2 and 3 and accompanying narrative). The first step is to determine whether matters of fit or coherence are pertinent to the evaluation. If they are not, the evaluation might exclude consideration of the coherence criterion providing a detailed rationale for this decision (see also paragraph 60). If coherence is to be considered by the evaluation, the next step would be to define the entity or entities that are the subject(s) of the evaluation and to identify other actors/entities with whom an intervention may need to be coherent.

72. The next steps seek to identify in which areas of coherence the strategic evaluation question on coherence might be applied. For example, if the evaluation is interested in coherence within a sector, such as health, social protection, livelihoods etc., the evaluation might include a question such as "*to what extent does the response align to ensure coherent approaches within a sector?*" Similarly, if the evaluation is interested in coherence across sectors, the evaluation might include a question such as "*to what extent does the response align to ensure coherent approaches across sectors?*" Some details of how evaluations have addressed such questions to date are included in paragraph 64. However, as noted there, there are relatively few examples of assessing coherence within and across sectors and using the proposed flow chart in Figure 8 as a checklist may enable evaluations to avoid overlooking this area of coherence if it is of importance.

73. If the evaluation is interested in how the COVID intervention or response fits together with other interventions (activities, responses policies etc.) of the entity being evaluated, the evaluation might include a question such as "*to what extent does the response align with what else the entity or entities is/are doing?*" Some details of how evaluations have addressed such questions to date are included in paragraph 65. Finally, if the evaluation is interested in how the COVID intervention or response fits together with broader themes and topics, such as the SDGs, the triple nexus, human rights, equity and inclusion etc., the evaluation might include a question such as "*to what extent does the response align with [these] broader themes?*" Some details of how evaluations have addressed such questions to date are included in paragraph 69.

Figure 8: Proposed flow chart for adapting and applying the proposed strategic evaluation question on coherence for particular evaluations of responses to COVID-19
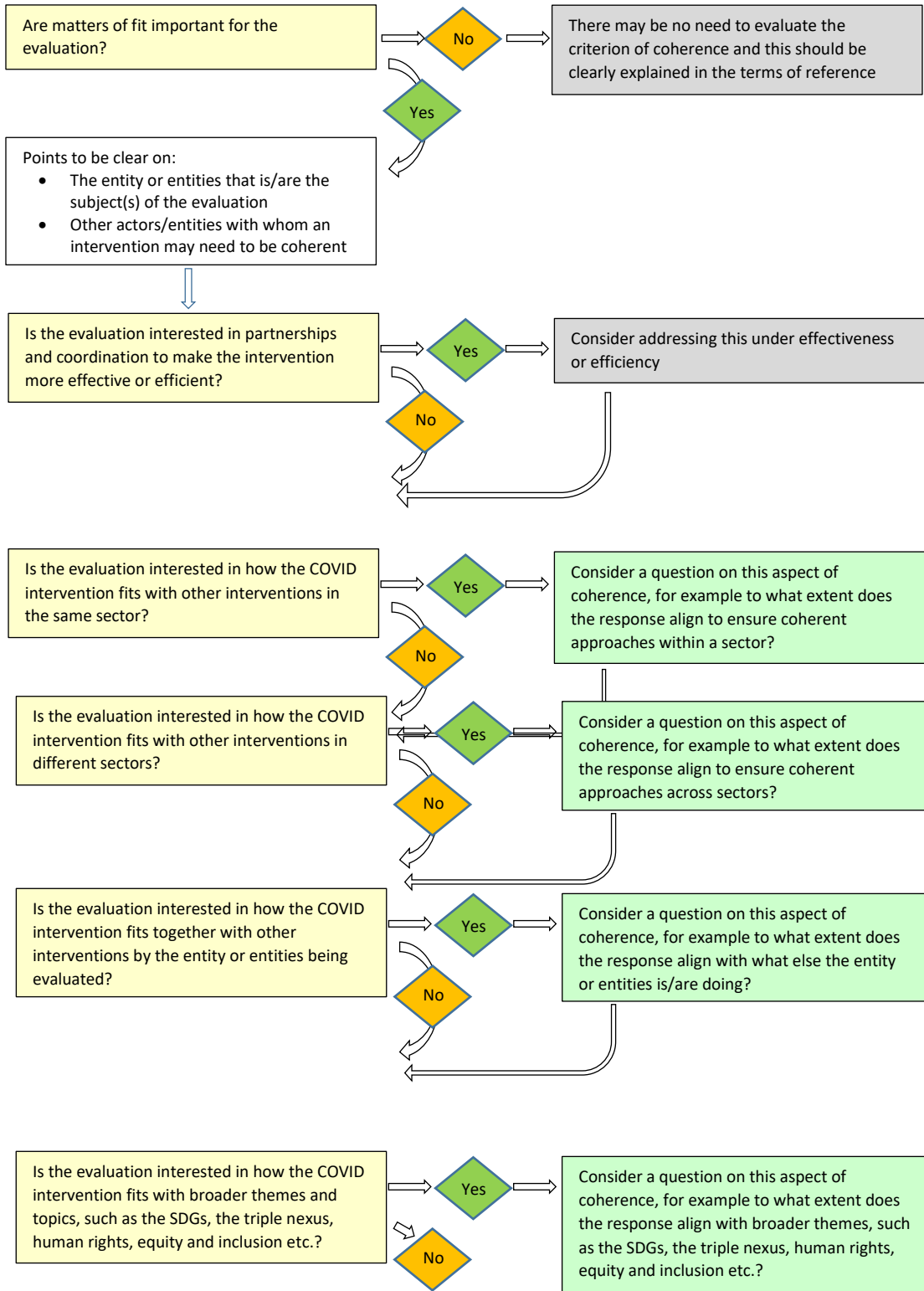
74. Nevertheless, there is strong interest from OECD and participants of the COVID-19 Global Evaluation Coalition in developing more detailed questions to guide evaluations of the coherence of COVID-19 responses. This has been done in other settings, e.g. footprint evaluation (BetterEvaluation, undated; Davidson and Rowe, 2021). This work seeks to identify a small number of key evaluation questions that can guide footprint evaluations, that is evaluations focused on the "*footprint*" that human systems make on natural systems. There are seven such questions focused on the criteria of relevance and coherence (together); design and adaptation; implementation; outcomes and impact; patterns, outliers and links; durability; and overall value. The identified question for relevance and coherence is "*how relevant is the evaluand[13] to the population/sector and the natural environment – and how well does it complement other efforts in the context?*" It includes a number of sub-questions and considerations including one that pertains specifically to coherence, namely "*how well does it complement other initiatives or change efforts that affect this population/sector and the natural environment?*" It may be important to note that these seven key evaluation questions cover a wide range of criteria and there is only one that pertains to coherence, which would be broadly analogous to the strategic question on coherence identified by the COVID-19 Global Evaluation Coalition.

75. Table 5 presents this study's best attempt to meet this expectation. It is based on the structure used by Davidson and Rowe of a table with one column for key evaluation questions and another for sub-questions and considerations. It builds from the COVID-19 Global Evaluation Coalition's strategic question on coherence and seeks to apply it across the elements identified in Figure 7 which themselves are derived from the OECD DAC definition of coherence, the COVID-19 Global Evaluation Coalition's strategic question on coherence and the analytical framing used for this study. In order to maximise specificity, the table has been developed with a specific context in mind, i.e. a joint evaluation of a number of bilateral agencies' responses to COVID-19. However, an attempt has been made to balance this with keeping the material generic enough to be relevant in other contexts.

---

[13] This term is not explained or defined. It is not widely used or understood outside evaluation circles and there may be different understandings among evaluators. Given these issues, this term has not been used in this scoping study except where directly citing a source. Broadly, it is the intervention being evaluated, the subject or object of an evaluation and may be a programme, a system, a person, an idea, a policy, an object, performance or any other entity.

Table 5: Proposed key evaluation questions for evaluating coherence of responses to COVID-19

	Key evaluation questions	Sub-questions and considerations
Strategic evaluation question: To what extent are responses aligning to ensure coherent approaches…		
1	… in a country?	To what extent is there coherence across different aid modalities, for example, humanitarian and development aid, bilateral aid allocated centrally as compared to bilateral aid allocated by a country office, multilateral as compared to bilateral aid?
		To what extent is there coherence between aid spending and other activities of the donor governments involved in the evaluation, e.g. diplomacy, trade, security, health etc?
		Who are the other key actors in particular countries with whom coherence should be assessed (e.g. national and local government, other development partners, civil society, academic institutions, private sector, communities etc.)? To what extent is there coherence between the responses being evaluated and other relevant responses?
		To what extent is there coherence within and beyond the COVID-19 response? For example, are there other actions/responses/policies/interventions beyond specific COVID-19 responses which have had an effect, positive or negative, on COVID responses? Similarly, have COVID-19 responses had an effect, positive or negative on other actions/responses/policies/interventions beyond specific COVID-19 responses?
2	… within a sector?	What specific sectors are included in the responses to COVID-19 being evaluated? Which of these, if any, are within the scope of the evaluation?
		In each of those sectors, what effects, positive or negative, have there been on other parts of the sector? Conversely, have there been particular elements (activities, interventions, structures, systems) within the sector that have had a particular effect, positive or negative, on COVID-19 responses?
3	… across sectors?	What specific sectors are included in the responses to COVID-19 being evaluated? Which of these, if any, are within the scope of the evaluation?
		Across these sectors, what effects, positive or negative, have there been on other sectors? Conversely, have there been particular elements (activities, interventions, structures, systems) in another sector that have had a particular effect, positive or negative, on COVID-19 responses in a particular sector?
4	…within an institution?	To what extent are the domestic COVID-19 responses of governments involved in the evaluation coherent with their support for international COVID-19 responses and vice versa?
		To what extent are the different elements of support to the international COVID-19 response provided by governments involved in the evaluation coherent with each other?

	Key evaluation questions	Sub-questions and considerations
		To what extent are activities (policies/interventions/financial support etc.) of the governments involved in the evaluation coherent with their support for international COVID-19 responses and vice versa?
		To what extent do the governments involved in the evaluation have a coherent whole-of-government approach to COVID-19?
5	…internally?	If this analysis is desired[14], it will be important to have a clear and documented understanding of where entities' boundaries are. For example, is internal, within a particular ministry, e.g. Foreign Affairs, or is it internal to a particular government?
		To what extent are the entities' approaches to COVID-19 coherent? Are those approaches coherent with activities beyond COVID-19?
6	…externally?	If this analysis is desired, it will be important to have a clear and documented understanding of where entities' boundaries are. For example, is external, beyond a particular ministry, e.g. Foreign Affairs, or is it external to a particular government?
		Who are the other key actors with whom coherence should be assessed? Is it best to consider these at country, regional and global level? Might such an approach miss any key actors, for example domestic actors, such as local government, media, civil society, private sector, academic institutions etc.?
		To what extent is there coherence between the responses being evaluated and relevant responses (considering both COVID-specific and other responses) of others?
7	… at global level?	To what extent is there coherence in terms of engagement of the governments participating in the evaluation on the "*global stage*"? This might include engagement specific to COVID and international development but may also be broader.
		Who are the other key actors globally with whom coherence should be assessed (e.g. United Nations, G7, G20, OECD etc.)? To what extent is there coherence between the responses being evaluated and other relevant responses?
		To what extent is there coherence within and beyond the COVID-19 response? For example, are there other actions/responses/policies/interventions beyond specific COVID-19 responses which have had an effect, positive or negative, on COVID responses? Similarly, have COVID-19 responses had an effect, positive or negative on other actions/responses/policies/interventions beyond specific COVID-19 responses?
8	… at regional level?	To what extent is there coherence in terms of engagement of the governments participating in the evaluation in particular regions? This might include engagement specific to COVID and international development but may also be broader.
		Who are the other key actors in the region with whom coherence should be assessed (e.g. groupings of countries, regional structures of

---

[14] In this particular hypothetical context, this seems to overlap with #4 … within an institution?

Key evaluation questions	Sub-questions and considerations
	global bodies)? To what extent is there coherence between the responses being evaluated and other relevant responses? Might there be need to consider coherence between regions, particularly where one regional body, such as the European Union, is active in another region?
	To what extent is there coherence within and beyond the COVID-19 response? For example, are there other actions/responses/policies/interventions beyond specific COVID-19 responses which have had an effect, positive or negative, on COVID responses? Similarly, have COVID-19 responses had an effect, positive or negative on other actions/responses/policies/interventions beyond specific COVID-19 responses?
9 … with a broader theme?	Are there particular themes or policies with which responses might expect to be coherent? If yes, what are they and how will they be evaluated?
	To what extent are responses coherent with responses focused on supporting achievement of the SDGs? This needs to go beyond simply mapping how any intervention might theoretically contribute to the SDGs. Rather, there may be need to collect detailed evidence of the effects (positive or negative) of COVID responses on progress towards achieving the SDGs. There may be some need for the reverse of this, for example, are there effects (positive or negative) of other activities and interventions to reach the SDGs on COVID responses?
	In humanitarian contexts (including fragile and conflict-affected contexts), has there been consideration of coherence between humanitarian interventions, development interventions and activities focused on peace and stabilisation? This consideration might include the actions of governments involved in the evaluation but would likely include others also.
	To what extent are responses coherent with fundamental human rights?
	To what extent are responses coherent with fundamental principles, such as equity, inclusion and leaving no-one behind?

Evaluating coherence: Approaches, methods, tools and ways of working

76. In addition to identifying questions that an evaluation of coherence in COVID-19 responses might ask, there is need to determine how such questions might be answered. This section looks at this in a number of ways. First, it considers suitable approaches to evaluation, then it identifies methods and tools and finally it considers ways of working that might be best suited to exploring issues of coherence.

Approaches to evaluation

77. Most of the evaluations were similar in overall approach in that they took place sometime after an intervention had begun and looked back retrospectively on what had happened. Some of these occurred at the end of intervention while others occurred at a point where the intervention would or might continue.[15] By contrast, some evaluations explicitly used real-time approaches which are designed to give immediate (real-time) feedback to those planning or implementing a project or programme. Most real-time evaluations have been of emergency or humanitarian responses. They tend to take place much earlier in the life of responses than other approaches to evaluation so it would be unsurprising if real-time evaluations were disproportionately represented in COVID-19 evaluations identified to date. Box 23 presents details of real-time evaluations identified in this study.

Box 23: Real-time evaluations identified in this study

Unsurprisingly, all the real-time evaluations identified pre-COVID related to humanitarian crises including the Indian Ocean tsunami and earthquakes in Haiti and Nepal. In 2005, Abhijit Bhattacharjee conducted the second round of a real-time evaluation of the tsunami response in Asia and East Africa for IFRC (see Box 6). In 2010, François Grünewald and others conducted an inter-agency real-time evaluation in Haiti three months after the earthquake (see Box 4). In 2011, Alam and Balthazar conducted a real-time evaluation of ActionAid's emergency response programme to the earthquake in Haiti. However, this did not explicitly consider coherence (see Box 4). In 2015, Jock Baker and others carried out for IFRC a real-time evaluation of the Nepal Earthquake Response Operation (see Box 5).

There have been a number of real-time evaluations (and assessments) planned or carried out. These include for ADB, IOM and particularly UNICEF. In 2020, Asquith and Bloom produced an approach paper for a real-time evaluation being conducted on ADB's response to the pandemic (see Box 16). There were also a number of real-time COVID assessments for UNICEF including in Malaysia, Mongolia, the Middle East and North Africa and South Asia (see Box 20). In 2020, IOM called for applications to conduct a real-time evaluation of the IOM COVID-19 response in East and Horn of Africa (see Box 21).

In January 2021, Ricardo Arqués and others produced a report of a real-time evaluation of UNICEF's response to COVID-19 in Malawi. Findings are presented according to a number of evaluation criteria, including coordination and connectedness but not coherence. The findings presented in relation to coordination and connectedness are pertinent to coherence as defined as an OECD DAC evaluation criterion. The questions asked under these criteria were:

Coordination
- Are existing coordination mechanisms (both internal and external) functioning effectively and efficiently to facilitate effective emergency response?
- What was the value added of UNICEF's coordination role in the national COVID-19 response in Malawi?

Connectedness
- To what extent does the UNICEF COVID-19 Response plan take into account the long-term country programme document and build on the integrated service provision envisaged?
- Was UNICEF Malawi response able to quickly learn from and apply best practices used by other organizations, neighbouring UNICEF country offices and/or Eastern and Southern Africa Regional Office in responding to the COVID-19 crisis?

---

[15] A variety of terms are used to describe some or all of such evaluations including mid-term, end-of-term and ex-post.

78. Almost all of the evaluations reviewed used similar data collection methods such as key informant interviews and review of different types of documents, such as grey literature, management information etc. This is in keeping with the findings of Keijzer and Oppewal in 2012 when they reviewed methodological approaches for evaluating coherence in the field of international cooperation for the European Centre for Development Policy Management (see Box 11). They reviewed 22 studies and observed that the studies most commonly used were interviews, document analysis and descriptive statistics. The study also noted that, "*analysing the respective strengths and weaknesses of different methods was challenging due to two reasons. First of all, few if any studies offered much reflection on the benefits and limitations of their methodological approaches. Secondly, methods are not intrinsically useful as the usefulness depends on the way in which the method was applied*".

79. The vast majority of evaluations reviewed analyse and present findings on coherence using descriptive and analytical narrative. While this is undoubtedly useful, there has been interest in supplementing this with other methods and tools. Some of these will briefly be reviewed here:

- In 2012, the European Centre for Development Policy Management (ECDPM) was commissioned by the Dutch and German governments to examine the feasibility and potential design of a "*development-friendliness*" index to evaluate non-aid donor policies affecting developing countries (King et al., 2012). It considered a number of political and technical considerations and concluded that a political agreement about a coherence or development friendliness index was a necessary first step. It is unclear what progress if any has been made on such an agreement since this study was commissioned. Given this, it appears unlikely that any such index is available which might prove useful to an evaluator of the coherence of COVID responses.

- The study by King et al (2012), did refer to the Commitment to Development Index (CDI) maintained by the Center for Global Development (CGD, 2021). This was commented on positively by some respondents and described as a system of composite indicators as compared to a portfolio of indicators. One of the main advantages of the composite approach is for presentational purposes. The CDI assigns points in seven policy areas – aid, trade, investment, migration, environment, security and technology (see Figure 9). The seven components are averaged to give a final score. The CDI might be useful to evaluators in a number of ways. First, it provides useful contextual data. Second, it identifies seven policy areas that might be useful to en evaluator seeking to assess coherence across whole-of government. Finally, aspects of the detailed methodology (Robinson et al, 2020) might be useful but potentially less so than the first two uses.

Source: CGD

Figure 9: Areas covered in the Commitment to Development Index

- In 2013, OECD and ECDPM produced *a methodology for country-level impact assessments of policy coherence for development on food security*. This essentially provided an analytical framework consisting of a set of principles and five modules namely important considerations when launching the impact assessment, deconstructing food security to develop country profile, linking that country profile to aid and non-aid policies, empirical research at country level and developing a communication strategy and follow-up actions. While this method is specific to food security, it could be adapted and potentially this five module (or five step) approach could be useful for an evaluation seeking to assess how coherent approaches to COVID-19 have been with the aid and non-aid policies of bilateral development partners and their governments.

- In 2007, Duraiappah and Bhardwaj produced a paper for the International Institute for Sustainable Development on measuring policy coherence among the Multilateral Environmental Agreements and Millenium Development Goals. They **analysed the degeee of coherence between policies by identifying key concepts in each policy through a process of content analysis and then assessing how often those concepts are mentioned in different policies through creating a policy coherence matrix.** The paper lacks any consideration of limitations of this approach and in our context they may be considerable. The method appears to be most suited to analysing policies on paper and it is unclear how it would be used to assess implementation of policy in practice or non-policy initiatives. Also, it seems to be based on the idea that a policy that mentions a concept from another policy more often is more coherent with that policy than one that mentions the concept less often. Is that true? Even if it is, could the method be gamed and might it create an incentive for longer documents? On balance, it seems that this approach is likely to be of limited value in the context we are considering here. However, if there were to be an evaluation of a COVID policy and there was interest in the degree of coherence with other policies, this approach could potentially be of some value.

- In 2020, the Dara evaluation of IOM's regional response to the flows of refugees and migrants from Venezuela (Bugnion and Durand, 2020) explored the issue of 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. In reporting, they used a simple **traffic light (red, amber, green) system** which they applied to individual criteria per country and which they referred to as a "*benchmark*". They also had a detailed analytical narrative and the traffic light system certainly helps as a communication tool to quickly identify which areas were going well in which countries and which needed attention. It would have been better if the basis for the different gradings and the recommended remedial actions were clearer. While there are limitations of such systems, they can be useful for communication purposes and they can be nuanced beyond the system used here. For example, ICAI uses a four-level system (green, green-amber, amber-red, red) in its reviews to illustrate findings based on a rubric-based evaluative approach (see ICAI, 2014[16]).

- In 2019, an evaluation of forced displacement and Finnish Development Policy (Zetter et al, 2019 and Box 12) used nine *judgement criteria* as opposed to evaluation sub-questions. So, for example, one of the judgement criteria for policy coherence is that mechanisms to promote policy coherence within the Ministry for Foreign Affairs are in place and operate effectively. The evaluators then seek evidence to support whether this criterion is met or not from various data sources including document analysis, key informant interviews and case studies.

---

[16] Although this did not have a particularly strong focus on coherence.

- In 2019, J Burrett of Haiku Analytics conducted a nexus system mapping for Global Affairs Canada in the Democratic Republic of the Congo (Burrett, 2019). This used *social network analysis* and produced a number of network maps which helped visualise and analyse networks involved in the humanitarian system in the country.

- OECD has a website which provides a *policy coherence for sustainable development toolkit* (OECD, undated). This contains:

    - Guidance across three core pillars (strategic vision, commitment and leadership; coordinated action across sectors and government levels; and impacts and informed decision-making) and eight building blocks (political commitment; strategic long-term vision; policy integration; policy coordination; local and regional involvement; stakeholder engagement; policy and financing impacts; and monitoring, reporting and evaluation)
    - Self-assessment checklists with two to five questions across each of eight building blocks. For evaluations with a focus on policy coherence, these questions may be useful to consider
    - Good practice examples organised by country and building block
    - Tools organised by three core pillars including an integrated simulation tool abd a tool to disentangle interactions between the SDGs. These tools may be worthy of consideration in planning evaluations particularly where there is a focus on policy coherence, especially where this relates to broader themes such as the SDGs and the Paris Climate Agreement
    - Sections on analysis, institutional mechanisms, monitoring and thematic areas. The first three sections contain sets of potentially useful questions. The final section links to a report entitled Better Policies for Sustainable Development 2016

- Bond and Coherent Europe for Sustainable Development (undated) have produced *a toolkit for being an effective policy coherence in sustainable development watchdog*. While this is focused on PCSD watchdogs, some of the material, e.g. the questions for assessing policy impact are potentially useful for evaluations as are the suggestions for data sources (see Figure 10).

Figure 10: Questions for assessing impact of policy and suggested data sources for policy coherence in sustainable development watchdogs

Source: Bond

Note: these images are graphics so the links are not live. To access these, please follow the link to the Bond website.

80. In general, the evaluations identified did not use quantitative, statistical methods and where they did (for example IPA 2020a/b), these methods were not used to evaluate coherence. Given the nature of coherence and the questions being asked, it seems unlikely that quantitative statistical methods would be of much value. In addition, no studies were identified that explicitly used methods, such as contribution analysis.

81. Table 7 seeks to assess the relative strengths and weaknesses of the tools and methods identified along with some suggestions as to where these might be useful. However, the same constraints apply that were noted by Keijzer and Oppewal in 2012 (see paragraph 78).

Table 7: Summary of emerging reflections about strengths and weaknesses of approaches, tools and methods to evaluate coherence including approaches used by participants of the Coalition

Method/Tool	Strengths	Weaknesses	Potential situations for use
Indices	Allows quantification of qualitative data Allows comparisons between organisations and countries Visually compelling	Those being scored often dispute the process and the scores Not always clear why certain criteria or indicators have been chosen nore how the scores have been weighted or calculated	CDI could be useful as contextual data and seven policy areas may be useful
Policy coherence matrix	Gives appearance of rigour by generating quantitative data	No consideration of limitations Could be easily gamed Of limited relevance for assessing policies in practice or non-policy initiatives Based on assumption that the frequency of mention of a concept is linked to coherence Incentivises longer documents	Limited
Evaluative Rubrics: Applying criteria with or without RAG ratings A rubric is a framework that sets out criteria and standards for different	Credible way of assessing performance against agreed criteria and rating scales which are defined by stakeholders – when method is explained and justified	May be criticised as subjective Those being assessed may dispute the process. Some degree of self-assessment may mitigate this but may increase	Judgement criteria could probably be widely applicable to evaluations of coherence. RAG ratings may be useful to communicate findings

levels of performance and describes what performance would look like at each level.	appropriately and in detail Visually compelling	perceptions of subjectivity Requires evaluation commissioner to support an evaluation approach with scales and assessments driven by stakeholders.	succinctly, e.g. in policy briefs
Social network analysis This involves quantitative and qualitative analysis of a social network to measure and map the flow of relationships and relationship changes between individual actors, organisations and systems.	Useful evaluation tool when wanting to understand networks and connections Visually compelling	Requires evaluators with the appropriate skills and suitable software	Probably could be used more, particularly where relationships and interactions need to be understood so potentially could be very useful in evaluating coherence
Real time evaluation A type of participatory evaluation that intends to provide immediate (real-time) feedback about delivery and results during fieldwork#_ftn1[1] Examples: Evaluation of ENABEL's response to the COVID-19 pandemic Asian Development Bank's response to the COVID-19 pandemic	Provides rapid evidence about needs and outcomes to inform decision-making[17] and programme management. A dynamic tool that can supports review of implementation and identification of appropriate solutions.	Requires the evaluation commissioner being comfortable with questions emerging from data and working iteratively. Risks providing partial evidence about the coherence of an intervention which in turn may provide decision-makers with an incomplete and mis-leading understanding of full data sets.	To gather real-time or rapid feedback about the coherence of implementation, early results and lessons learned. Findings can subsequently be used to inform adaptive management, review, reflection and management

---

[17] https://odihpn.org/magazine/real-time-evaluations-contributing-to-system-wide-learning-and-accountability/

Syntheses A broad overarching term to describe various approaches to combining, integrating, and synthesizing different research and/or evaluation findings[18] COVID-19 Global Evaluation Coalition's evaluation of responses to the pandemic	Provides a comprehensive understanding about delivery, outcomes and lessons learned from a diverse range of contexts in comparison to individual evaluations Provide comparative analysis of the coherence of different approaches in different contexts.	Does not provide real-time learning or data about coherence to inform programme management and adaptation.	To support the development of plans that maximise opportunities to be coherent, based on comprehensive evidence about what has worked well, less well, how and why in different contexts.

Ways of working

82. In principle, it seems reasonable that evaluations conducted jointly by multiple actors would be a good way of assessing coherence and could in themselves be seen as modelling coherence of actions by evaluation commissioners. While it does not follow that joint evaluations will always look at coherence (see Box 24), it seems that perhaps they have greater potential to do so given that commissioners have decided to join up with others rather than just commissioning a single-agency evaluation. Certainly, this was the conclusion of an ALNAP review of the advantages and disadvantages of joint evaluations. One of the advantages of joint evaluations, as identified by participants in an ALNAP meeting on the topic, was that they allow broader questions to be answered than can be done by one actor alone and the examples of coordination and coherence were given specifically (see Table 8).

---

[18] Schich-Makaroff, K; MacDonald, M; Plummer, M; Burgess, J; Neander, W, (2016) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5690272/

83. There are examples of joint evaluations which have looked at matters of coherence (e.g. Telford et al, 2006; Wood et al, 2008; UNICEF and UNFPA, 2020) or which are being planned or implemented with respect to COVID-19 (e.g. UNHCR et al, 2020; Freeman et al, 2021). More details of these are given in Box 24.

Box 24: Examples of joint evaluations including evaluation of coherence

In 2006, John Telford and his colleagues produced a synthesis report for the Tsunami Evaluation Coalition based on the joint evaluation of the international response to the Indian Ocean tsunami. Coherence was a central part of some of the recommendations, for example that all actors should strive to increase their disaster response capacities and to improve the linkages and coherence between themselves and other actors in the international disaster response system, including those from the affected countries themselves (see Box 6).

In 2008, Bernard Wood and others produced a synthesis report of the first phase of the evaluation of the implementation of the Paris Declaration. While understandably this focused more on alignment and harmonisation (as the five commitments of the declaration were ownership, alignment, harmonisation, managing for results and mutual accountability), there was also a section on coherence under the general heading of partner country assessments of the Paris Declaration as a tool for aid effectiveness and this focused on the internal coherence of the Declaration itself. This focus was reflected in the evaluation questions on coherence which were:

- What is the perceived *coherence* of the PD?
- Are any of the commitments and indicators and/or derived implications experienced as contradictory?

But, there were also some broader questions on coherence, such as "*do the Development Partners work coherently to support nationally led development frameworks such as the national plan, Poverty Reduction Strategy or United Nations Development Assistance Framework?*" There were some scattered findings related to coherence throughout the report. For example, the Bolivia evaluation highlighted the value of basket funding to increase coherence between policies, expenditures and real results.

In 2020, UNICEF and UNFPA conducted a joint evaluation of a joint programme on female genital mutilation. While it did not explicitly include coherence as one of the evaluation criteria, it did include coordination and this was defined as the "*extent to which the cooperation between United Nations agencies, national partners and implementing partners has been optimized to support efficient and effective implementation and expanded reach and influence of the overall programme to reach those furthest behind*". Based on this, it does seem this criterion has more to do with efficiency and effectiveness than coherence. Indeed, for the questions, coordination was combined with effectiveness and the relevant question was "*to what extent do the Joint Programme country, regional and global initiatives and its holistic approach create synergies that accelerate efforts to end FGM?*"

With respect to COVID-19, there is a joint evaluation planned by UNHCR, the Ministry for Foreign Affairs of Finland and the COVID-19 Global Evaluation Coalition (see Box 17). While it is possible that this evaluation may assess coherence, the concept note does not state this explicitly and the language of the concept note is predominantly framed in terms of effectiveness. The evaluation of the UN COVID-19 Response and Recovery Multi-Partner Trust Fund is a joint evaluation and one of the three main areas of investigation (see Box 24).

---

84. As noted above, following the joint evaluation of the response to the Indian Ocean Tsunami, ALNAP (2006) convened a half-day workshop looking at the pros and cons of joint evaluation in the humanitarian sector. This is summarised in Table 8. Some guidance for evaluations in the light of the COVID-19 pandemic recommends greater use of joint evaluations (Office of Internal Oversight Services Inspection and Evaluation Division, 2020a) with the intention of reducing burden on those being evaluated and of increasing efficiencies. It does appear that more COVID-19 evaluations are now being conducted jointly based on data available through the OECD landscaping exercise discussed earlier (see paragraph 41).

Table 8: Pros and cons of joint evaluations in the humanitarian sector as identified in a half-day ALNAP workshop following the joint evaluation of the response to the Indian Ocean Tsunami

Pros of joint evaluation	Cons of joint evaluation
• Broader scope: answers questions that cannot be addressed by one actor alone, such as coordination and coherence; also enables sensitive issues to be addressed. • Objectivity and legitimacy: increased weight of the evaluation if it is undertaken with partners. • Advocacy tool: opportunity to influence at the highest level, and can contribute to ongoing reform initiatives • Rigour: joint evaluations generally demand a higher water mark of rigour than single agency evaluations. • Attribution: it is usually easier to capture attribution in a joint evaluation. • Efficiency: rationalisation, harmonisation and reduced transaction costs for all partners (except the lead agency). • Participation and alignment between agencies: there is an opportunity for peer review and peer learning. • Evaluation capacity: it is a way of developing evaluation capacity within the sector. • Beneficiary voice: opportunities for doing large beneficiary surveys are usually greater in joint evaluations than single agency ones. • Social capital: builds social capital amongst the agencies involved.	• Complexity: can hamper joint evaluations of humanitarian assistance. • Lack of agreed common standards: this can get in the way of joint evaluation within the humanitarian sector. • Time: it takes much longer to plan and execute joint vs single agency evaluations. • Management: it requires a complicated management structure to work. • Transaction costs: these are usually higher for participating agencies (especially for the lead agency) compared with single agency evaluations, although this needs to be balanced against the value that participating agencies gain from the exercise. • Focus of recommendations: need to guard against inadequately targeted recommendations which can reduce their impact. • Detail required for a single agency: a joint evaluation may not provide the detail required to fulfil accountability requirements on the part of a single agency.

85. Synthesis approaches may be used as part of joint evaluations (e.g. Telford et al, 2006; Wood et al, 2008) or where joint evaluations have not been possible (Darcy, 2016). In the latter case, this synthesis sought to pull together the findings of 24 publicly available evaluative studies concerning the international response to the Syria crisis. Most (18) of these studies were from individual agencies although some were joint studies. This synthesis was able to pull together findings in a wide range of areas. Synthesis may be a useful approach within a particular evaluation and it may be useful as an evaluative study in its own right to pull together findings from different evaluations. However, synthesis is only possible where evaluations have already collected data in ways which permit comparative and analytical review, aggregation and comparison. Table 9 presents practical steps to support the development of evaluations and evaluation portfolios that allow synthesis to occur.

86. In conclusion, it may be helpful for those planning a COVID-19 evaluation to ask if it is possible to do this jointly with others particularly if part of the evaluation is to focus on coherence of responses. In situations where this is not possible, it might be helpful to incorporate some form of synthesis where findings from different evaluations conducted by different agencies are pulled together.

87. **Practical steps to support the development of evaluations and evaluation portfolios that can be synthesised**

Designing evaluations in ways that can be included in future synthesis studies.

What to do	How	Why
Develop and apply a common analytical framework to inform the scope of evaluations	Work collaboratively to develop a common analytical framework that respond to gaps in global evidence and collective interests. This should inform the high level direction and focus of evaluations. Specific and contextually relevant evaluation questions and plans can be subsequently developed.	Supports the generation of evaluative evidence that can be synthesised based on these questions. Without an analytical framework there is a high risk that evaluative evidence will be disparate and not conducive to comprehensive analysis.
Clarify how quality of evidence will be assessed and/or minimum quality standards will be applied to future syntheses	Adhering to relevant global, national and/or institutional evaluation standards.	Supports development of high quality evidence that is conducive to synthesis as it meets recognised standards.
Draft evaluation reports that document the evaluation approach, methods of data collection and ways of working.	Ensure evaluation reports document strengths and limitations of the study.	Supports analysis and review of the quality of evidence and informed decisions about if and how evidence is included within a study.
Support sharing and understanding of evaluation plans across agencies evaluating specific topics and geographical areas	Develop online portal or system for sharing evaluation plans	Supports complementarity rather than duplication of evaluations and development of a comprehensive evidence base.
Develop work plans that include sufficient time and facilitate collaboration and participatory reflection across different agencies	Plan timelines and ways of working that facilitate collaboration, participatory reflection, co-design of synthesis plans and scope of work.	Allows opportunities to co-create syntheses to meet needs of audiences and ensure evidence is interpreted accurately.

94: Equity and inclusion

The OECD/DAC (2021) publication 'Applying Evaluation Criteria Thoughtfully' emphasises the importance of considering inclusion when developing evaluation plans focused on coherence. This document notes that coherence can be considered in relation to human rights commitments, norms and standards as well as compatibility with inclusion and equality norms and standards. The Coalition's

strategic evaluation questions also highlight the importance of the leave no one behind agenda and understanding if and how responses to and recovery from the pandemic have supported the most marginalised and vulnerable groups. Reflecting on the relevance of the leave no one behind agenda, inclusion, equity within evaluation design, implementation and communication of findings will support evaluation teams to design and deliver ethical and credible evaluations that focus on coherence.

The Terms of Reference for a COVID-19 Global Evaluation Coalition evaluation about the protection of the rights of refugees during the pandemic [19] includes coherence as one of the three areas of evaluative inquiry:

'To what extent have national government, development partners and global responses aligned to ensure coherent approaches for the international protection of refuges during COVID-19 at the global, regional and country levels? To what extent was there synergy and coherence across the humanitarian/development/peace nexus? What were the drivers and barriers to alignment?'

These questions demonstrate how the criteria of coherence can be applied and interpreted to explore if and how vulnerable and marginalised groups have been protected. The international refugee protection regime is the framework for this evaluation.

Evaluators should also be explicit about whose definitions and interpretation of coherence informs design and implementation of activities, which groups have been consulted about their experiences of efforts to ensure coherence of strategies and activities and be explicit about any assumptions made.

Discussion

The Terms of Reference identifies a number of questions for the study to respond to. This section reflects on key findings and identifies a small number of recommendations for participants of the COVID-19 Global Evaluation Coalition to consider in their future work.

What are the most relevant evaluation questions related to Coherence, and which stakeholders are interested in these questions?

The most relevant evaluation questions have been identified throughout the document and are summarised in Table 5. In almost all cases, these are content- and context-specific applications of the strategic questionidentified by the COVID-19 Global Evaluation Coalition, namely "*to what extent are responses aligning to ensure coherent approaches…*".

Based on the work carried out for this study, it seems that a range of stakeholders, including bilateral and multilateral development partners, are interested in questions of coherence as evidenced by them including questions on coherence in their evaluations. There has been particular interest in coherence among humanitarian actors. Some interest in coherence appears to have been related to commitment to the principles of the Paris Declaration on Aid Effectiveness. So, where actors' interest in that declaration has waned, so may interest in questions of coherence. In terms of interest in particular types of questions on coherence, there is some evidence that actors, particularly funders, may be most interested in how they can involve others to increase their results, although this paper argues that questions related to this are about effectiveness and not coherence. Similarly, some actors

---

[19] ToR-Refugees.pdf (covid19-evaluation-coalition.org)

wish to limit consideration of coherence to areas they can directly control. However, this is potentially contradictory thinking as the very nature of coherence means that it is likely to mean thinking about activities and approaches beyond one's own control, either within another department of the same instituion of with other external actors. A key element in ensuring coherence may be recognising the importance of influencing (and being influenced by) others and not only focusing on activities or approaches that are directly controlled or funded. The OECD DAC definition of coherence as an evaluation criterion and the strategic question on coherence proposed by the COVID-19 Global Coaltion on Evaluation are both useful in framing issues of coherence for stakeholders who wish theur evaluations to consider this criterion. Some important stakeholders are largely absent from the evaluations reviewed. These include national and local governments and the private sector. While civil society evaluations are included, they are pretty separate from other evaluations conducted by bilateral and multilateral development partners.

There would be scope to determine specifically which stakeholders are interested in which of the questions identified in Table 5. This could be done, for example, by circulating Table 5 to stakeholders (bilateral development partners, multilateral development partners, civil society, academic institutions, private sector, national and local governments etc) asking them to rank their interest in the relevant questions. Of course, such interest may not be fixed and may vary depending on particular contexts.

Which coherence issues are raised at different levels of analysis: institutional, country-level and global?

Attempts have been made to summarise these in Table 5. Briefly, if there is interest in understanding coherence within an institution, there is need to understand the boundaries of that institution and whether all stakeholders understand those boundaries in the same way. For example, in the case of a bilateral development partner, is the institution, the relevant ministry of the whole-of-government? In the case of a UN agency, is the institution defined as that agency's country office, the entirety of that agency or the UN as a whole? Key coherence issues to consider within bilateral development partners as institutions in relation to COVID may be:

- To what extent are the domestic COVID-19 responses of governments coherent with their support for international COVID-19 responses and vice versa?
- To what extent are the different elements of support to the international COVID-19 response provided by governments coherent with each other?
- To what extent are activities (policies/interventions/financial support etc.) of governments coherent with their support for international COVID-19 responses and vice versa?
- To what extent do the governments have a coherent whole-of-government approach to COVID-19?

In terms of coherence at country level, perhaps the main coherence issues relate to who is doing what in relation to COVID-19 and how this fits together. This requires thinking about COVID-specific responses but also about other activities, policies etc. that may not appear to be specifically about COVID but have an effect on the coherence of COVID responses nonetheless. The specificities of this issue are likely to differ between low-, middle- and high-income countries. Key coherence issues to consider within bilateral development partners in relation to aid provided to other countries to support COVID responses may include:

- To what extent is there coherence across different aid modalities, for example, humanitarian and development aid, bilateral aid allocated centrally as compared to bilateral aid allocated by a country office, multilateral as compared to bilateral aid?

- To what extent is there coherence between aid spending and other activities of the donor governments, e.g. diplomacy, trade, security, health etc.?
- Who are the other key actors in particular countries with whom coherence should be assured (e.g. national and local government, other development partners, civil society, academic institutions, private sector, communities etc.)? To what extent is there coherence between particular COVID responses and other relevant responses?
- To what extent is there coherence within and beyond the COVID-19 response? For example, are there other actions/responses/policies/interventions beyond specific COVID-19 responses which have had an effect, positive or negative, on COVID responses? Similarly, have COVID-19 responses had an effect, positive or negative on other actions/responses/policies/ interventions beyond specific COVID-19 responses?

In terms of coherence at global level, perhaps the main coherence issues relate to how everything that is being done fits together. This includes not only what might be considered aid and humanitarian efforts but also efforts to support domestic responses in high-income countries and what might be seen as non-COVID interventions. Key coherence issues to consider within bilateral development partners in relation to the global level may include:

- To what extent is there coherence in terms of engagement of governments on the "*global stage*"? This might include engagement specific to COVID and international development but may also be broader.
- Who are the other key actors globally with whom coherence should be assessed (e.g. United Nations, G7, G20, OECD etc.)? To what extent is there coherence between government responses and other relevant responses?
- To what extent is there coherence within and beyond the COVID-19 response? For example, are there other actions/responses/policies/interventions beyond specific COVID-19 responses which have had an effect, positive or negative, on COVID responses? Similarly, have COVID-19 responses had an effect, positive or negative on other actions/responses/policies/ interventions beyond specific COVID-19 responses?

More detail is provided in Table including of levels not specified in the question, e.g. regional level.

What aspects of the COVID-19 response and recovery effort (the immediate health response, secondary effects including on education and livelihoods, or building back sustainably and equitably) are the most pertinent when it comes to evaluating the coherence criterion?

All of these are relevant and the precise balance between these will vary depending on the specific context of an evaluation. However, in an evaluation of the overall COVID response by a group of bilateral development agencies, it may be good to start with the immediate health response – given that COVID is a disease so is, at its roots, a health issue - and then look out to other sectors. In general, evidence of focus on coherence within and across sectors was difficult to ascertain from the evaluations reviewed from this study and these may be areas to which future evaluations wish to pay particular attention.

What are the boundaries between coherence and other criteria including relevance and effectiveness?

These matters are covered in Table 1. In brief:

- While consideration of fit between national government interventions and interventions of other actors may be considered to be coherence, consideration of alignment to national

policies falls under, and forms part of the assessment of, the OECD DAC evaluation criterion of *relevance*.

- While many aspects of partnership and coordination fall under external coherence. However, where the focus on partnership and coordination is focused on maximising an intervention's results, this study considers such partnership and coordination to be instrumental and part of the OECD DAC criterion of *effectiveness*.
- Connectedness between the humanitarian, development and peace nexus is considered to be an element of coherence but this also relates to the OECD DAC criterion of *sustainability*.

There are some other elements which may have been treated as separate criteria previously, for example in the humanitarian field, but this study considers as part of coherence based on the OECD DAC definition. These include *connectedness* and *coordination*.

Which evaluation approaches and methods will enable agencies to evaluate coherence – and answer the identified questions – in ways that are meaningful, feasible and manageable?

It is difficult to answer this question definitively for the same reasons given to ECDPM in 2012, namely "*analysing the respective strengths and weaknesses of different methods was challenging due to two reasons. First of all, few if any studies offered much reflection on the benefits and limitations of their methodological approaches. Secondly, methods are not intrinsically useful as the usefulness depends on the way in which the method was applied*" (Keijzer and Oppewal, 2012 and see paragraph 78). However, subject to that caveat, the following observations can be made:

- If evaluations are seeking to obtain data quickly so that management decisions and course corrections can be made, real-time evaluations may be useful. As a methodology, much could be learned about real-time evaluations from the humanitarian field. Real-time evaluations could probably be used more in development interventions and not only in the humanitarian field.
- There is unlikely to be much scope for quantitative, statistical methods in evaluations of coherence. Evaluations of coherence are likely to continue to rely predominantly on rigorous and well-documented qualitative methods including review of secondary data and collection of primary data from key informant interviews.
- There may be scope for more use of evaluative rubrics which identify criteria based on the perspectives of stakeholders to seek to bring more rigour to such qualitative methods. These could be combined with RAG ratings where there is a need to communicate with senior people in visually-compelling ways.
- There is potentially scope for more use of social network analysis where there is need to map and understand interactions and relationships between entities, including people and organisations.

Is there sufficient, relevant data being collected and/or available to evaluate coherence of the response to the pandemic (dimensions identified above)?

It is difficult to assess this at this stage because not many COVID-related evaluations have yet been completed and it is difficult to get full information about what is being planned and what will be delivered.

Nevertheless, two observations can be made. First, it is not only about collecting sufficient data but when such data collection takes place and how the information generated is used. There is probably scope for more information soon which is another argument in favour of real-time evaluations.

Second, perhaps the biggest problem remains the one noted by Darcy and Dillon that "*the great majority* [of humanitarian evaluations[20]] *deal with context-specific crisis responses by individual agencies; relatively few are concerned with system-wide performance or organisational performance across a range of different contexts*". A key issue would be to try to get evaluations to focus less on micro projects and more on big picture issues. It is disappointing for example that no evaluations of national responses to COVID have been carried out and none were identified in the study. In addition, it is unclear if the United Nations are planning to evaluate the totality of their responses to COVID, e.g. through some form of synthesis approach. If that were done, it could be an essential piece in seeking to evaluate the coherence of the overall, international response to COVID. If OECD and some bilateral development agencies do decide to conduct an overall evaluation of the coherence of responses to COVID, it would be great if this could seek to evaluate (or synthesise evidence of) the extent to which the global response to COVID has been coherent and the extent to which the response to COVID has been coherent in some case study countries.

Challenges of evaluating coherence

First, one of the main challenges to evaluating coherence as a criterion has been the lack of a clear, shared understanding of what coherence is. Clearly, this should be helped by having the OECD DAC definition but there may be challenges because this differs from earlier definitions, e.g. in the humanitarian space. In addition, there may be need to publicise and explain the definition more. There are quite a lot of questions around as to what the definition means in particular settings and simply referring back to the definition may not be sufficient. Developing a set of frequently asked questions may be one option.

The second challenge will be getting evaluations to include and use the criterion of coherence. It is striking that, even in the humanitarian space, most evaluations (around two thirds according to Darcy and Dillon in 2020) have not evaluated coherence. So, just because there is a criterion on coherence does not necessarily mean evaluations will use it.

Third, perhaps the biggest challenge is what might be called "*micro*" thinking that is wanting to focus on a particular project and/or on areas that the commissioning agency can control directly. Again, this has been a common finding when exploring these matters. Such thinking is a major barrier to evaluating coherence and will end up with much evaluation of "*coherence*" focused on how partnerships, coordination etc. can be set up to maximise an intervention's results, i.e. to increase its

---

[20] Although this would apply to different types of evaluation reviewed for this study.

effectiveness. There is a pressing need for "*big picture*" studies, for example an evaluation of the coherence of the entire global response to COVID and/or evaluations of the entire response to COVID in particular countries.

Finally, joint evaluations and approaches to synthesis, based on common frameworks and people-based methods, offer the best ways of overcoming these challenges. .

Recommendations

CR1. The newly-adopted OECDC DAC evaluation criterion on coherence provides a good basis for evaluations wishing to assess the coherence of responses to COVID-19. The criterion is helpful as it emphasises the importance of evaluating an intervention's fit and the delineation of internal and external coherence is particularly useful for evaluations of well-defined entities. However, the distinction may be less useful where the entity being evaluated is complex or amorphous, e.g. a United Nations agency or where an evaluation is not of a specific entity, e.g. a thematic evaluation.

CR2. Evaluations wishing to include the coherence criterion should first identify what elements of coherence will be evaluated in the specific context under consideration. This will involve understanding the entity or entities being evaluated and identifying other relevant actors. It will also involve being clear as to "*coherence with what*" is being evaluated, for example, as illustrated in Figure 3. Is the evaluation interested in fit with other interventions in the same sector or another sector? Or is the main focus on how the intervention fits with what other parts of the entity are doing? Or is the main focus on fit with broader themes and topics, such as the SDGs, the humanitarian-peace-development triple nexus, human rights, equity and inclusion.

CR3. Working under the umbrella of the Coalition's strategic question on coherence, each evaluation will need to carefully develop evaluation questions that are relevant to the intervention(s) being evaluated and the context. The focus of the coherence questions, including any definitions or explanations of the dimensions explored, should be made explicit to avoid confusion or misunderstanding.

CR4. Where possible evaluations should go beyond simple descriptive and analytical narratives as a way of describing and communicating findings concerning coherence. The use of simple tools, such as red amber green (RAG) ratings may be useful to aid communication but they have to date been used by relatively few evaluations relating to coherence.

CR5. Where possible, joint evaluations may have particular advantages in answering questions related to coherence. Where these are not possible, synthesis approaches may be useful in pulling together findings from evaluations conducted by different actors.

Annex 1: Terms of Reference

1: Background

The COVID-19 Global Evaluation Coalition

The COVID-19 Global Evaluation Coalition is an independent collaborative project. The Coalition includes the development evaluation units of countries, United Nations organisations and multilateral institutions. The Coalition provides credible evidence to inform international co-operation supporting non-clinical responses to and recovery from the COVID-19 pandemic in developing countries - helping to ensure that lessons are learned and that the global development community delivers on its promises. The Coalition will support and communicate both individual participants' evaluations, and joint work involving multiple participants. This collaborative approach will maximize synergies and learning, while reducing duplication of effort in evaluating different elements of the COVID-19 pandemic response. The DAC EvalNet serves as Secretariat for the Coalition.

The overall purpose of the COVID-19 Global Evaluation Coalition is to foster collaboration, create synergies, and avoid duplication. The Coalition will improve the speed and quality of evaluation, analysis, and communication in ways that provide credible evaluative evidence that will:

- Inform policy makers and implementers to support a more effective (collective) response to and recovery from the COVID-19 pandemic.
- Generate lessons and good practices, and facilitate learning for future global crises.
- Provide a basis for accountability of development partners, including the provision of information to the public.

Revision of the criteria definitions and inclusion of coherence

The Organisation for Economic Co-operation and Development (OECD) Development Assistance Committee (DAC) first laid out the evaluation criteria (relevance, effectiveness, efficiency, impact and sustainability) in the 1991 OECD DAC Principles for Evaluation of Development Assistance, and later defined the terms in the 2002 Glossary of Key Terms in Evaluation and Results Based Management. These five criteria have come to serve as the core reference for evaluating international development and humanitarian projects, programmes and policies. Beyond development co-operation, evaluators and commissioners also use the criteria in other areas of public policy.

Building from learning gathered over 25 years of applying the criteria, the global evaluation community began discussing revisiting the criteria following the 2015 agreement of the 2030 Agenda for Sustainable Development, including the Sustainable Development Goals (the 2030 Agenda), and the Paris Agreement within the United Nations Framework Convention on Climate Change (Paris Agreement).

The consultation and process has led to the development of new and improved criteria and the additional of one major new criterion – coherence – to better capture linkages, systems thinking, partnership dynamics, and complexity. The inclusion of coherence as one of the six OECD-DAC Evaluation criteria builds on evaluative guidance and work led by the humanitarian and peacebuilding sectors and from major international donor agencies including the European Commission and UNHCR.

2: Objective

The EvalNet Secretariat is commissioning a rapid scoping and initial evaluability assessment that responds to the following areas and questions outlined in this Terms of Reference. Using the OECD-DAC definition of "coherence", the Contractor will conduct a scoping study to support COVID-19 Global Evaluation Coalition participants to plan evaluations of the coherence of the international response to the COVID-19 pandemic. Through the assessment the Contractor will provide strategic and practical advice to support participants to develop evaluation plans that are feasible, credible and useful.

3: Context

Coherence has emerged as a key topic of interest for evaluation among participants of the Coalition, who recognize that COVID-19 is a global challenge that calls for joined up responses. This is an area that is best addressed in collaboration, which can provide broader scope and coverage, compared to individual evaluations.

The Study will apply the definition of coherence within the [OECD-DAC Evaluation Criteria](#) (OECD, 2019) and aims to support better understanding of how this criterion can be applied within evaluations. The OECD-DAC (2019) definition of coherence is:

How well does the intervention fit?

The compatibility of the intervention with other interventions in a country, sector or institution.

Note: The extent to which other interventions (particularly policies) support or undermine the intervention, and vice versa. 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.

This definition includes the following key elements of analysis:

- Synergies and interlinkages between the intervention and other interventions carried out by the same institution/government.
- Consistency of the intervention with the relevant international norms and standards to which that institution/government adheres.
- Consistency of the intervention with other actors' interventions in the same context.
- Complementarity, harmonisation and co-ordination with others
- Adding value while avoiding duplication of effort.

4: Audience and intended use of The Study

The primary audience is Evaluation Specialists working in bilateral agencies, multilateral agencies and partner countries who are participants of the COVID-19 Global Evaluation Coalition. The document will also be used by the OECD-DAC's EvalNet Secretariat to support work planning and additional research.

The Study is intended to be an initial step to support the Coalition in conceptualizing the topics, identifying evaluation needs, developing terms of reference, and planning evaluation work.

5: Tasks and scope of work

The research will involve carrying out interviews with 8-12 key stakeholders, including EvalNet members, OECD-DAC Secretariat analytical team leaders and COVID-19 Coalition participants, and a review of academic and grey literature relevant to the COVID-19 response.

The assessment shall cover the following tasks, and answer these questions:

TASK 1. Scoping the topic of coherence in the context of COVID-19 and Identification of strategic evaluation questions.

This task involves understanding what dimensions of coherence are of most interest to key evaluation stakeholders. Scoping of the topic should consider coherence in relation to the following areas. Consultants are invited to refine this list and propose additional relevant topics and questions. The paper will answer the following questions:

10. What are the most relevant evaluation questions related to Coherence, and which stakeholders are interested in these questions. Question might include, for example:
    a. If and how international responses are aligned with national (recipient) governments plans and the domestic actions of provider countries
    b. Coherence of the response with international human rights norms and standards
    c. Coherence of the response and recovery efforts with environmental sustainability and low-carbon transition.
    d. Whether and how there has been a coherent response across the humanitarian, development, peacebuilding nexus, and across government in fragile contexts.
    e. If, how and why coherence has led to more efficient and effective responses
    f. Whether coherence has contributed to more agile and adaptive global responses
11. Which coherence issues are raised at different levels of analysis: institutional, country-level and global?
12. What aspects of the COVID-19 response and recovery effort (the immediate health response, secondary effects including on education and livelihoods, or building back sustainably and equitably) are the most pertinent when it comes to evaluating the coherence criterion?
13. What are the boundaries between coherence and other criteria, including relevance and effectiveness?

TASK 2. Analysis of feasibility of answering the identified questions

This task involves exploring the extent to which the questions identified (TASK 1) can feasibly be answered through evaluation. The Paper will answer the following questions:

14. Which evaluation approaches and methods will enable agencies to evaluate coherence – and answer the identified questions – in ways that are meaningful, feasible and manageable?
15. Is there sufficient, relevant data being collected and/or available to evaluate coherence of the response to the pandemic (dimensions identified above)?
16. Which data systems and resources need to be developed to maximise the feasibility and success of evaluations?

TASK 3. Identification of appropriate, credible processes and ways of working

Building on Tasks 1 and 2, this task involves looking at evaluation processes, and making recommendations to the Coalition participants on ways forward. This may include identifying or proposing several options for ways of working and processes for future evaluations. The Paper will answer:

17. Which overall ways of working will support Coalition participants to design credible, ethical, timely and appropriate evaluations?
18. Which processes will ensure plans and deliverables are inclusive, credible and maximise opportunities for audiences to use findings to inform decisions?

6: Methodology

It is expected that the contractor will complete a review and analysis of relevant documents, lead and analyse interviews with experts, facilitate a discussion about emerging findings with a small number of participants of the COVID-19 Global Evaluation Coalition and subsequently drafting of the paper.

We expect the review to cover:

- Organisations: Bilateral donors, multilateral regional banks, UN organisations including COVID-19 multi-partner trust funds and other COVID-19 related special funds. Institutions will be included selectively (purposive sample) to ensure coverage of the key issues and experiences, rather than complete coverage.
- Documents: available strategies, institutional policies, statements, papers and articles.
- Interviewees: The Contractor will work with the OECD-DAC Secretariat to identify 8-12 interviewees. These should include experts such as senior practitioners working in development co-operation agencies and in partner countries; and evaluation experts in the COVID-19 Coalition. Interviewees should have an in-depth understanding of trends and narratives in the COVID-19 response as it relates to the topic of coherence.

Outside of scope: The Contractor is *not* expected to begin assessing coherence or applying the coherence criteria for evaluative analysis, nor to gather evidence on the extent of coherence.

Deliverables and Indicative performance schedule

Note: Timeframe will be adjusted according to the start date of the contract

Deliverable 1. One page summary of inception/kick off call	25 January 2021
Deliverable 2. List of proposed interviewees	
Deliverable 3. First draft Paper, which includes description of methodology and analysis of data, main findings and structure of discussion at EvalNet/Coalition meeting	1st March 2021
Deliverable 4. Presentation of the draft at EvalNet/Coalition meeting and facilitation of discussion to support finalisation of paper	31st March 2021
Deliverable 5. Final draft Paper (25- 30 pages with 1 page summary)	15th April 2021

6: Standards for deliverables

- The Paper will be approximately 25-30 pages long, including a one page, clear summary of key findings and conclusions. Additional information can be included in annexes.
- It will be written in British English in clear language suitable for a technical audience, avoiding unnecessary jargon and abbreviations.
- It will adhere to evaluation and research quality standards, including describing methods and any limitations (and the implications of these on the findings and conclusions). There should be a logical flow from findings to conclusions and recommendations.
- The paper and references will be drafted in accordance to the OECD Style Guide (https://www.oecd.org/about/publishing/OECD-Style-Guide-Third-Edition.pdf).

Annex 2: People Interviewed

Ole Andersen, Danish Institute for International Studies
Catherine Anderson, OECD
Kevin Andrews, Foreign and Commonwealth Development Office (FCDO)
Julia Betts, Consultant
Wendy Asbeek Brusse, Netherlands Ministry of Foreign Affairs
Alexandra Chambrel, UNFPA
James Darcy, Consultant
Kelly David, OCHA
Richard Jones, UNDP
Megan Kennedy-Chouane, EvalNet Secretariat
Anne-Claire Luzot, WFP
Susanna Morrison-Metois, ALNAP
Santhosh Persaud, OECD
Alison Pollard, EvalNet Secretariat
Mathew Varghese, United Nations
Anna Williams, Consultant

Annex 3: Documents Reviewed

ACAPS (2020) *ACAPS COVID-19 Analytical Framework* available on
https://www.acaps.org/sites/acaps/files/resources/files/20200327_acaps_covid-19_analytical_framework_0.pdf (accessed 10.02.21)

ADB Independent Evaluation (2020a) *Responding to COVID-19: Lessons from Previous Support to Micro, Small and Medium-Sized Enterprises* available on
https://www.adb.org/sites/default/files/evaluation-document/610716/files/synthesis-note-3.pdf
(accessed 11.02.21)

ADB Independent Evaluation (2020b) *Responding to the Novel Coronavirus Crisis: 13 Lessons from Evaluation* available on https://www.adb.org/sites/default/files/evaluation-document/565391/files/ll-covid-19.pdf (accessed 11.02.21)

Alam, K. and Balthazar, E. (2011) *ActionAid's International Haiti Emergency Response Programme: Rebuilding Lives and Livelihood: A Report on Real-Time Evaluation: Final Report* available on
https://actionaid.org/sites/default/files/evaluation_of_actionaids_response_to_haiti_earthquake_-_dec_2011.pdf (accessed 05.03.21)

Ali, M. and Hutton, K. (2016) *Evaluation of the DFID Ebola Emergency Response Fund (DEERF) in Sierra Leone* available on https://aidworks.org.uk/wp-content/uploads/2014/01/DEERF-Evaluation-by-Aid-Works.pdf (accessed 05.03.21)

ALNAP (2006a) *Evaluating Humanitarian Action Using the OECD-DAC Criteria: An ALNAP Guide for Humanitarian Agencies* Available on https://www.alnap.org/help-library/evaluating-humanitarian-action-using-the-oecd-dac-criteria (accessed 15.02.21)

ALNAP (2006b) *Report on Half-Day Workshop on Joint Evaluations* available on http://www.tsunami-evaluation.org/NR/rdonlyres/C29F927D-D035-44C7-B739-9BDFF7274308/0/ReportonjointevaluationsworkshopALNAPBiannual.pdf (accessed 22.02.21)

ALNAP (2016) *Evaluation of Humanitarian Guide* available on https://www.alnap.org/help-library/evaluation-of-humanitarian-action-guide (accessed 17.05.21)

ALNAP (2018) *The State of the Humanitarian System 2018* available on https://www.alnap.org/help-library/the-state-of-the-humanitarian-system-2018-full-report (accessed 17.05.21)

ALNAP (2021) *Covid-19 Response Portal* available on https://covid19.alnap.org/ (accessed 11.02.21)

Anderson, L.R. (2016) *International Lessons from Integrated Approaches in Afghanistan* available on
http://www.oecd.org/derec/denmark/afghan_lessons_part_i.pdf (accessed 26.02.21)

Arqués, R.S., Thakwalakwa, C. and Durand, J. (2021) *Real Time Evaluation of UNICEF's Response to the COVID-19 Outbreak Crisis in Malawi* available on
https://evaluationreports.unicef.org/GetDocument?fileID=16462 (accessed 04.03.21)

Asquith, J. and Bloom, E. (2020) *Real-time Evaluation: ADB's Response to the COVID-19 Pandemic: Evaluation Approach Paper* available on https://www.adb.org/sites/default/files/evaluation-document/652326/files/eap-rte-adb-response-covid-19-pandemic.pdf (accessed 10.02.21)

Austrian Development Agency (2021) *The Whole of Government Approach in Austrian Development Policy: Evaluation and Evidence Synthesis* available on https://www.entwicklung.at/en/projects/detail-en/the-whole-of-government-approach-in-austrian-development-policy-evaluation-and-evidence-synthesis (accessed 04.03.21)

Avdeenko, A. and Heesemann, E. (2020) *Evidence on COVID-19 Pandemic Control Interventions and their Impacts on Health-Related Outcomes* available on http://www.covid19-evaluation-coalition.org/documents/Pandemic-Control-Interventions-impacts-on-Health-Related.pdf (accessed 11.02.21)

Baker, J., Palkovits, K., Abeywickrama, T., Lee, C. and Keen, P. (2015) *Real Time Evaluation of the Nepal Earthquake Response Operation* available on https://www.alnap.org/help-library/real-time-evaluation-of-the-nepal-earthquake-response-operation (accessed 05.03.21)

Barry, F., King, M. and Matthews, A. (2010) *Policy Coherence for Development: Five Challenges* Irish Studies in International Affairs: Vol 21 (2010), 207-223 available on https://www.jstor.org/stable/41413183?seq=1 (accessed 23.02.21)

Bastøe, P.Ø. Brusse, W.A. and Faust, J. (2020) *COVID-19 and Development Cooperation: We Know a Lot about What Works, Let's Use the Evidence* available on http://www.developmentresearch.eu/?p=694 (accessed 22.02.21)

BetterEvaluation (2016) *Specify the Key Evaluation Questions* available on https://www.betterevaluation.org/en/rainbow_framework/frame/specify_key_evaluation_questions (accessed 26.06.21)

BetterEvaluation (undated) *Footprint Evaluation* available on https://www.betterevaluation.org/en/themes/footprint_evaluation (accessed 27.06.21)

Bhattacharjee, A. (2005) *Real Time Evaluation of Tsunami Response in Asia and East Africa, Second Round* available on https://reliefweb.int/report/india/real-time-evaluation-tsunami-response-asia-and-east-africa-second-round-final-report (accessed 05.03.21)

Bhattarcharjee, A. and Lossio, R. (2011) *Evaluation of OCHA Response to the Haiti Earthquake: Final Report* available on https://www.unocha.org/sites/unocha/files/dms/Documents/Evaluation%20of%20OCHA%20Response%20to%20the%20Haiti%20Earthquake.pdf (accessed 05.03.21)

Bhattacharjee, A. (2017) *Islamic Relief Worldwide: Independent Evaluation of Nepal Earthquake Response* available on https://www.islamic-relief.org/wp-content/uploads/2019/01/IRW-Nepal-Response.pdf (accessed 05.03.21)

Bond and Coherent Europe for Sustainable Development (undated) *Ensuring Policy Coherence in Sustainable Development (PCSD): A Toolkit for Being an Effective PCSD Watchdog* available on https://www.bond.org.uk/sites/default/files/resource-documents/ensuring_policy_coherence_in_sustainable_development_pcsd_0.pdf (accessed 19.11.21)

Bugnion, C. and Durand, J. (2020) *Interim Evaluation of IOM's Regional Response to the Flow of Refugees and Migrants from Venezuela* available on https://daraint.org/wp-content/uploads/2020/07/DARA-IOM-Interim-Final-Evaluation-Report.pdf (accessed 09.03.21)

Burrett, J. (2019) *Democratic Republic of the Congo: Nexus System Mapping* for Global Affairs Canada International Assistance Evaluation Division

Caparini, M. and Reagan, A. (2019) *Connecting the Dots on the Triple Nexus* available on https://www.sipri.org/commentary/topical-backgrounder/2019/connecting-dots-triple-nexus (accessed 04.03.21)

Center for Global Development (2021) *The Commitment to Development Index* available on https://www.cgdev.org/cdi#/ (accessed 26.02.21)

Center of Resources for Innovation and Development (CRID) (2017) *Typhoon Haiyan Response Program: Summative Report* available on https://reliefweb.int/report/philippines/typhoon-haiyan-response-program-summative-evaluation-report (accessed 05.03.21)

Chazaly, C. and Goldman, E. (2021) *MENA Real Time Assessment: COVID 19 Response* available on https://evaluationreports.unicef.org/GetDocument?fileID=16410 (accessed 04.03.21)

Cornish, L. (2021) *Interactive: Who's Funding the COVID-19 Response and What are the Priorities?* Available on https://www.devex.com/news/interactive-who-s-funding-the-covid-19-response-and-what-are-the-priorities-96833 (accessed 23.03.21)

Coronakommissionen (2020) *The Elderly Care in the Pandemic* available on https://coronakommissionen.com/wp-content/uploads/2020/12/summary.pdf (accessed 26.02.21)

Cota (2021) *Enabel – Service Contract to Support the Evaluation of Enabel's Response to COVID-19* available on http://www.cota.be/?p=6625&lang=en (accessed 10.02.21)

COVID-19 Global Evaluation Coalition (2020a) *Lessons from Evaluation: Food Security* available on http://www.covid19-evaluation-coalition.org/documents/Lessons-from-evaluation-issue-1.pdf (accessed 11.02.21)

COVID-19 Global Evaluation Coalition (2020b) *Lessons from Evaluation: Gender Equality in Education* available on http://www.covid19-evaluation-coalition.org/documents/Lessons-from-evaluation-issue-2.pdf (accessed 11.02.21)

COVID-19 Global Evaluation Coalition (2020c) *Lessons from Evaluation: The Use of Cash Transfers in Humanitarian and Development Settings* available on http://www.covid19-evaluation-coalition.org/documents/Lesson-from-evaluation-issue-4.pdf (accessed 11.02.21)

COVID-19 Global Evaluation Coalition (2020d) *Lessons from Evaluation: Gender Equality* available on http://www.covid19-evaluation-coalition.org/documents/Lesson-from-evaluation-issue-5.pdf (accessed 11.02.21)

COVID-19 Global Evaluation Coalition (2021a) *Draft Strategic Evaluation Questions* This is a draft document for discussion

COVID-19 Global Evaluation Coalition (2021b) *Home* This is the Coalition's website available on http://www.covid19-evaluation-coalition.org/ (accessed 10.02.21). More details of the Coalition are provided on the *About* page – see http://www.covid19-evaluation-coalition.org/about/ (accessed 10.02.21)

COVID-19 Global Evaluation Coalition (2021c) *COVID-19 Evaluation Landscape: Initial Analysis for Discussion.* A presentation to a Coalition meeting held on 26 February 2021

Dara (2016) *Evaluation of UNICEF's Response and Recovery Efforts to the Gorkha Earthquake in Nepal: Final Evaluation Report* available on https://daraint.org/dara_evaluations/evaluation-unicefs-response-recovery-efforts-gorkha-earthquake-nepal/ (accessed 05.03.21)

Darcy, J. (2016) *Syria Coordinated Accountability and Lesson Learning (CALL): Evaluation Synthesis and Gap Analysis: A Synthesis and Analysis of 24 Publicly Available Evaluative Studies concerning the International Response to the Syria Crisis Commissioned by the Steering Group for Inter-Agency Humanitarian Evaluations* available on https://interagencystandingcommittee.org/system/files/ocha_syria_web_interactive_2.pdf (accessed 05.03.21)

Darcy, J. and Dillon, N. (2020) *Missing the Point? Reflections on Current Practice in Evaluating Humanitarian Action: Discussion Paper* available on https://www.alnap.org/help-library/missing-the-point-reflections-on-current-practice-in-evaluating-humanitarian-action (accessed 18.05.21)

Davidson, J. and Rowe, A. (2021) *Key Evaluation Questions to Guide Footprint Evaluations* available on https://www.betterevaluation.org/sites/default/files/Footprint%20KEQs%20v2.pdf (accessed 27.06.21)

Duraiappah, A.K. and Bhardwaj, A. (2007) *Measuring Policy Coherence among the MEAs and MDGs* A Report for the International Institute for Sustainable Development (IISD) available on https://www.iisd.org/system/files/publications/measuring_policy.pdf (accessed 19.02.21)

Dy, P. and Stephens, T. (2016) *The Typhoon Haiyan Response: Strengthening Coordination among Philippine Government, Civil Society and International Actors* available on https://www.hks.harvard.edu/sites/default/files/centers/research-initiatives/crisisleadership/files/Dy_and_Stephens.pdf (accessed 05.03.21)

ECDC (2017) *Evaluation of ECDC Ebola Deployment in Guinea* available on https://www.ecdc.europa.eu/en/publications-data/evaluation-ecdc-ebola-deployment-guinea (accessed 05.03.21)

ECDC (2020) *Monitoring and Evaluation Framework for COVID-19 Response Activities in the EU/EEA and the UK* available on https://www.ecdc.europa.eu/sites/default/files/documents/covid-19-framework-monitor-responses.pdf (accessed 10.02.21)

EPYPSA (2011) *An Evaluation of the Haiti Earthquake 2010 Meeting Shelter Needs: Issues, Achievements and Constraints* available on https://www.ifrc.org/docs/Evaluations/Evaluations2011/Global/HTShelterClusterReview11.pdf (accessed 05.03.21)

European Commission (2006) *Evaluation Methods for the European Union's External Assistance: Methodological Bases for Evaluation* available on http://www.oecd.org/dac/evaluation/dcdndep/47469160.pdf (accessed 04.03.21)

European Commission (2018) *Mid-Term Evaluation of the EU's Generalised Scheme of Preferences (GSP)* available on https://trade.ec.europa.eu/doclib/docs/2018/october/tradoc_157434.pdf (accessed 26.02.21)

European Commission (2019) *2019 EU Report on Policy Coherence for Development* Commission Staff Working Document available on https://ec.europa.eu/international-partnerships/system/files/swd-2019-20-pcdreport_en.pdf (accessed 23.02.21)

European Commission (2020a) *COVID-19: Commission Sets Out European Coordinated Response to Counter the Economic Impact of Coronavirus* available on https://ec.europa.eu/commission/presscorner/detail/en/IP_20_459 (accessed 26.02.21)

European Commission (2020b) *Fast-Track Assessment of the EU's Initial Response to COVID-19 Crisis in Partner Countries and Regions: Specific Terms of Reference – Part A*

European Commission (undated) *Policy Coherence for Development* available on https://ec.europa.eu/international-partnerships/policy-coherence-development_en (accessed 23.02.21)

European Investment Bank (2020) *Africa's Digital Solution to Tackle COVID-19* available on https://www.eib.org/attachments/country/africa_s_digital_solutions_to_tackle_covid_19_en.pdf (accessed 10.02.21)

European Union (2007) *Evaluating Coordination, Complementarity and Coherence in EU Development Policy: A Synthesis* Studies in European Development Cooperation Evaluation No. 8 available[21] on http://www.three-cs.net/images/ccc_8.pdf (accessed 25.06.21)

European Union (undated) *INTPA/ESS Initiatives* available on https://europa.eu/capacity4dev/devco-ess (accessed 26.02.21) - covers Evaluation in Crisis, Disseminating Evaluations and Evaluation in Hard-to-Reach Areas

Fellesson, M. and Román, L. (2016) *Sustaining a Development Policy_ Results and Responsibility for the Swedish Policy for Global Development* available on https://eba.se/en/rapporter/sustaining-a-development-policy-results-and-responsibility-for-the-swedish-policy-for-global-development/4505/ (accessed 26.02.21)

Ferretti S. (2021) personal communication on the topic *How to write terms of reference*

Food and Agriculture Organization of the United Nations (FAO) (2020) *Risk Analysis and Guidance for the Management and Conduct of Evaluations during International and National Level COVID-19 Crisis and Restrictions* available on http://uneval.org/document/download/3558 (accessed 19.02.21)

Freeman, T., Esser, A.L., Chatterjee, C. and Vela, P. (2021) *Early Lessons and Evaluability of the UN COVID-19 Response and Recovery MPTF: Inception Report* A report for the SG's Designate COVID19 Recover Better Fund, Executive Office of the Secretary General, United Nations

Freudenreich, H., Demmler, K.M., Fongar, A., Jäckering, L. and Rudolf, K. (2020) *Effective Interventions to Increase Food and Nutrition Security in Response to COVID-19* available on

---

[21] This study is also found on Three-Cs.net (http://www.three-cs.net/triple-c-evaluations-8-evaluating-co-ordination-complementarity-and-coherence-in-eu-development-policy-a-synthesis.html) which focuses on evaluations of coordination, complementarity and coherence. There are details of six ongoing studies here (http://www.three-cs.net/ongoing-studies.html) and a number of publications (http://www.three-cs.net/3cs-publications.html).

http://www.covid19-evaluation-coalition.org/documents/Effective-interventions-to-increase-food-nutrition-security-in-response-to-Covid-19.pdf (accessed 11.02.21)

Glenton, C. and Lewin, S. (2020) *Communicating with the Public about Vaccines: Implementation Considerations* Brief prepared for Norad. Oslo: Norwegian Institute of Public Health. October 2020. Available on http://www.covid19-evaluation-coalition.org/documents/VACCINES-Brief-1.pdf (accessed 11.02.21)

Global Affairs Canada (2020) *Evaluation of Diplomacy, Trade, and Development Coherence in the Asia-Pacific Branch 2015-16 to 2020-21: Evaluation Design* PowerPoint

Global Affairs Canada (2021) *Evaluation of Diplomacy, Trade, and Development Coherence in the Latin American and Caribbean Region* PowerPoint

Gold, J. and Hutton, S. (2020) *Three Lessons from Past Public Health Crises for the Global Response to COVID-19 (coronavirus)* available on https://ieg.worldbankgroup.org/blog/3-lessons-past-public-health-crises-global-response-covid-19-coronavirus (accessed 11.02.21)

Government Offices of Sweden (2020) *Swedish Strategy to Combat Antimicrobial Resistance 2020-2023* available on https://www.government.se/499178/globalassets/government/dokument/socialdepartementet/amr_strategi_eng_web.pdf (accessed 26.02.21)

Grünewald, F., Binder, A. and Georges, Y. (2010) *Inter-agency Real-time Evaluation in Haiti: 3 Months after the Earthquake* available on https://www.humanitarianresponse.info/sites/www.humanitarianresponse.info/files/2019/08/IARTE---Haiti-%28English%29---2010.pdf (accessed 05.03.21)

Hallam, A. (1998) *Evaluating Humanitarian Assistance Programmes in Complex Emergencies* Good Practice Review for the Relief and Rehabilitation Network available on https://odihpn.org/resources/evaluating-humanitarian-assistance-programmes-in-complex-emergencies/ (accessed 15.02.21)

Hanley, T., Binas, R., Murray, J. and Tribunalo, B. (2014) *IASC Inter-Agency Humanitarian Evaluation of the Typhoon Haiyan Response* available on https://interagencystandingcommittee.org/system/files/evaluation_report_iahe_haiyan_december_2016.pdf (accessed 05.03.21)

Hodge-Mitchell, S., Amirebrahimi, H. and Rahmany, S. (2014) *MENARID: Institutional Strengthening and Coherence for Integrated Natural Resources Management in Iran* available on https://erc.undp.org/evaluation/evaluations/detail/5938 (accessed 04.03.21)

The Humanitarian Coalition (2012) *Haiti 2010 Earthquake Response: Final Evaluation Report* available on https://reliefweb.int/report/haiti/haiti-2010-earthquake-response-final-evaluation-report (accessed 05.03.21)

ICAI (2014) *Rapid Review of DFID's Humanitarian Response to Typhoon Haiyan in the Philippines* available on https://icai.independent.gov.uk/wp-content/uploads/ICAI-Philippines-report-FINAL.pdf (accessed 05.03.21)

ILO Evaluation Office (2020a) *ILO's Response to the Impact of the COVID-19 Pandemic on Workers and Enterprises: What Evaluative Lessons can be Drawn from the ILO's Past Responses to an Economic and Financial Crisis?* Available on http://www.covid19-evaluation-coalition.org/documents/ILO-Covid19.pdf (accessed 11.02.21)

ILO Evaluation Office (2020b) *Implications of COVID-19 on Evaluations in the ILO* available on http://uneval.org/document/download/3557 (accessed 19.02.21)

IMF (2021) *Policy Response to COVID-19* available on https://www.imf.org/en/Topics/imf-and-covid19/Policy-Responses-to-COVID-19 (accessed 26.02.21)

The Independent Panel for Pandemic Preparedness & Response (2021) *Core Documents* available on https://theindependentpanel.org/documents/ (accessed 19.02.21) includes:

- Terms of reference – available on https://theindependentpanel.org/wp-content/uploads/2020/10/TheIndependentPanel_TermsofReference.pdf
- World Health Assembly Resolution 73.1 - available on https://theindependentpanel.org/wp-content/uploads/2020/10/A73_R1-en.pdf
- Program of Work – available on https://theindependentpanel.org/wp-content/uploads/2020/10/The-Independent-Panel-Program-of-Work-October-20-2.pdf
- Reports of three meetings – available on:
  - 1st meeting - https://theindependentpanel.org/wp-content/uploads/2020/10/Final_TheIndependentPanel_MeetingReport_Sept17.pdf
  - 2nd meeting - https://theindependentpanel.org/wp-content/uploads/2020/10/TheIndependentPanel_2ndMeetingReport_Final.pdf
  - 3rd meeting - https://theindependentpanel.org/wp-content/uploads/2021/01/TheIndependentPanel_3rdMeetingReport.pdf
- Second report on progress - available on https://theindependentpanel.org/wp-content/uploads/2021/01/Independent-Panel_Second-Report-on-Progress_Final-15-Jan-2021.pdf
- Mapping of Key Commissions and Documents, Past and Present, on Pandemic Preparedness and Response and Broader Health Governance Issues – available on https://theindependentpanel.org/wp-content/uploads/2020/10/IndependentPanel_Mapping-Exercise.pdf
- State of the Pandemic – available on https://theindependentpanel.org/wp-content/uploads/2020/10/IndependentPanel_State-of-the-Pandemic_final.pdf

Innovations for Poverty Action (IPA) (2020a) *The Impact of Emergency Cash Assistance During the COVID-19 Pandemic in Colombia: Study Summary* available on https://www.poverty-action.org/printpdf/41851 (accessed 10.02.21)

Innovations for Poverty Action (IPA) (2020b) *The Impacts of Expanding Cash Transfers During an Emergency: Evidence from Colombia's Ingreso Solidario Program: Study Summary* available on https://www.poverty-action.org/printpdf/45101 (accessed 10.02.21)

Inter-Agency Standing Committee (2020) *Looking at the Coronavirus Crisis through the Nexus Lens – What Needs to be Done?* Available on https://interagencystandingcommittee.org/grand-bargain-official-website/looking-coronavirus-crisis-through-nexus-lens-what-needs-be-done (accessed 11.02.21)

International Federation of Red Cross and Red Crescent Societies (IFRC) (2021) *IFRC-Wide COVID-19 Evaluation: Terms of Reference* available on https://europeanevaluation.org/wp-content/uploads/2021/02/FW-Evaluation-COVID-19-Terms-of-Reference_27Jan202FINAL.pdf (accessed 10.02.21)

International Organization for Migration (2017) *The Caribbean: Hurricanes Irma, Maria and Jose Response: Situation Report* available on https://reliefweb.int/report/dominica/caribbean-hurricanes-irma-and-jose-response-situation-report-2-20-september-2017 (accessed 05.03.21)

IOM (2020) *Call for Applications for Consultancy Services: Real-time Evaluation of the IOM East and Horn of Africa COVID-19 Response* available on http://www.uneval.org/resources/images/vacancies/TOR-RealtimeEvaluation_Final.pdf (accessed 09.03.21)

IOM (undated) *Request for Proposal: The International Organization for Migration (IOM) is inviting interested Consultant to conduct an End term Evaluation of the Project on Addressing Trafficking and Smuggling through improved border management and Counter Trafficking Responses in Cox's Bazar in Bangladesh as per below Terms of Reference (ToR)* available on https://bangladesh.iom.int/sites/default/files/documents/ToR_Endline_Evaluation_INL_Project_2020_11_05.docx (accessed 09.03.21)

Ishida, Y. (2020) *How does the Newly Added DAC Evaluation Criterion "Coherence" Contribute to Achieving the SDG Target 4c for Teachers?* Journal of International Cooperation in Education Vol 22-2/23-2 (2020) pp15-29 available on https://cice.hiroshima-u.ac.jp/wp-content/uploads/2021/02/3.Ishida.pdf (accessed 25.06.21)

Itad (2015) *Evaluation of DFID's Humanitarian Response to Typhoon Haiyan (Yolanda): Final Report* available on https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/501225/Eval-Humanitarian-Response-Typhoon-Haiyan.pdf (accessed 05.03.21)

Jordana, J. and Triviño-Salazar, J.C. (2020) *Where are the ECDC and the EU-wide Responses in the COVID-19 Pandemic?* The Lancet, Volume 395, Issue 10237, p1611-1612, May 23 2020, available on https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)31132-6/fulltext (accessed 11.02.21)

K4D (2020) *K4D HelpDesk Reports and Emerging Issues Papers on COVID-19* available on http://www.covid19-evaluation-coalition.org/documents/k4d-helpdesk-%20reports-covid-19.pdf (accessed 26.02.21) including:
- *The Informal Sector and COVID-19 in Sub-Saharan Africa* available on https://opendocs.ids.ac.uk/opendocs/bitstream/handle/20.500.12413/15725/890_informal_sector_and_COVID-19_Final.pdf?sequence=5&isAllowed=y (accessed 26.02.21)
- *The COVID-19 Pandemic and the Future of Global Value Chains* available on https://opendocs.ids.ac.uk/opendocs/bitstream/handle/20.500.12413/15668/874_Covid-19_and_Global_Value_Chains.pdf?sequence=5&isAllowed=y (accessed 26.02.21)
- *COVID-19 and the Participation of Women and Women's Rights Organisations in Decision-Making* available on https://opendocs.ids.ac.uk/opendocs/bitstream/handle/20.500.12413/15532/850_COVID-19_and_the_participation_of_women_in_decision_making.pdf?sequence=1&isAllowed=y (accessed 26.02.21)

Keijzer, N. and Oppewal, J. (2012) *Learn to Walk Before You Run? Review of Methodological Approaches for Evaluating Coherence in the Field of International Cooperation* Discussion Paper No.132 of the European Centre for Development Policy Management (ECDPM) available on https://ecdpm.org/publications/review-methodological-approaches-evaluating-coherence-international-cooperation/ (accessed 09.02.21)

Key Aid Consulting (2018) *Final Evaluation: Nepal Earthquake Recovery Programme* available on https://www.alnap.org/help-library/british-red-cross-final-evaluation-nepal-earthquake-recovery-programme (accessed 05.03.21)

King, M., Keijzer N., Spierings E. and Matthews, A. (2012) *Measuring Policy Coherence for Development* A report for the European Centre for Development Policy Management (ECDPM) available on https://www.oecd.org/gov/pcsd/ECDPM%20Paper_Measuring%20PCD.pdf (accessed 10.02.21)

Kluve, J., Langbein, J. and Weber, M. (2020) *Protecting Workers and Firms in Times of Crisis: Key Labour Market Policies for Low-and Middle-Income Countries* available on http://www.covid19-evaluation-coalition.org/documents/Protecting-workers-firms-in-crisis.pdf (accessed 11.02.21)

Koff, H., Challenger, A. and Portillo, I. (2020) *Guidelines for Operationalizing Policy Coherence for Development (PCD) as a Methodology for the Design and Implementation of Sustainable Development Strategies* Sustainability 12, 4055 doi:10.3390/su12104055 available on https://www.mdpi.com/2071-1050/12/10/4055 (accessed 19.02.21)

Kuhnt, J. and Schüttler, K. (2020) *Economic Integration into Host Communities in Times of Crisis: How to Ease the Impact of COVID-19 on Displaced Populations in Low- and Middle-Income Countries* available on http://www.covid19-evaluation-coalition.org/documents/Economic-integration-into-host-communities-in-crisis.pdf (accessed 11.02.21)

Lamoure, G. and Juillard, H. (2020) *Responding to Ebola Epidemics: An ALNAP Lessons Paper* available on https://www.alnap.org/help-library/alnap-lessons-paper-responding-to-ebola-epidemics (accessed 18.05.21)

Lewin, S. and Glenton, C. (2020) *Effects of Digital Interventions for Promoting Vaccination Uptake*. Brief prepared for Norad. Oslo: Norwegian Institute of Public Health. October 2020. Available on http://www.covid19-evaluation-coalition.org/documents/VACCINES-Brief-2.pdf (accessed 11.02.21)

Lilly, D. (2020) *What Happened to the Nexus Approach in the COVID-19 Response?* Available on IPI Global Observatory https://theglobalobservatory.org/2020/06/what-happened-to-nexus-approach-in-covid-19-response/ (accessed 11.02.21)

Lockhart, C. (2005) *From Aid Effectiveness to Development Effectiveness: Strategy and Policy Coherence in Fragile States* ODI paper for the Senior Level Forum on Development Effectiveness in Fragile States available on https://inee.org/system/files/resources/doc_1_ODI-From_aid_effectiveness_to_dev_effectiveness.pdf (accessed 25.06.21)

Ministry of Foreign Affairs of Denmark (2006) *Evaluation Guidelines* available on http://www.netpublikationer.dk/um/7571/pdf/Guideline_web.pdf (accessed 23.02.21)

Ministry of Foreign Affairs of Finland (2021) *Evaluation Plan for 2021-2023* available on https://um.fi/documents/35732/0/Evaluation+Plan+2021-2023_English.pdf/d3d3f632-9369-040c-7e85-52f64b95ea82?t=1611584887747 (accessed 09.03.21)

Ministry of Foreign Affairs of the Netherlands (2019) *Mind the Governance Gap, Map the Chain: Evaluation of the Dutch Government's Policy on International Responsible Business Conduct (2012-2018)* available on https://english.iob-evaluatie.nl/binaries/iob-evaluatie-eng/documents/evaluations/2019/09/01/433-%E2%80%93-iob-%E2%80%93-evaluation-of-the-dutch-governments-policy-on-international-responsible-business-conduct-2012-2018-%E2%80%93-mind-the-governance-gap-map-the-chain/IOB_Evaluation_of_the_Dutch_government_s_policy_on_international_responsible_business_conduct_201909.pdf (accessed 26.02.21)

Momoh, H.B., Lamin, F. and Samai, I. (2016) *Final Report: Evaluation of DEC Ebola Response Program Phase 1 and 2* available on https://cafod.org.uk/content/download/33045/386590/version/4/file/CAFOD%20DEC%20Ebola%20Reponse%20Evaluation%20Report.pdf (accessed 05.03.21)

Murtaza, N., Arnborg, A. and Gautam, D. (2016) *Nepal Earthquake Response – NPL151: Final Evaluation Report* available on https://actalliance.org/wp-content/uploads/2018/03/Annexure-6-ACT_FINAL-EVALUATION-NPL151.pdf (accessed 05.03.21)

Nebehay, S. and Miller, J. (2020) *WHO Promises "Honest Evaluation" of how World Handled COVID-19* available on https://www.reuters.com/article/us-health-coronavirus-who-idUSKBN24A1O9 (accessed 11.02.21)

Núñez-Borja, C., Bandelet, E. and Picarello, T. (2018) *External Evaluation of the European Union's Policy Coherence for Development (2009-2016)* This report consists of two volumes, the main report and annexes. It was carried out on behalf of the European Commission by a consortium of Economisti Associati, BKP Development Research & Consulting, TRANSTEC and Consorzio Italiano Consulenti (C.I.C.) – available on https://ec.europa.eu/international-partnerships/external-evaluation-eus-policy-coherence-development-2009-2016_en (accessed 10.02.21)

Observatory of Public Sector Innovation (OPSI) (2020) *Linking the SDGs with COVID-19 Recovery: How can Policy Coherence by Applied to Ensure a Sustainable Recovery?* available on https://gov-after-shock.oecd-opsi.org/event/policy-coherence/ (accessed 22.02.21)

OECD (1999) *Guidance for Evaluating Humanitarian Assistance in Complex Emergencies* available on https://www.oecd.org/dac/evaluation/2667294.pdf (accessed 17.05.21)

OECD (2002) *Glossary of Key Terms in Evaluation and Results Based Management* available on https://www.oecd.org/development/evaluation/2754804.pdf (accessed 12.03.21)

OECD (2003) *Policy Coherence: Vital for Global Development* OECD Policy Brief available on http://www.oecd.org/governance/pcsd/20202515.pdf (accessed 19.02.21)

OECD (2008) *Policy Coherence for Development – Lessons Learned* available on https://www.un.org/en/ecosoc/newfunct/pdf/hls_finland-policy_coherence(oecd).pdf (accessed 19.02.21)

OECD (2017) *OECD Development Co-operation Peer Reviews: The Netherlands 2017* available on https://read.oecd-ilibrary.org/development/oecd-development-co-operation-peer-reviews-the-netherlands-2017_9789264278363-en#page28 (accessed 26.02.21)

OECD (2019) *Recommendation of the Council on Policy Coherence for Sustainable Development* OECD/LEGAL/0381 available on http://www.oecd.org/gov/pcsd/recommendation-on-policy-coherence-for-sustainable-development-eng.pdf (accessed 22.02.21)

OECD (2020a) *Better Criteria for Better Evaluation: Revised and Updated Evaluation Criteria* available on https://www.oecd.org/dac/evaluation/evaluation-criteria-flyer-2020.pdf (accessed 10.02.21)

OECD (2020b) *Evaluating the Initial Impact of COVID-19 Containment Measures on Economic Activity* available on https://www.oecd.org/coronavirus/policy-responses/evaluating-the-initial-impact-of-covid-19-containment-measures-on-economic-activity-b1f6b68b/ (accessed 11.02.21)

OECD (2020c) *Building a Coherent Response for a Sustainable Post COVID-19 Recovery* available on http://www.oecd.org/governance/pcsd/PRELIMINARY%20VERSION_PCSD_Policy-Response-Covid19_13%20July%202020.pdf (accessed 11.02.21)

OECD (2020d) *OECD Development Co-operation Peer Reviews: Ireland 2020: Policy Coherence for Sustainable Development* available on https://www.oecd-ilibrary.org/sites/c20f6995-en/1/3/2/1/index.html?itemId=/content/publication/c20f6995-en&_csp_=791813851d5b851636d9ca5e2f8f6dac&itemIGO=oecd&itemContentType=book#section-d1e1005 (accessed 26.02.21)

OECD (2020e) *OECD Development Co-operation Peer Reviews: Austria 2020: Executive Summary* available on https://www.oecd-ilibrary.org/sites/03b626d5-en/index.html?itemId=/content/publication/03b626d5-en (accessed 04.03.21)

OECD (2020f) *Development Co-operation Report 2020: Learning from Crises, Building Resilience* available on https://doi.org/10.1787f6d42aa5-en (accessed 23.03.21)

OECD (undated) *Policy Coherence for Sustainable Development* Website available on http://www.oecd.org/gov/pcsd/ (accessed 22.02.21) includes *Policy Coherence for Sustainable Development Toolkit* available on http://www.oecd.org/governance/pcsd/toolkit/ (accessed 22.02.21)

OECD DAC (1991) *Principles for Evaluation of Development Assistance* available on https://www.oecd.org/dac/evaluation/2755284.pdf (accessed 12.03.21)

OECD DAC (1999) *Guidance for Evaluating Humanitarian Assistance in Complex Emergencies* available on https://www.oecd.org/dac/evaluation/2667294.pdf (accessed 15.02.21)

OECD DAC Governance Network (2020) *Situation Report on Governance and COVID-19*

OECD DAC Network on Development Evaluation (2019) *Better Criteria for Better Evaluation: Revised Evaluation Criteria Definitions and Principles for Use* available on http://www.oecd.org/dac/evaluation/revised-evaluation-criteria-dec-2019.pdf (accessed 10.02.21)

OECD DAC Network on Development Evaluation and the Independent Evaluation Office of UNDP (2020) *Good Practices during COVID-19: IEO/UNDP and OECD/DAC Evalnet Joint Guidance Note for*

Evaluation Units available on
http://web.undp.org/evaluation/guideline/documents/covid19/IEOOECD_DAC_Joint-Guidance_COVID19.pdf (accessed 11.02.21)

OECD and ECDPM (2013) *Towards a Methodology for Country-level Impact Assessments of PCD on Food Security* Slides available on https://www.slideshare.net/ecdpm/ecdpm-ppt-pcd-methodology-oecd-mtng-230513-final (accessed 23.03.21)

Office of Evaluation and Oversight, Inter-American Development Bank (IADB) (2020) *COVID-19: What can we Learn from Past Responses to Public Health Crises?* Available on https://www.iadb.org/en/ove/covid-19 (accessed 10.02.21)

Office of Internal Oversight Services Inspection and Evaluation Division (2019) *Evaluation of United Nations Entities' Preparedness, Policy Coherence and Early Results Associated with their Support to Sustainable Development Goals* available on http://www.uneval.org/evaluation/reports/detail/11340 (accessed 04.03.21)

Office of Internal Oversight Services Inspection and Evaluation Division (2020a) *Synthesis of Guidelines for UN Evaluation under COVID-19* available on file:///C:/Users/Roger/Downloads/IED%20-%20OIOS%20Synthesis%20of%20Guidelines%20for%20UN%20Evaluation%20under%20COVID.pdf (accessed 19.02.21)

Office of Internal Oversight Services Inspection and Evaluation Division (2020b) *COVID-19 Response Evaluation Protocol* available on http://www.uneval.org/document/download/3687 (accessed 04.03.21)

Patrick, J. (2011) *Haiti Earthquake Response: Emerging Evaluation Lessons* for OECD Evaluation Insights Number 1, June 2011 available on http://www.oecd.org/dac/evaluation/dcdndep/48165069.pdf (accessed 22.02.21)

Patton, M.Q. (2020) *Evaluation Criteria for Evaluating Transformation Implications for the Coronavirus Pandemic and the Global Climate Emergency* American Journal of Evaluation 1-37 available on https://doi.org/10.1177/1098214020933689 (accessed 10.02.21)

Picciotto, R. (2005) *The Evaluation of Policy Coherence for Development* Evaluation, Vol 11(3):311-330, available on https://journals.sagepub.com/doi/10.1177/1356389005058479 (accessed 10.02.21)

Raab, M. (2020) *Know what You Need to Know* available on https://www.developblog.org/2020/09/know-what-you-need-to-know.html (accessed 26.06.21)

Ramalingam, B., Singh, N.S., Mahieu, A. and Blanchet, K. (2020) *Responding to COVID-19: Guidance for humanitarian agencies*. ALNAP Rapid Learning Review. London: ODI/ALNAP. Available on http://www.covid19-evaluation-coalition.org/documents/alnap-responding-to-covid-19-review_0.pdf (accessed 11.02.21)

Robinson, L., Cichocka, B., Richie, E. and Mitchell, I. (2020) *The Commitment to Development Index: 2020 Edition: Methodological Overview Paper* available on https://www.cgdev.org/sites/default/files/CDI-Methodology-2020.pdf (accessed 22.03.21)

Scottish Government (2020) *Coronavirus (COVID-19): Review of International Development Programme* available on https://www.gov.scot/policies/international-development/2020-covid-19-review/ (accessed 11.02.21)

Scottish Government (2021) *Summary Report on the Review of Scottish Government's International Development Programme in Light of COVID-19* available on https://www.gov.scot/binaries/content/documents/govscot/publications/progress-report/2021/03/summary-report-review-scottish-governments-international-development-programme-light-covid-192/documents/summary-report-review-scottish-governments-international-development-programme-light-covid-19/summary-report-review-scottish-governments-international-development-programme-light-covid-19/govscot%3Adocument/summary-report-review-scottish-governments-international-development-programme-light-covid-19.pdf?forceDownload=true (accessed 26.06.21)

Select Committee on International Development (undated) *Policy Coherence for Development* available on https://publications.parliament.uk/pa/cm200405/cmselect/cmintdev/123/12307.htm (accessed 23.02.21)

Sida, L., Trombetta, L. and Panero, V. (2016) *Evaluation of OCHA Response to the Syria Crisis* available on https://www.unocha.org/sites/dms/Documents/OCHA%20Syria%20Evaluation%20Report_FINAL.pdf (accessed 05.03.21)

Stave, S.E., Nilsen, M. and Dalen, K. (2018) *Evaluation of Norwegian Efforts to Ensure Policy Coherence for Development* Norad Report 8/18 available on https://www.norad.no/en/toolspublications/publications/2018/evaluation-of-norwegian-efforts-to-ensure-policy-coherence-for-development/ (accessed 15.02.21). There is also a policy brief and additional annexes.

Thormar, S. (2013) *Evaluation of Ebola Response: Uganda* available on https://adore.ifrc.org/Download.aspx?FileId=42478&.pdf (accessed 05.03.21)

Telford, J. Cosgrave, J. and Houghton, R. (2006) *Joint Evaluation of the International Response to the Indian Ocean Tsunami: Synthesis Report* available on http://www.tsunami-evaluation.org/NR/rdonlyres/2E8A3262-0320-4656-BC81-EE0B46B54CAA/0/SynthRep.pdf (accessed 22.02.21)

UNDP Independent Evaluation Office (2020a) *Lessons from Evaluations: UNDP Governance Support to Countries in Crisis* available on http://web.undp.org/evaluation/documents/reflections/governance.pdf (accessed 11.02.21)

UNDP Independent Evaluation Office (2020b) *Lessons from Evaluations: UNDP Support to Social Protection* available on http://web.undp.org/evaluation/documents/reflections/social_protection.pdf (accessed 11.02.21)

UNDP Independent Evaluation Office (2020c) *Lessons from Evaluations: UNDP Support to Livelihoods Restoration and Job Creation in Crisis Countries available* on http://web.undp.org/evaluation/documents/reflections/livelihoods.pdf (accessed 11.02.21)

UNDP Independent Evaluation Office (2020d) *Lessons from Evaluations: UNDP Support to the Health Sector available* on http://web.undp.org/evaluation/documents/reflections/health-sector.pdf (accessed 11.02.21)

UNDP Independent Evaluation Office (2021) *Reflections: Lessons from Evaluations: Learning from Past Crises for Recovering from COVID-19* available on http://web.undp.org/evaluation/documents/reflections/book/reflections-crisis-series-02-21.pdf (accessed 19.02.21)

UNDP Independent Evaluation Office (undated) *Evaluation during COVID-19* available on http://web.undp.org/evaluation/guideline/covid19.shtml (accessed 26.2.21)

UNFPA Evaluation Office (2019) *Evaluation of the UNFPA Response to the Syria Crisis (2011-2018)* available on https://www.unfpa.org/admin-resource/evaluation-unfpa-response-syria-crisis-2011-2018 (accessed 05.03.21)

UNFPA Evaluation Office (2020a) *Adapting Evaluations to the COVID-19 Pandemic* available on http://uneval.org/document/download/3559 (accessed 19.02.21)

UNFPA Evaluation Office (2020b) *Formative Evaluation of UNFPA Approach to South-South and Triangular Cooperation* available on https://www.unfpa.org/sites/default/files/admin-resource/1._SSC_Eval_Report_Final.pdf (accessed 09.03.21)

UNHCR (2021) *UNHCR's Approach to Evaluating the COVID-19 Response for Refugees and Other Persons of Concern* available on https://reliefweb.int/sites/reliefweb.int/files/resources/601293564.pdf (accessed 11.02.21)

UNHCR, Ministry for Foreign Affairs of Finland and the COVID-19 Global Evaluation Coalition (2020) *Joint Evaluation of the Protection of the Fundamental Rights of Refugees during the COVID-19 Pandemic: Concept Note* available on http://www.covid19-evaluation-coalition.org/documents/Concept%20Note%20-%20Refugee%20Rights%20and%20COVID%20SEP2020_with%20Coalition%20logo.pdf (accessed 11.02.21)

UNICEF (2017) *Evaluation of UNICEF's Response to the Ebola Outbreak in West Africa 2014-2015* available on https://reliefweb.int/report/liberia/evaluation-unicef-s-response-ebola-outbreak-west-africa-2014-2015-enesfr (accessed 05.03.21)

UNICEF (2018) *Evaluation of the UNICEF Level 3 Response to the Cholera Epidemic in Yemen: Crisis Within a Crisis* available on https://digitallibrary.un.org/record/1637498?ln=en (accessed 005.03.21)

UNICEF (2020a) *Response of the UNICEF Evaluation Function to the COVID-19 Crisis: Technical Note* available on http://uneval.org/document/download/3560 (accessed 19.02.21)

UNICEF (2020b) *Review of Risk Communication and Community Engagement Initiative for COVID-19 Prevention Behaviours in Cambodia* available on https://evaluationreports.unicef.org/GetDocument?fileID=16456 (accessed 04.03.21)

UNICEF (2020c) *After Action Review: Thailand Country Office Response to COVID-19 Crisis: Final Report* available on https://evaluationreports.unicef.org/GetDocument?fileID=16212 (accessed 04.03.21)

UNICEF (2020d) *EAPRO COVID-19 Real Time Assessment: Initial Report for UNICEF Malaysia Country Office* available on https://evaluationreports.unicef.org/GetDocument?fileID=16364 (accessed 04.03.21)

UNICEF (2021a) *UNICEF Mongolia: Findings from the Regional COVID-19 Response Real Time Assessment – Country Office Report* available on https://evaluationreports.unicef.org/GetDocument?fileID=16342 (accessed 04.03.21)

UNICEF (2021b) *Real-Time Assessment of the UNICEF South Asia Response to COVID-19* available on https://evaluationreports.unicef.org/GetDocument?fileID=16489 (accessed 04.03.21)

UNICEF and UNFPA (2020) *Joint Evaluation of the UNFPA-UNICEF Joint Programme on the Abandonment of Female Genital Mutilation: Accelerating Change* available on https://www.unfpa.org/admin-resource/joint-evaluation-unfpa-unicef-joint-programme-abandonment-female-genital-mutilation (accessed 04.03.21)

United Nations (2020a) *The Impact of COVID-19 on Food Security and Nutrition* available on https://www.un.org/sites/un2.un.org/files/sg_policy_brief_on_covid_impact_on_food_security.pdf (accessed 11.02.21)

United Nations (2020b) *United Nations Comprehensive Response to COVID-19: Saving Lives, Protecting Societies, Recovering Better* available on https://www.un.org/sites/un2.un.org/files/un_comprehensive_response_to_covid-19_june_2020.pdf (accessed 26.02.21)

United Nations, The World Bank, European Union and African Development Bank (2015) *Recovering from the Ebola Crisis* available on https://www.undp.org/content/undp/en/home/librarypage/crisis-prevention-and-recovery/recovering-from-the-ebola-crisis---full-report.html (accessed 19.03.21)

UNODC (undated) *COVID-19: UNODC's Approach to Evaluation During Crisis* available on http://uneval.org/document/download/3568 (1ccessed 19.02.21)

UN Women (2020) *Pocket Tool for Managing Evaluation during the COVID-19 Pandemic* available on https://genderevaluation.unwomen.org/-/media/files/un%20women/gender%20evaluation/resourcefiles/pocket%20tool.pdf?la=en&vs=4604%20http://web.undp.org/evaluation/media-centre/infographics/evaluation_covid19.shtml%20https://www.iom.int/sites/default/files/about-iom/evaluation/me-continuity-in-covid-08-april-20-version1.pdf (accessed 19.02.21)

WFP (2017) *An Evaluation of WFP's Response to the Ebola Virus Disease (EVD) Crisis in West Africa* available on https://docs.wfp.org/api/documents/789c0eb95e5d4773884d920e9f605673/download/?_ga=2.112391172.1615512947.1614339293-379844604.1613051255 (accessed 26.02.21)

WFP (2018a) *Evaluation of the WFP Regional Response to the Syrian Crisis (2015-2018)* available on https://www.wfp.org/publications/evaluation-wfps-regional-response-syrian-crisis-2015-2017 (accessed 26.02.21)

WFP (2018b) *Decentralized Evaluation: WFP's General Food Assistance to Syrian Refugees in Jordan 2015 to mid-2018: Evaluation Report* available on https://docs.wfp.org/api/documents/WFP-

0000101797/download/?_ga=2.11613236.1615512947.1614339293-379844604.1613051255 (accessed 26.02.21)

WFP (2019) *Evaluation of WFP's Corporate Emergency Response in Northeast Nigeria (2016-2018)* available on https://www.wfp.org/publications/wfps-corporate-emergency-response-northeast-nigeria-evaluation (accessed 26.02.21)

WFP (2020a) *Technical Note for Planning and Conducting Evaluations during COVID-19* available on http://www.covid19-evaluation-coalition.org/documents/Technical-Note-evaluations-during-Covid-19.pdf (accessed 11.02.21)

WFP (2020b) *Evaluation of WFP's Response to the COVID-19 Pandemic (2019-2020): Terms of Reference* available on https://www.wfp.org/publications/evaluation-wfps-response-covid-19-pandemic and https://docs.wfp.org/api/documents/WFP-0000122257/download/?_ga=2.126070733.1583616632.1613051255-379844604.1613051255 (accessed 11.02.21)

WFP (2020c) *Strategic Evaluation of WFP's Capacity to Respond to Emergencies* available on https://docs.wfp.org/api/documents/WFP-0000112253/download/?_ga=2.120683656.1615512947.1614339293-379844604.1613051255 (accessed 26.02.21)

WFP (2020d) *Evaluation of Democratic Republic of Congo WFP Interim Country Strategic Plan 2018-2020* available on https://www.wfp.org/publications/democratic-republic-congo-interim-country-strategic-plan-evaluation-2018-2020 (accessed 26.02.21)

White, G. (2019) *IFRC Hurricane Irma Response Operation (Antigua & Barbuda and St Kitts & Nevis): Final Evaluation* available on https://reliefweb.int/report/antigua-and-barbuda/final-evaluation-ifrc-hurricane-irma-response-operation-antigua-barbuda (accessed 05.03.21)

WHO (2020) *COVID-19 Strategic Preparedness and Response (SPRP) Monitoring and Evaluation Framework* available on https://www.who.int/docs/default-source/coronaviruse/who-ncov-me-framework-web.pdf?sfvrsn=656e430f_1&download=true (accessed 19.02.21)

Wood, B., Kabell, D., Sagasti, F. and Muwanga, N. (2008) *Synthesis Report on the First Phase of the Evaluation of the Implementation of the Paris Declaration* available on http://www.oecd.org/development/evaluation/dcdndep/40888983.pdf (accessed 23.03.21)

World Bank (2015) *When Institutions Work: Nigeria's Ebola Response* available on https://openknowledge.worldbank.org/bitstream/handle/10986/22775/Doing0developm00delivery0case0study.pdf?sequence=1&isAllowed=y (accessed 19.03.21)

World Bank (2020) *Coherent Policy Response Needed to Overcome Coronavirus Crisis in Latin America and the Caribbean* available on https://www.worldbank.org/en/news/press-release/2020/04/12/coronavirus-crisis-latin-america-and-the-caribbean (accessed 26.02.21)

World Bank Independent Evaluation Group (2020) *Lessons from Evaluation: Support and Financing to the Formal Private Sector in Response to COVID-19* available on https://ieg.worldbankgroup.org/sites/default/files/Data/Topic/COVID19Lessons_CrisisResponse_FormalPrivateSector.pdf (accessed 11.02.21). This is within the IEG Lesson Library available on https://ieg.worldbankgroup.org/topic/covid-19-coronavirus-response (accessed 26.02.21).

Other materials in here include:
- Knowledge Series on Using Technologies and Tools for Remote Data Collection
- Thematic Evaluations and Syntheses including:
  - *IDA's Crisis Response Window: Lessons from IEG Evaluations* available on https://ieg.worldbankgroup.org/sites/default/files/Data/reports/syn_idacrisisresponse.pdf (accessed 26.02.21)
  - *Crisis Response and Resilience to Systemic Shocks: Lessons from IEG Evaluations* available on https://ieg.worldbankgroup.org/sites/default/files/Data/reports/building-resilience.pdf (accessed 26.02.21)
  - *Responding to Global Public Bads: Learning from Evaluation of the World Bank Experience with Avian Influenza 2006-13* available on https://ieg.worldbankgroup.org/sites/default/files/Data/Evaluation/files/avian_flu1.pdf (accessed 26.02.21)
  - *The World Bank Group's Response to the Global Economic Crisis: Phase 1 & 2* available on https://ieg.worldbankgroup.org/sites/default/files/Data/Evaluation/files/Global_Econ_Crisis-full.pdf and https://ieg.worldbankgroup.org/sites/default/files/Data/Evaluation/files/crisis2_full_report.pdf (accessed 26.02.21)
  - *The World Bank Group and the Global Food Crisis* available on https://ieg.worldbankgroup.org/sites/default/files/Data/reports/food_crisis_eval_1.pdf (accessed 26.02.21)
- Reference Guides including:
  - *Lessons from Fragility Conflict and Violence (FCV) for COVID-19* available on https://ieg.worldbankgroup.org/sites/default/files/Data/Topic/IEG_COVID19_LessonsFromFCV.pdf (accessed 26.02.21)
  - *Lessons from Mobilizing Finance for Development (MFD) for COVID-19: A Reference Guide* available on https://ieg.worldbankgroup.org/sites/default/files/Data/Topic/ReferenceGuide_COVID19_MFD.pdf (accessed 26.02.21)
  - *Human Development and COVID-19: A Reference Guide* available on https://ieg.worldbankgroup.org/sites/default/files/Data/Topic/ReferenceGuide_COVID19_HD.pdf (accessed 26.02.21)
- Blogs and Blog Series – a number of blogs including:

  - *What do Past Crises Tell Us about Coping with the Economic Shocks of COVID-19 (Coronavirus)* available on https://ieg.worldbankgroup.org/blog/what-do-past-crises-tell-us-about-coping-economic-shocks-covid-19-coronavirus (accessed 26.02.21)
  - *The World Bank and Global Collaboration: Lessons for the COVID-19 (coronavirus) response* available on https://ieg.worldbankgroup.org/blog/world-bank-and-global-collaboration-lessons-covid-19-coronavirus-response (accessed 26.02.21)

Wright, S. and Roche, J.M. (2018) *Assessing Aid and Policy Coherence* available on https://www.savethechildren.org.uk/blogs/2018/assessing-aid-and-policy-coherence (accessed 26.02.21)

Zetter, R., Mackie, J., Ruaudel, H., Seppänen, M. and Cascone, N. (2019) *Evaluation on Forced Displacement and Finnish Development Policy: Final Report* available on http://www.oecd.org/derec/finland/finland-2019-forced-displacement.pdf (accessed 26.02.21)