

Back to the Drawing Board

How to improve monitoring of outcomes

Neil Dillon and Amelie Sundberg



**MONITORING &
EVALUATION**

 **ALNAP**

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

www.alnap.org

About the authors

Neil Dillon is a Research Fellow, and Amelie Sundberg is a Senior Research Officer at ALNAP.

Acknowledgements

A special thank you goes to the individuals who gave their time to discuss design and review drafts of this research. Insights and comments from Anne Langston (IRC) and OFDA were invaluable. Thank you as well to the many individuals who generously gave their time to share the experience, insights and ideas synthesised and recorded in this report.

Within the ALNAP Secretariat, thanks are owed to the stellar Communications team, Maria Gili, Tim Harcourt-Powell and Cara Casey-Boyce, for their ability to distil and portray our ideas for the paper and animation. Catriona Foley and Charlotte Skinner provided instrumental research assistance in organisation logistics for research travel and workshops, without which this report would not have been possible. Finally, thank you also to John Mitchell and Alice Obrecht for helping to orient the research through discussion.

Suggested citation

Dillon, N. and Sundberg, A. (2019) *Back to the Drawing Board: How to improve monitoring of outcomes*. ALNAP Paper. London: ODI/ALNAP.

ISBN: 978-1-910454-92-3

© ALNAP/ODI 2019. This work is licensed under a Creative Commons Attribution-non Commercial Licence (CC BY-NC 4.0).

Design by Soapbox, www.soapbox.co.uk
Layout by Cara Casey-Boyce
Copyediting by Aaron Griffiths
Bibliographic editing by Renée Goulet

Cover image

Photo credit: Isabel Coello/ECHO.

Contents

Introduction	3
Process and method	6
<hr/>	
1 The outcomes monitoring landscape	9
<hr/>	
2 Cracks in the landscape	13
2.1 Understanding what things need to be measured	13
Defining the concept of 'outcomes'	13
Formulating specific outcomes	14
Handling sectoral silos	17
2.2 Learning how to measure them	18
Aggregating programme outcomes from project outcomes	18
Aggregation	18
Funding cycles	22
The need for more open-ended enquiry	23
Data management	25
2.3 Doing what it takes to get this done	26
Resourcing	26
Cultural shifts	28
<hr/>	
3 Conclusion	30
Bibliography	31

Abbreviations and acronyms

ALNAP	Active Learning Network for Accountability and Performance
CRS	Catholic Relief Services
CAFOD	Catholic Agency for Overseas Development
CCCM	Camp Coordination and Camp Management
DAC	Development Assistance Committee
Danida	Danish International Development Agency
DFID	UK Department for International Development
DRC	Danish Refugee Council
ECHO	European Civil Protection and Humanitarian Aid Operations
FAO	United Nations Food and Agriculture Organisation
FFO	Federal Foreign Office
ICLA	Information, counselling and legal assistance
IRC	International Rescue Committee
M&E	Monitoring and evaluation
NGO	Non-governmental organisation
NRC	Norwegian Refugee Council
OECD	Organisation for Economic Co-operation and Development
OFDA	US Foreign Disaster Assistance
SC	Save the Children
ToC	Theory of Change
UNICEF	United Nations Children's Fund
WCH	War Child Holland
WCUK	War Child UK
WFP	World Food Programme

Key to design features

Icons



Definition

Introduction

Measuring the results of humanitarian action is extremely important. It is essential for any assessment of whether a humanitarian agency's projects and programmes are achieving what they set out to achieve, and whether they are having unintended consequences on the lives of people affected by crisis (Obrecht, 2018a; Warner, 2017; Hofmann et al., 2004).

Most humanitarian organisations know this. A number invest in systems and tools to help monitor outcomes, defined as the 'likely or achieved short-term and medium-term effects of an intervention's outputs' (OECD-DAC, 2002). Catholic Relief Services (CRS), United Nations Food and Agriculture Organisation (FAO), International Rescue Committee (IRC), Norwegian Refugee Council (NRC), Oxfam, Save the Children (SC), UNICEF (United Nations Children's Fund), the World Food Programme (WFP) and War Child UK (WUK) have all published monitoring frameworks or guidance within the last decade that delineate their approach to outcome monitoring. And interest is growing (Warner, 2017).

This is in part a response to the changing nature of humanitarian action. Organisations are increasingly being asked to respond to protracted crises (OCHA, 2018; Knox-Clarke, 2018; Bennett, 2015). Many non-emergency-based interventions, such as livelihoods and resilience programmes, seek to achieve quite long-lasting behavioural change (FAO, 2018; Peters et al., 2016; IFRC, 2011). This type of work can straddle the divide between humanitarian and development work and create a demand to understand medium and longer effects of an intervention (Agenda for Humanity, 2016; Bennett et al., 2016). Likewise, the increasing use of multi-sectoral cash-based interventions by many actors has put the spotlight on the capacity of humanitarian actors to meet the broad package of basic household needs (CaLP, 2018).

But it is also driven by a need felt by organisations to demonstrate the effectiveness of their work. The drive towards evidence-based programming and accountability to affected populations has increased the emphasis on outcomes monitoring (Darcy et al., 2013; UNICEF, 2017; CHS, 2014). And most organisations now want to be able to make claims about the effects of their outputs, how these effects compare between similar interventions in different places, and how they change and develop over time.

So, the need to monitor outcomes is understood, and investment in outcomes monitoring is growing. But there is a risk that monitoring systems, as they currently stand, are being asked to do too much. Initial scoping work conducted by the ALNAP Secretariat suggests that monitoring systems are often pulled in different directions by different stakeholders. What started, for most organisations, as a system to measure the outputs of individual projects, is now being asked to do much more. Donors want to use monitoring data to compare projects between different NGOs in their funding portfolio; NGOs want to use the same data to compare projects across different donors; and both are seeking to do this at both the country and global levels.

“What started, for most organisations, as a system to measure the outputs of individual projects, is now being asked to do much more.”

The result is often an overstretched monitoring system, trying to achieve different goals with the same tool, and producing a kaleidoscope of information of different and incompatible types. The data needs for country and global-level comparisons, for example, are often not the same – and even sometimes incompatible (GPPi, 2016). Good quality country-level analysis requires highly granular data tailored to the specific context of operation. In contrast, global-level analysis requires generalisable data and a significant degree of standardisation of data parameters between contexts. Likewise, reporting deadlines and project cycle lengths vary considerably between donors, NGOs and country contexts (Gaston, 2017). This makes it very hard to measure outcomes across a project portfolio using a monitoring system that is itself tied to an individual project funding cycle (Ramalingam et al., 2019; Mayne, 2007; Hatton and Schroeder, 2007; Hofmann et al., 2004).

Invariably, the kaleidoscope does not provide the big picture. How can organisations provide holistic, meaningful analysis of the outcomes achieved by their activities? How can they situate their programming among the changes taking place around them? How can they understand the full set of changes experienced by an individual affected by crisis? It makes little sense to monitor outcomes of a livelihoods programme without taking account of changes in a host government’s policy on refugees’ rights to work, for example. Likewise, it is of little use to monitor sanitation outcomes for an individual without tracking how their shelter arrangements have changed during the implementation period. The risk of continuing with the status quo is that outcome monitoring systems fail to capture this broader picture, and ultimately fail to provide anyone with the information they need.

“The risk of continuing with the status quo is that outcome monitoring systems fail to capture this broader picture, and ultimately fail to provide anyone with the information they need.”

This paper aims to encourage humanitarian agencies to step back and reflect on what is currently being done to measure outcomes and how it can be improved in the future. It starts by identifying core assumptions and foundational thinking behind current monitoring systems. It then outlines issues arising from current practice, and concludes by raising questions about to think about this differently.



Photo credit: Samiel Marie-Fanon/ECHO.

Process and method

This paper is one part of a series of research products developed by the ALNAP Secretariat on the subject monitoring of humanitarian action. The series began with a scoping paper describing current practice and identifying challenges (Warner, 2017). This work helped identify a range of issues for improvement in the monitoring systems observed. In 2017–2018, the ALNAP Secretariat consulted the membership to select the critical challenges for further research. Four issues were identified:

- Limited ability to measure outcomes in a meaningful way.
- Capacity constraints regarding the capture and use of qualitative data by monitoring teams.
- An absence of tools for sharing good monitoring practice within and across organisations.
- Limited use of monitoring and evaluation (M&E) information to support project decision-making and learning.

Each of these issues was investigated further through independent research components. The results of each research area are available on the ALNAP website at alnap.org/me.

This paper tackles the first of these four challenges. It is based on a literature review, key informant interviews and case studies. The literature review covered 28 organisational M&E guidance documents, policies, toolkits, and frameworks; 36 published grey and academic literature focusing on humanitarian outcomes or related topics; 7 internal ‘project packages’ from participating organisations covering theories of change, project results frameworks, project proposal narratives, monitoring reports, evaluation reports and mid-term review reports.

A total of 42 key informant interviews were conducted, including:

- 12 interviews with HQ M&E and relevant technical staff
- 10 regional or country-level M&E staff
- 5 donor agencies: FFO Germany, the Office of US Foreign Disaster Assistance (OFDA), the European Commission, the Danish International Development Agency (Danida) and the Ikea Foundation
- 15 sectoral humanitarian experts or academic/technical experts, including the Water, Sanitation and Hygiene (WASH), Nutrition, Camp Coordination and Camp Management (CCCM) and Shelter Clusters.

In addition, four case studies were conducted to provide context-rich examples of the issues identified by HQ interviews and documentation. The four case studies were the Danish Refugee Council (DRC – global results framework), War Child Holland (WCH) and WCUK (global results framework), Catholic Agency for Overseas Development (CAFOD) in Zimbabwe and NRC in Lebanon.

Data collection was structured around two core research questions, identified during the scoping phase:

1. How close are we to monitoring outcomes in humanitarian action, and what are the key issues that need to be resolved to realise the benefits of outcome monitoring?
2. What has worked well so far and what feasible recommendations can be made for further progress towards operational use of outcome monitoring in humanitarian action?



The outcomes monitoring landscape

1 The outcomes monitoring landscape

In the past few years, several organisations have invested significant time and resources at a central level in improving outcomes monitoring. They have approached this in one or several of the following ways:

- articulating global-level outcomes
- setting policies on monitoring and reporting practice for country offices to follow
- establishing strategic or operational frameworks or guidance specific to outcomes for staff to use
- drafting toolkits for outcomes design and measurement, and/or developing data management systems.

Invariably, the choice of which approach to follow – and how – reflects the underlying structure and business model of the organisation, as much as the core monitoring objectives. But, regardless of which route organisations take, this study finds a common set of four core assumptions underlying the way in which outcomes monitoring is conceptualised.

First, all organisations have worked on the basis that they should be measuring planned-for outcomes of specific projects (Hofmann et al., 2004). To a degree, this is a symptom of the prioritisation donor accountability over learning (Mayne, 2007; Dillon, 2019; Ramalingam et al., 2019). The Results-Based Management (RBM) approach typically requires projects or programmes to be designed against pre-defined objectives and targets, against which success is then measured (Hatton and Schroeder, 2007; OECD-DAC, 2002). The widespread use of RBM means that most organisations design their programmes with specific outcomes in mind and then set up their monitoring activities to assess whether these outcomes have been achieved.

Second, organisations widely use a ‘Theory of Change’ (ToC) framework to help anticipate these effects and how they might occur as a result of programming. The ToC methodology is supposed to define long-term goals (such as impact or outcomes) and then map backwards the chain of preconditions necessary for achieving these goals (Rogers, 2014). Often, donors ask implementing partners to present an intervention logic for their proposed projects using ToC principles (DFID, Danida, European Civil Protection and Humanitarian Aid Operations (ECHO), United States Agency for International Development, Global Affairs Canada, Swedish International Development Cooperation Agency). The different stages in these ToCs are often mirrored in – or even expressed as – the logframe indicators that partners need to report against under the RBM approach. Over the past ten years organisations have even applied ToC thinking to develop global strategic objectives and country strategic objectives (IRC, 2015; NRC, 2018; SC, 2016). In some organisations this has led to the application of centralised and standardised ToCs across country offices and contexts.

“Donors and organisations have explored various takes on the same principle of pulling together a list of vetted or acceptable outcome indicators.”

Third, most efforts to improve the monitoring of outcomes have started with an indicator-based approach. Donors and organisations have explored various takes on the same principle of pulling together a list of vetted or acceptable outcome indicators from which country offices can choose from when they are designing their projects (IRC, NRC, SC, ECHO; OFDA; Indikit¹). At the same time, significant resources have also been put into the development of data management platforms based on indicator reporting. This is, in part, an effort to help country teams design and measure outcomes. But it is also an attempt to lay the groundwork for more harmonised measurements, which could eventually be looked at over time or compared across projects and programmes. For example, NRC has a few selected mandatory outcomes for all of its sectors, linked to the programmatic ToC, which must be reported against for a country office engaging in that area of work. They also have a list of suggested indicators that can be selected as relevant. In turn, Humanity and Inclusion (HI) has a list of recommended outcome indicators and SC is working on developing a similar list that will become a part of their Global Results Framework. In comparison, the United Nation’s WFP has articulated a few mandatory outcome indicators for specific programming that correspond to their overarching Evidence Frameworks, and IRC is exploring how to do the same (WFP, 2018; IRC). WCH and WCUK have also together developed a

1. See <https://www.indikit.net/>.

specific list of project indicators that correspond to their Global Monitoring Framework. Each project seeks to include at least one relevant outcome indicator per thematic area.

And fourth, outcome measurements are primarily project-focused. As humanitarian intervention is largely modelled on project-level funding for individual organisations for a distinct time period, there is currently little incentive to measure collective or response-level outcomes. A small number of recent consortium interventions do attempt to jointly measure a handful of outcome indicators, but these are still collectively measuring and reporting against the same project. In addition, where some coordination bodies, such as the Global Shelter Cluster, are exploring ways in which to measure outcomes at the cluster-level across multiple partners, these are often built upon the idea of adding together the results from different projects. As there are differences between how individual partners in the cluster have designed their outcomes and gone about collecting them, coordination bodies are currently struggling to aggregate data beyond the project level.



Photo credit: ECHO.

Cracks in the landscape

2 Cracks in the landscape

While it is encouraging to see a degree of political will and commitment at the strategic level to finding solutions to monitoring of outcomes, organisations that have been working on these systems have found that certain issues arise. These occur across three dimensions:

- Understanding what things need to be measured: looking at definition, formulation of outcomes, and sectoral silos
- Learning how to measure them: looking at possibilities for aggregation, open-ended inquiry and data management systems
- Doing what it takes to get this done: including the levels of investment and parallel internal advocacy required.

Each of these issues need to be resolved in order to monitor outcomes in a way that yields relevant and useful information for decision-making in humanitarian action.

2.1 Understanding what things need to be measured

Defining the concept of ‘outcomes’

‘Outcomes’ can mean different things to different people. Most humanitarian organisations use the OECD-DAC definition of ‘likely or achieved short-term and medium-term effects of an interventions’ outputs’ (OECD-DAC 2002; NRC, 2016; UNICEF, 2011). While it can be difficult to draw a hard distinction between short and medium-term outcomes, there are certainly differences between them. On the one hand, there are immediate, directly attributable effects of an activity; and on the other, there are indirect results that take longer to develop and can be influenced by external factors. For example, the Shelter Cluster (GSC, 2013) uses both of the following metrics to measure outcomes:

- ‘the number of a target population using the non-food assistance they received to meet other household needs’
- ‘the number of a target population that are more resilient as a result of the non-food assistance they received’.

The current definition also groups together both simple and complex outcomes. For example, a reduction in medical complications during birth as a result of women being provided with antenatal care is a direct result. In comparison, a reduction in the occurrence of gender-based violence as a result of men and women being provided with psychosocial care is more complex, indirect behavioural change.

The problem with this broad definition is that it allows organisations to meet donor accountability requirements by focusing on short-term outcome measurements to the detriment of understanding medium-term or complex outcomes of their work. In practice, when timelines are short and resources stretched, the shorter-term outcomes are seen as the cheaper and easier option to measure (Hatton and Schroeder, 2007). As a result, the sector fails to accumulate enough data on indirect effects to contextualise and understand the story behind and around these outcomes. As expressed by one interview, ‘We only evaluate our own little area and focus which doesn’t take into consideration the whole picture. For example when looking at a wash programme, the project may be considered a success because everyone has clean water – but no-one has a house.’

In truth, it is important to monitor short-term outcomes, but the type of outcome needs to be determined by context, not convenience. And while sometimes it can be hard to understand the difference between short-term, medium-term, simple and complex outcomes at the design stage, they usually require different measurement approaches. So, one way to tease them apart might be to start by looking at the measurement tools at the outset, to identify what type of outcome is being talked about.

Formulating specific outcomes

Interviews with M&E practitioners at country level suggested that teams do not always have the time and capacity to develop a detailed ToC for each outcome they commit to, so the exercise can often be reduced to a tokenistic part of project funding proposals. This can lead to formulation of outcomes that are difficult to measure or may not say much about the project.

Some organisations have tried to address this by developing standardised or recommended ToCs to cover their main areas of intervention at a global level. For example, IRC dedicated three years to the development of its ‘Outcomes and Evidence Framework’, which has been used globally since 2016. Country teams must select ToCs from this framework when designing their programming. These have been designed to be broad enough to be applicable in most contexts. At NRC, global teams have developed recommended sectoral ToCs that country teams can adapt or contextualise when they design their country-level strategic outcomes. But they still have several mandatory outcomes for various levels of the ToC. Likewise, the DRC global office carried out ToC workshops in 2017 and 2018 to support programme development and proposal design in Iraq, Somaliland and the Sahel.

Although this work appears to be helpful, there are challenges on the consistency of application and updating underlying assumptions. IRC spoke of difficulties ensuring that the recommended ToC is aligned with suitable indicators at country level. Even when indicators do match up well to the ToC, they aren't always reported on by country teams as requested – which in turn means there's not enough data to test or update the validity of the full programmatic ToCs. NRC has concluded that it takes a minimum of two years to test global ToCs, because of the time-lag in asking country offices to collect data, run analysis, incorporate revised indicators into their programmes, and feed back new data.

Moreover, as organisations start to accumulate data for these ToCs, they might struggle to interpret or understand the results. This is because they are still typically only systematically measuring one or two outcomes per ToC, without pairing them with other information to address their decision needs or use. This may in part be linked to the use of linear logframe models in humanitarian programme design. Typically, these 'RBM-style-ToCs' have only a few steps between the activity and the final impact. Networked ToCs remain quite rare, despite the fact that change in humanitarian contexts is so often complex and non-linear (Few et al., 2014; Scheers, n.d.); Valente and Lasker, 2015; Ramalingam, 2015). While it is clearly never possible to measure everything, practitioners are struggling to determine the appropriate balance between complexity and simplicity required to generate meaningful analysis.

“As humanitarian actors tackle more complex measures, they struggle to identify what to measure.”

This problem becomes even more important as humanitarian action graduates from purely life-saving activities and adapts to more protracted crises. Terms such as 'resilience', 'well-being', 'self-reliance', 'social cohesion' and 'safety' are often used but are difficult to define as they can mean such different things to different people. As humanitarian actors tackle more complex measures, they struggle to identify what to measure. For example, one case in Lebanon is trying to identify which variables need to be captured in order to measure to what extent WASH infrastructure repair has contributed to a reduction in social tension in the targeted communities.

To tackle this, a number of organisations are investing in specific efforts to formulate outcomes for protracted crises. For example, multiple organisations are coordinating at the global level to conceptualise meeting 'basic needs'. The FAO has set up a specific task team at the global level in part to look at how to measure resilience outcomes, and WFP's Vulnerability Analysis and Mapping unit is working with the R4 Resilience Project to understand how project outcomes reach into longer-term impact. The Refugee Self-Reliance Initiative is a joint effort by a coalition of



Photo credit: Kibae Park/UN Photo.

organisations, government agencies, foundations, research institutes and other partners that is testing ways to best measure and monitoring self-reliance outcomes. Similarly, HI is piloting a multi-sector survey tool to understand people's 'quality of life', based on a global vision regardless of different technical dimensions of interventions. Although this tool is being tested for development programmes, HI plans to investigate if this can be applied or adapted to humanitarian settings – such as camp environments.

Handling sectoral silos

Much of the work on outcomes to date has been organised at the sectoral level. This has shaped the design of both ToCs, indicators and measurement tools. And it applies both within and across organisations working within the same sector.

For example, DRC now has a global framework with recommended indicators for economic recovery and CRS has a regional guidance for measuring child psychosocial wellbeing in Europe, Middle East and Central Asia. In some cases, organisations adapt and apply the work of other actors, such as HI's adaptation of the United Nations World Health Organisation's health outcome scales, or DRC's adaptation of SC's child protection psychosocial outcomes alongside the Inter-Agency Standing Committee indicators for protection. Over time, certain tools and measurement approaches have become more and more widely accepted as good practice across organisations. WFP's food security tools are now used across its 80 country interventions and have been adopted by implementing partners and other food security actors across the sector. And in some specific sectors, practitioners also borrow tried and tested tools from outside the humanitarian field and contextualise them slightly to their own needs, such as the universal measurements for health and rehabilitation outcomes which are taken directly from medical science.

Although these developments are promising, it is important to remember those aspects that overlap sectors...

To a degree, sectoral thinking of this type has encouraged coordinated efforts to improve outcomes measurements. Examples include the Cash Learning Partnership's work to pool together, test and recommend outcomes measurements for cash-based interventions and the Shelter Cluster efforts to finance a longitudinal outcomes monitoring study.

Although these developments are promising, it is important to remember those aspects that overlap sectors. For example, WFP now recognises that livelihoods and capacity to meet essential needs has a big influence on food security outcomes and therefore needs to be understood as a cross-cutting issue (WFP, 2018). Similarly, NRC's updated accelerated education ToC highlights the contributions of shelter, WASH, food security and information, counselling and legal assistance (ICLA) to education outcomes. This was also evidenced in a recent thematic review of the data.

2.2 Learning how to measure them

Aggregating programme outcomes from project outcomes

Many organisations have tried to think about measuring longer-term, programmatic outcomes beyond projects. Most have done so by developing country and global-level strategic plans (SC, 2016; IRC, 2015; NRC, 2018; WFP, 2017; UNICEF, 2018). Understandable as this approach is, it also brings its own set of challenges, particularly where project-level outcomes prove impossible to ‘add up’ to the global level, or where the project funding timeframes don’t align with strategic reporting timeframes.

Aggregation

The most appealing route for many organisations to take has been to try to add up data from project level indicators to the country or global level. Often this is done by harmonising indicators and sometimes even measurement tools. The idea is attractive as a ‘quick win’, which can maximise pre-existing project monitoring mechanisms, avoid overburdening country teams with reporting and save on the costs of developing entirely new systems. It is now common practice for organisations to either mandate a few standard project indicators for areas of intervention or to allow country teams to choose from a larger menu of ‘approved’ indicator options when designing their projects. The idea is that this will save country teams reinventing the wheel, and also allow country or global M&E teams to access this data and perform higher level analysis. This has been particularly popular among directors and senior management requesting outcome-level data. M&E staff feel that they often need to temper these senior-level expectations with what is technically feasible, while simultaneously providing solutions that can help country teams understand their results story.

“Without consistent application of tools and methods, simply combining indicators fails to produce useful information and can even be misleading.”

Indeed, the road to harmonisation has proven bumpy. Organisations that have managed to get agreement on a set of standard indicators are still struggling with how to ensure that they are being measured in the same way. Without consistent application of tools and methods, simply combining indicators fails to produce useful information and can even be misleading. Country offices often require significant guidance, instruction and mentoring to measure these indicators in a robust and timely way, but few organisations have clear definitions and protocols for outcome indicators. WCH and WCUK have developed specific guidance and tools

for their new monitoring framework – parts of which cannot be adapted by country teams, so as to ensure consistency in the structure of the data. In another approach, NRC has sample data collection forms for some of their sectoral assessment areas on their Kobo library that country offices can use or adapt. Indikit is also an interesting resource of indicator menus that provide measurement guidance. Even where guidance and training are clear and comprehensive, results may not always be valid. In order to manage this challenge, a few organisations conduct data audits or systematic data quality checks. However, many organisations also have limited capacity at the global level to undertake analysis (typically just one or two staff who look at this data once a year), so most global level M&E counterparts said that they have limited oversight of data quality.

While several organisations are still working towards global harmonisation, such as IRC and SC, others have made the conscious decision to go down a different route. Some agencies even feel that not only is harmonisation a long way off – it may not even be desirable, as all emergency projects vary so much by context that harmonising measurements would not be useful for understanding real outcomes. Oxfam decided that outcome-level data on indicators on their own could not evidence Oxfam's effectiveness without being complemented by sufficiently rigorous evaluation designs. Oxfam undertook an indicator feasibility study in 2006, but:

ultimately it was felt that this would not deliver the information that the organisation needed. The costs involved in quality assuring global outcome data – ensuring that the indicators were commonly understood and measured consistently – was felt to be prohibitively high. And, perhaps more importantly, it was recognised that while tracking changes in outcome indicators would allow the organisation to understand and communicate changes in the contexts in which it is working, crucially it would not allow Oxfam GB to unpack what, if any, contribution its interventions had made to those changes... Additionally, there was concern that requiring programmes to collect data on pre-set global outcome indicators had the potential to distort programme design, and would be at odds with the value Oxfam GB places on developing programmes 'bottom-up', based on robust analyses of how change happens in the contexts in which it is working. (Hutchings, 2014: 1; Hughes and Hutchings, 2011)

NRC made a similar decision not to fully standardise outcome indicators at the global level, instead viewing outcome indicators as just one part of a more flexible and country-specific M&E framework. This means that country offices need to be able to adjust indicators where appropriate; for example, by tailoring indicators against the type of shelter provided to displaced populations and determining the appropriate amount of

time to wait to measure occupancy rates on that basis. However, NRC does undertake secondary analysis at the global level – for example of effectiveness overall, by measuring the percentage of education programmes which achieve target pass rates for exams by more or less than 10%. But crucially, to answer these global questions on effectiveness, NRC has developed additional data collection and learning mechanisms. This includes two types of ‘global learning moments’. At a programmatic level, NRC undertakes annual contribution mapping with each of its sectors to analyse both output and outcome data (using secondary analysis or case study analysis of outcomes alongside evidence from country office evaluations). At a senior management level, NRC conducts three thematic reviews per year to dig into existing monitoring and evaluation data to answer global programme or strategic questions.



Photo credit: Albert Gonzalez Farran/UN Photo.

“

While several organisations are still working towards global harmonisation... some agencies even feel that not only is harmonisations a long way off - it may not even be desirable.

”

Funding cycles

A second challenge presented by aggregating programme outcomes from project-level measurements arises when programme-level outcomes do not align with project funding cycles and procedures. Individual humanitarian grants often do not last long enough for medium-term outcomes to be seen within the project lifespan. But in protracted crises, where one project often rolls into the next, some organisations have developed country-level strategic outcomes that are measured on a bi-annual or annual basis, in addition to existing donor-required outcome measurements. For example, NRC and SC use the concept of a ‘master’ or ‘macro’ logframe at the country level in a number of humanitarian responses. SC finds that this concept is gaining traction across the organisation, but systematic use is still a work in progress.

“...programme-level outcomes do not align with project funding cycles and procedures.”

This approach has two benefits. First, if one project is not long enough to achieve outcome-level change, the intended outcome is not lost but is still relevant and transferable to other projects with other longer-term sources of funding. Second, this allows complementary projects to be assessed together (such as if one donor funds the rehabilitation of classrooms, while another donor funds teacher training, both projects work towards improved education). This requires all project indicators to be harmonised at least at the country level with the overarching programmatic or country indicator, and to be replicated across projects.

Others have attempted to de-link more strategic outcome measurement from indicator reporting. For example, UNICEF does not approach outcome monitoring on a project-by-project basis, but rather looks at a broader ToC at the ‘programme’ level across multiple partners and projects. Oxfam conducts a number of ‘Effectiveness Reviews’ every few years to complement global output monitoring (Oxfam GB, 2014); these consist of ‘intensive evaluations that consider the extent to which projects have contributed to change in relation to the selected global outcome indicator.’ WFP also synthesises individual evaluations into country portfolio pieces, which is seen as measuring outcomes at aggregate level. NRC also conducts evaluation synthesis on an ad hoc basis for global programme analysis. Since 2015, they have done two for education, one for cash and one for ICLA. This synthesis work need not be limited to evaluations. NRC also completed an in-depth contribution analysis for education by looking at all monitoring and evaluation data across three years. DRC also piloted a de-centralised programme outcome measurements approach, where three countries volunteered to take part. They were mandated to choose one programmatic outcome that they would measure over a year’s time, with access to a small amount of central funding. Although a step in the right direction, the study required more mentoring and follow-up from the global team than expected, to ensure that country teams completed the data collection and analysis.

The need for more open-ended enquiry

The focus on reporting against indicators outlined in [chapter 1](#) has led in many cases to a prioritisation of quantitative methods of data collection. While valuable in their own right, quantitative measurement tools present two principle constraints for outcomes monitoring: limited explanatory power regarding why changes have happened, and blind-spots regarding changes that were not expected.

The focus on quantitative data has made it hard to provide a broader analysis of behavioural changes and why they occur (ALNAP, 2019; Brikci and Green, 2007; WFP, 2019; ACAPS, 2012). As discussed in Sundberg (2019), there is a growing acceptance across the sector that mixed methods with qualitative approaches are necessary for a wider understanding of context, culture and the changes caused by humanitarian programming. For example, after many years of collecting food consumption scores, WFP have noticed challenges in interpreting why changes in results have occurred over time and across different regions. To address this, they are investing in the development of a corporate technical guidance on qualitative monitoring, as well as designing and piloting new qualitative training for field-level staff.

In addition, quantitative data can make it hard to measure changes not anticipated in the original ToC. By restricting monitoring approaches to pre-defined and closed data collection methods, organisations are struggling to identify unintended effects, whether positive or negative. Saferworld and Oxfam have both recognised that achieving outcomes is often unpredictable. Both organisations have attempted to address this challenge by implementing outcome mapping and outcome harvesting approaches in some programmes (Saferworld, 2016; Oxfam, 2017).



Definition: Outcome harvesting

Outcome harvesting seeks to cast a wider net than standard monitoring practices, by looking to capture things that are beyond the control of the individual organisation:

Outcome Harvesting does not measure progress towards predetermined outcomes or objectives, but rather collects evidence of what has been achieved, and works backward to determine whether and how the project or intervention contributed to the change. (Wilson-Grau and Heather Britt, 2012: 1)



Definition: Outcome Mapping

Outcome Mapping is slightly different in that it is ‘concerned with results – or “outcomes” – that fall strictly within the programmes sphere of influence’ and focuses on people and their relationships (Earl et al., 2001; Jones and Hearn, 2009; Young et al., 2018: v).

Both organisations sought to provide a flexible approach to keep up with dynamic and uncertain humanitarian environments. Saferworld committed to a five-year investment process to embed the approach through practice of new routines and systems. They have found that the approach now enables staff to be better at spotting change as it happens. For example, in South Sudan the team keep ‘Outcome Watch’ notes about things that are not quite yet outcomes but could develop (Saferworld, 2016: 8). As more actors test similar approaches, it will be interesting to see the extent to which outcome mapping will be applicable across different sectors, countries and contexts. It should be remembered that, at the moment, these tools are still only an addition to existing reporting requirements.

As more actors test similar approaches, it will be interesting to see the extent to which outcome mapping will be applicable across different sectors, countries and contexts.



Photo credit: M. Guthrie/UN Photo.

Other organisations have carried this point even further. The Shelter Cluster is exploring a new longitudinal approach which inverts the traditional monitoring lens by following specific households for a period of five to ten years to generate case study evidence. DRC, with support from Danida, is also applying an expanded methodology in the Middle East region. Their three-year 'Livelihoods Learning Journey' aims to better understand the role that NGOs can play in livelihoods support in middle-income contexts and interrogates programmatic assumptions through the application of more qualitative methods over time. Likewise, CRS has been piloting the use of the SenseMaker methodology in more complex operating environments since 2015, which has allowed them to integrate the analysis of affected populations on why and how change has occurred.

Data management

Just as organisations have sought to aggregate and harmonise indicators, many have also attempted to use one central data management system to track data across output and outcome levels. However, this has proven to be more complex than first assumed. For example, IRC is currently in the process of redesigning their global data platform. They need to balance many competing requirements, including being standardised and manageable for country teams to implement, while meeting competing country, donor and HQ information needs, and doing so in a sustainable way with available resources. For similar reasons, NRC found that they had to pilot different systems in several country offices in order to find the most appropriate system. This process took several years and significant investment.

Indeed, several M&E teams have expressed frustration that senior management do not understand the complexity behind developing these systems to get them to work. Data systems at this level demand a high level of discipline and rigid data architecture. Tailoring platforms from scratch to an organisation's needs is expensive and requires a significant investment of time from the global as well as country teams (for design and testing). Data systems can be burdensome in emergency contexts (both to set up and maintain), and capacities to understand and implement monitoring systems vary between countries: just because you have a great system does not mean that people know how to use it or will use it consistently.

Recognising these difficulties, DRC has conducted a feasibility study for regional data management solutions in the Middle East and North Africa, considering a country-level system as the foundation to be replicated elsewhere. They invested in an external data mapping assessment across the region to understand which data systems already existed and how they could be best brought together. Despite external support and enthusiasm in country and regional office, the proposed solution did not prove to be viable for scale-up as a global investment project. Instead, the local system continued to be supported or even replicated in certain countries based on country-level initiative and country-specific assessment.

2.3 Doing what it takes to get this done

Resourcing

When thinking about what it takes for organisations to develop measurement tools or data management systems, significant resourcing and staff time are necessary. Mayne (2007) found that it can take an average four to five years to embed a RBM framework within a humanitarian organisation. Indeed, organisations interviewed for this study had spent between three to five years designing, testing and rolling out their current frameworks for monitoring outcomes – and are still in the process of further developing and fine-tuning these. For example, after a year of development, it took two full-time members of NRC staff eight months to ‘roll out’ the global M&E framework, and then a further six months to embed the global reporting system. They now conduct training every few years at the regional level. Even in cases where organisations have tried to resort to less ‘indicator reporting’ and more alternative, open-ended approaches, such as outcome harvesting, this took several years to conceptualise, train and implement. And ultimately, organisations are constrained by the annual reporting cycles, which only offer one window per year to test or roll out a new system.

Most organisations only have small M&E units at the global level, typically of around two to three persons. In order for them to be able to spend the amount of time necessary to see an initiative through to the end, their senior management need to support them to spend a significant of their time on this over a couple of years. In some cases, consultants were hired to provide technical inputs. Most organisations then asked country offices to dedicate some time to providing inputs and testing the systems. This ‘surge’ of time is needed in addition to – or as a priority over – other pre-existing M&E tasks. Some teams are able to absorb these tasks better than others. To address this gap, many organisations are looking to improve their M&E capacity, such as by expanding their global teams to support country offices, designing and implementing continuous training or scaling up financing of human resources for M&E significantly in order to have stronger staff.



Photo credit: Johnathan Hymens/UN Photo.

Cultural shifts

Even once all the resources are in place, it is imperative to instil a culture of learning in order for an outcomes system to work (Behn, 2002; Mayne, 2007; Saferworld, 2016). Behn (2002) emphasises just how much it takes to achieve a cultural shift of the necessary scale. Many initiatives at this scale require significant internal advocacy efforts to ensure that colleagues understand why such systems are worth the investment.

“Refocusing humanitarian M&E systems to produce more meaningful data has to be linked to a system-wide rethink of the incentives for providing such data.”

Some donors expressed an interest in moving towards creating learning environments between donors and between donors and implementing partners. ‘Incentives are a significant part of this environment. Refocusing humanitarian M&E systems to produce more meaningful data has to be linked to a system-wide rethink of the incentives for providing such data’ (Guerrero et al., 2013: 11).

With such a heavy commitment required from both global and country-level teams, strong leadership and drive from senior management has been critical. Senior champions are particularly important to influence the allocation of resources, which often entails political negotiation between stakeholders with competing priorities. To address problems of this type, many organisations have formally endorsed strategic planning documents at the highest level.



Photo credit: Tobin Jones/UN Photo.



Conclusion

3 Conclusion

Humanitarian organisations have done a significant amount of work to improve outcomes monitoring, with notable progress at the sector level. Nevertheless, practitioners are still expressing frustration that – for all the advances made – the sector struggles to provide meaningful and holistic analysis of the results of humanitarian action.

This paper has demonstrated that much of the preceding work on outcomes monitoring has been based on four underlying assumptions, which each bring their own constraints. First, all organisations have worked on the basis that they should be measuring planned-for outcomes of specific projects – but outcomes in humanitarian settings can be unpredictable. Second, organisations widely use the ToC framework to help anticipate these outcomes – but outcomes cannot be reduced to a linear process. Third, most organisations focus on measuring pre-defined indicators – and so struggle to systematically capture unintended outcomes. Fourth, outcome measurements have been designed with the project at their core – there is currently little incentive to measure outcomes between projects or over time.

“The result is a kaleidoscope of information that – while responding to the felt need for more data – doesn’t provide a meaningful analysis of results over time and in context.”

Part of the problem with the work that has been done to date, is that the system is now overstretched. Many organisations are hoping to understand country- and sector-wide outcomes by using pre-existing monitoring systems that were originally designed to track shorter-term indicators and assess single-project performance. The result is a kaleidoscope of information that – while responding to the felt need for more data – doesn’t provide a meaningful analysis of results over time and in context.

Clearly, overcoming all of the problems discussed in this paper will be hard to achieve through incremental change. Emerging thinking on the alternative structure and funding available for M&E systems needs to be given due recognition and investment. This could mean many things: moving away from pre-defined indicators, using more cross-sectoral measurements, breaking down the definition of outcomes, building long-term evidence-gathering models or creating space for more open-ended enquiry. It might even be time to look at different business models for monitoring and evaluation across the sector.

Bibliography

ACAPS. (2012) *Qualitative and quantitative research techniques for humanitarian needs assessment: An introductory brief*. Geneva: ACAPS. (www.alnap.org/help-library/qualitative-and-quantitative-research-techniques-for-humanitarian-needs-assessment-an).

Agenda for Humanity. (2016) 'New way of working'. [Webpage]. 17 May. New York: OCHA. (<https://www.alnap.org/help-library/new-way-of-working-0>).

ALNAP. (2018) *The state of the humanitarian system*. ALNAP Study. London: ALNAP/ODI. (www.alnap.org/help-library/the-state-of-the-humanitarian-system-2018-full-report).

Behn, R. (2002) 'The psychological barriers to performance management: Or why isn't everyone jumping on the performance-management bandwagon?' *Public Performance and Management Review*, 26(1): 5–25. (www.alnap.org/help-library/the-psychological-barriers-to-performance-management-or-why-isnt-everyone-jumping-on).

Benini, A. (2018) *Subjective measures in humanitarian analysis*. Geneva: ACAPS. (www.alnap.org/help-library/subjective-measures-in-humanitarian-analysis).

Bennett, C. (2015) *The development agency of the future. Fit for protracted crises?* London: ODI. (www.alnap.org/help-library/the-development-agency-of-the-future-fit-for-protracted-crises).

Brikci, N. and Green, J. (2007) *A guide to using qualitative research methodology*. Geneva: MSF. (www.alnap.org/help-library/a-guide-to-using-qualitative-research-methodology).

Brown, D. and Johnson, C. (2015) 'Setting a new research agenda for urban crisis and humanitarian response'. *Briefing*, June. (www.alnap.org/help-library/setting-a-new-research-agenda-for-urban-crisis-and-humanitarian-response).

Cash Learning Partnership. (2018) 'Guidance and toolbox for the basic needs analysis'. [Webpage]. 20 June. Oxford: CaLP. (www.alnap.org/help-library/guidance-and-toolbox-for-the-basic-needs-analysis).

Cash Learning Partnership. (2019) 'Monitoring & evaluation for cash and voucher assistance'. [Webpage]. 17 May. Oxford: CaLP. (www.alnap.org/help-library/monitoring-evaluation-for-cash-and-voucher-assistance).

Catholic Relief Service. (2010) *ProPack III: A guide to creating a SMILER M&E system*. Baltimore: CRS. (www.alnap.org/help-library/propack-iii-a-guide-to-creating-a-smiler-me-system).

Core Humanitarian Standard. (2014) *Core humanitarian standard on quality and accountability*. Geneva: CHS Alliance/Group Urd/the Sphere Project. (www.alnap.org/help-library/core-humanitarianstandard-on-qualityand-accountability).

Corlazzoli, V. and White, J. (2013) *Back to basics: A compilation of best practices in design, monitoring and evaluation in fragile and conflict-affected environments*. London: DFID. (www.alnap.org/help-library/back-to-basics-a-compilation-of-best-practices-in-design-monitoring-evaluation-in).

Danish Refugee Council. (2015) 'Monitoring, evaluation and learning minimal operational procedures'. [Website]. 17 May. Copenhagen: DRC. (www.alnap.org/help-library/monitoring-evaluation-and-learning-minimal-operational-procedures).

Danish Refugee Council. (2016) 'Striking the right balance for monitoring and evaluation in a complex organisation'. *Evaluation and Learning Brief*, 10. Copenhagen: DRC. (www.alnap.org/help-library/evaluationlearning-brief-10-striking-the-right-balance-for-monitoring-and-evaluation-in).

Darcy, J., Stobaugh, H., Walker, P. and Maxwell D. (2013) *The use of evidence in humanitarian decision making*. Geneva: ACAPS. (www.alnap.org/help-library/the-use-of-evidence-in-humanitarian-decision-making).

Development and Cooperation. (2011) 'Personal perceptions make a difference'. [Blog]. 17 May. Bonn: D+C. (www.alnap.org/help-library/personal-perceptions-make-a-difference).

Development Initiatives. (2016) *Humanitarian evidence systems mapping in East Africa*. Bristol: DI. (www.alnap.org/help-library/humanitarian-evidence-systems-mapping-in-east-africa).

Dillon, N. (2019) *Breaking the Mould: Alternative approaches to monitoring and evaluation*. ALNAP Paper. London: ALNAP/ODI.

Earl, S., Carden, F. Smutylo, T. (2001) *Outcome Mapping: A Guide to Building Learning and Reflection into Development Programs*. Ottawa: IDRC. (www.alnap.org/help-library/outcome-mapping-building-learning-and-reflection-into-development-programs).

ECHO. (n.d.) *List of key outcome indicators*. Brussels: ECHO. (www.alnap.org/help-library/list-of-key-outcome-indicators).

FAO. (n.d.) *Managing results at FAO orientation guide*. Rome: FAO. (www.alnap.org/help-library/managing-results-at-fao-orientation-guide).

FAO. (2019) *Savings livelihoods saves lives 2018*. Rome: FAO. (www.alnap.org/help-library/savings-livelihoods-saves-lives-2018).

Few, R., McAvoy, D., Tarazona, M. and Walden, V. (2014) *Contribution to change: An approach to evaluating the role of intervention in disaster recovery*. London: Oxfam. (www.alnap.org/help-library/contribution-to-change-an-approach-to-evaluating-the-role-of-intervention-in-disaster).

Gaston, E. (2017) *Harmonizing donor reporting*. Berlin: GPPi. (www.alnap.org/help-library/harmonizing-donor-reporting).

Global Affairs Canada. (n.d.) 'Annex 14: SRHR key performance indicators'. [Webpage]. 17 May. Toronto: GAC. (www.alnap.org/help-library/annex-14-srhr-key-performance-indicators).

Global Public Policy Institute. (2016) *Monitoring and evaluation in secure contexts: back to basics?* Berlin: GPPi. (www.alnap.org/help-library/monitoring-and-evaluation-in-insecure-contexts-back-to-basics).

Global Shelter Cluster. (2013) *Shelter cluster indicator guidelines: Version 2 – Draft*. October 2013. New York: IASC. (www.alnap.org/help-library/shelter-cluster-indicator-guidelines-version-2-%E2%80%93-draft).

Gray, I. (2017) *A quick guide to theory of change*. Edinburgh: Scotland's International Development Alliance. (<https://www.alnap.org/help-library/a-quick-guide-to-theory-of-change>).

Hallam, A. and Bonino, F. (2013) *Using evaluation for a change: Insights from humanitarian practitioners*. ALNAP Study. London: ALNAP/ODI. (www.alnap.org/help-library/using-evaluation-for-a-change-insights-from-humanitarian-practitioners).

Harvey, P. and Pavanello, S. (2018) *Multi-purpose cash and sectoral outcomes: a review of evidence and learning*. Geneva: UNHCR. (www.alnap.org/help-library/multi-purpose-cash-and-sectoral-outcomes-a-review-of-evidence-and-learning).

Hatton, M. J. and Schroeder, K. (2007) 'Results-based management: friend or foe?'. *Development in Practice*, 17:3: 426-432. (www.alnap.org/help-library/results-based-management-friend-or-foe).

Hearn, S. and Jones, H. (2009) *Outcome mapping: a realistic alternative for planning, monitoring and evaluation*. London: ODI. (www.alnap.org/help-library/outcome-mapping-a-realistic-alternative-for-planning-monitoring-and-evaluation).

Hofmann, C. A., Roberts, L., Shoham, J. and Harvey, P. (2004) 'Measuring the impact of humanitarian aid. A review of current practice'. *Humanitarian Policy Group Report*, 17. London: ODI. (www.alnap.org/help-library/measuring-the-impact-of-humanitarian-aid-a-review-of-current-practice).

Hughes, K. and Hutchings, C. (2011) *Can we obtain the required rigour without randomisation?*. London: 3ie. (www.alnap.org/help-library/can-we-obtain-the-required-rigour-without-randomisation).

Humanitarian Innovation Fund. (2014) *Impact guidelines – Humanitarian innovation fund*. London: ELRHA. (www.alnap.org/help-library/impact-guidelines-humanitarian-innovation-fund).

Humanitarian Policy Group. (2016) *Time to let go. Remaking humanitarian action for the modern era*. London: ODI. (www.alnap.org/help-library/time-to-let-go-remaking-humanitarian-action-for-the-modern-era).

Hutchings, C. (2014) 'Balancing accountability and learning: A review of Oxfam GB's global performance framework'. *Journal of Development Effectiveness*, 6(4). (www.alnap.org/help-library/balancing-accountability-and-learning-a-review-of-oxfam-gb%E2%80%99s-global-performance).

Indkit. (2019) 'Guidance on SMART indicators'. [Webpage]. 17 May. Vinohradý: People in Need. (www.alnap.org/help-library/guidance-on-smart-indicators).

IASC. (2015) *Guidance: Humanitarian response monitoring*. New York: IASC. (www.alnap.org/help-library/guidance-humanitarian-response-monitoring).

IASC. (2019) 'Grand bargain (hosted by the IASC)'. [Webpage]. 17 May. New York: IASC. (www.alnap.org/help-library/grand-bargain-hosted-by-the-iasc).

IFRC. (2011) *Project/programme monitoring and evaluation (M&E) guide*. Geneva: IFRC. (www.alnap.org/help-library/projectprogramme-monitoring-and-evaluation-me-guide).

INTRACT. (2014) 'Context and sustainability: Monitoring and evaluating humanitarian aid'. *The Newsletter of INTRAC*, 58 November. Oxford: INTRACT. (www.alnap.org/help-library/context-and-sustainability-monitoring-and-evaluating-humanitarian-aid-0).

IRC. (2015) *IRC strategy 2015-2020: Executive summary*. New York: IRC. (www.alnap.org/help-library/irc-strategy-2015-2020-executive-summary).

IRC. (2016) 'Outcomes and evidence framework'. [Webpage]. 17 May. New York: IRC. (www.alnap.org/help-library/outcomes-and-evidence-framework).

Jones, L. and Tanner, T. (2016) 'Subjective resilience: Using perceptions to quantify household resilience to climate extremes and disasters'. *Regional Environmental Change*, 17(1): 229-243. (www.alnap.org/help-library/subjective-resilience-using-perceptions-to-quantify-household-resilience-to-climate).

Knox-Clarke, P. and Darcy, J. (2014) *Insufficient evidence? The quality and use of evidence in humanitarian action*. ALNAP Study. London: ALNAP/ODI. (www.alnap.org/help-library/insufficient-evidence-the-quality-and-use-of-evidence-in-humanitarian-action-alnap-0).

Mayne, J. (2007) 'Challenges and lessons in implementing results-based management'. *Evaluation*, 13(1): 87-109. (www.alnap.org/help-library/challenges-and-lessons-in-implementing-results-based-management).

Mountfield, B. (2015) *Sphere unpacked: Sphere for monitoring and evaluation*. Geneva: Sphere. (www.alnap.org/help-library/sphere-for-monitoring-and-evaluation).

NRC. (2016) *Monitoring and evaluation: Minimum standards and guidelines*. Oslo: NRC. (www.alnap.org/help-library/monitoring-evaluation-minimum-standards-and-guidelines).

NRC. (2018) *Global strategy 2018-2020*. Oslo: NRC. (www.alnap.org/help-library/global-strategy-2018-2020).

Obrecht, A. (2017) *Using evidence to allocate humanitarian resources: Challenges and opportunities*. ALNAP Working Paper. London: ALNAP/ODI. (www.alnap.org/help-library/working-paper-using-evidence-to-allocate-humanitarian-resources-challenges-and).

Obrecht, A. (2018a) *Making humanitarian response more flexible: Challenges and questions*. ALNAP Background Paper. London: ALNAP/ODI. (www.alnap.org/help-library/making-humanitarian-response-more-flexible-alnap-workshop-background-paper).

Obrecht, A. (2018b) *Dynamic gridlock: Adaptive humanitarian action in the Democratic Republic of Congo*. ALNAP Country Study. London: ALNAP/ODI. (www.alnap.org/help-library/dynamic-gridlock-adaptive-humanitarian-action-in-the-drc).

OCHA. (2018) *World humanitarian data and trends 2018*. New York: OCHA. (www.alnap.org/help-library/world-humanitarian-data-and-trends-2018).

OECD. (2012) *Measuring regulatory performance*. Paris: OECD. (www.alnap.org/help-library/measuring-regulatory-performance-practitioner%E2%80%99s-guide-to-perception-surveys).

OECD/DAC. (2002) *Glossary of key terms in evaluation and results-based management*. Paris: OECD. (www.alnap.org/help-library/glossary-of-key-terms-in-evaluation-and-results-based-management).

Oxfam. (n.d.) *GB Evaluation guidelines*. Oxford: Oxfam. (www.alnap.org/help-library/oxfam-gb-evaluation-guidelines).

Oxfam. (2017) *Limited access humanitarian programming: operational guidance for managing programme quality*. Oxford: Oxfam. (www.alnap.org/help-library/limited-access-humanitarian-programming-operational-guidance-for-managing-programming).

Peters, K., Langston, L., Tanner, T. and Bahadur, A. (2016) *'Resilience' across the post-2015 frameworks: towards coherence?*. London: ODI. (www.alnap.org/help-library/%E2%80%98resilience%E2%80%99-across-the-post-2015-frameworks-towards-coherence).

Prowse, M. (2007) *Aid effectiveness: the role of qualitative research in impact evaluation*. London: ODI. (www.alnap.org/help-library/aid-effectiveness-the-role-of-qualitative-research-in-impact-evaluation).

Ramalingam, B. and Jones, H. (2008) *Exploring the science of complexity: Ideas and implications for development and humanitarian efforts*. London: ODI. (www.alnap.org/help-library/exploring-the-science-of-complexity-ideas-and-implications-for-development-and).

Ramalingam, B. and Mitchell, J. (2009) *Counting what counts: performance and effectiveness in the humanitarian sector: Chapter 1*. London: ALNAP/ODI. (www.alnap.org/help-library/alnaps-8th-review-of-humanitarian-action-counting-what-counts-performance-and).

Ramalingam, B. (2013) *Aid on the Edge of Chaos: Rethinking International Cooperation in a Complex World*. Oxford: Oxford University Press. (www.alnap.org/help-library/aid-on-the-edge-of-chaos).

Ramalingam, B. Wild, L. and Buffardi, A. (2019) *Making adaptive rigour work: Principles and practices for strengthening monitoring, evaluation and learning for adaptive management*. London: ODI. (www.alnap.org/help-library/making-adaptive-rigour-work-principles-and-practices-for-strengthening-monitoring).

Rogers, P. (2014) *Theory of change*. New York: UNICEF. (www.alnap.org/help-library/theory-of-change).

Saferworld. (2016) *Doing things differently*. London: Saferworld. (www.alnap.org/help-library/doing-things-differently).

Save the Children. (2014a) *2 Save the Children (MEAL) system overview*. Milton Keynes/London: The Open University/SC. (www.alnap.org/help-library/2-save-the-children-meal-system-overview).

Save the Children. (2014b) *18 measuring results in humanitarian work*. Milton Keynes/London: The Open University/SC. (www.alnap.org/help-library/18-measuring-results-in-humanitarian-work).

Save the Children. (2016) *Save the Children's global strategy: Ambition for children 2030 and 2016-2018 strategic plan*. London: SC. (www.alnap.org/help-library/save-the-children%E2%80%99s-global-strategy-ambition-for-children-2030-and-2016-2018-strategic).

Save the Children. (2019) 'Indicators'. [Webpage]. 17 May. London: SC. (www.alnap.org/help-library/indicators).

Scheers, G. (n.d.) *Measuring success in complex settings: GPPAC's experiences with planning, monitoring and evaluation*. The Hague: GPPAC. (www.alnap.org/help-library/measuring-success-in-complex-settings-gppac%E2%80%99s-experiences-with-planning-monitoring-and).

Simister, N. (2015) *Monitoring, evaluation and learning system review: Progress report #3*. Copenhagen: DRC. (www.alnap.org/help-library/monitoring-evaluation-and-learning-system-review).

Sundberg, A. (2019) *Beyond the numbers: How qualitative approaches can improve monitoring of humanitarian action*. ALNAP Paper. London: ALNAP/ODI.

Ticehurst, D. (2012) *Who is listening to who, how well and with what effect?*. London: ALNAP/ODI. (www.alnap.org/help-library/who-is-listening-to-who-how-well-and-with-what-effect).

UNHCR. (2006) *Practical guide to the systematic use of standards and indicators in UNHCR operations*. 2nd ed. Geneva: UNHCR. (www.alnap.org/help-library/practical-guide-to-the-systematic-use-of-standards-indicators-in-unhcr-operations).

UNICEF. (2011) *Humanitarian performance monitoring toolkit*. New York: UNICEF. (www.alnap.org/help-library/humanitarian-performance-monitoring-toolkit).

UNICEF. (2017) *Accountability to affected populations: Analytical paper on WHS self-reporting on the Agenda for Humanity*. New York: UNICEF. (www.alnap.org/help-library/accountability-to-affected-populations-analytical-paper-on-whs-self-reporting-on-the).

UNICEF. (2018) *UNICEF strategic plan 2018-2021*. New York: UNICEF. (www.alnap.org/help-library/unicef-strategic-plan-2018-2021).

USAID/OFDA. (2018) *Proposal guidelines*. Washington, D.C.: USAID/OFDA. (www.alnap.org/help-library/proposal-guidelines).

Valente, R. and Lasker, R. (2015) *An end in sight: multi-year planning to meet and reduce humanitarian needs in protracted crises*. New York: OCHA. (www.alnap.org/help-library/an-end-in-sight-multi-year-planning-to-meet-and-reduce-humanitarian-needs-in-protracted).

War Child UK. (2014) *Global monitoring and evaluation manual*. London: War Child UK. (www.alnap.org/help-library/global-monitoring-and-evaluation-manual).

Warner, A. (2017a) *What is monitoring in humanitarian action? Describing practice and identifying challenges*. London: ALNAP/ODI. (<https://www.alnap.org/help-library/what-is-monitoring-in-humanitarian-action-describing-practice-and-identifying>).

Warner, A. (2017b) *Monitoring humanitarian innovation*. HIF/ALNAP Working Paper. London: ALNAP/ODI. (www.alnap.org/help-library/monitoring-humanitarian-innovation-hif-alnap-working-paper).

WFP. (2017) *WFP strategic plan 2017-2021*. Rome: WFP. (www.alnap.org/help-library/wfp-strategic-plan-2017-2021).

WFP. (2018) *Revised corporate results framework (2017-2021)*. Rome: WFP. (www.alnap.org/help-library/revised-corporate-results-framework).

WFP. (2019) *Qualitative research guidance for WFP monitoring: Performance management and monitoring division*. Internal document. Rome: WFP.

Willitts-King, B., Assomull, R., Bryant, J., McCartney, C., Dhimi, T. and Llewellyn, D. (2019) *New financing partnerships for humanitarian impact*. London: ODI. (www.alnap.org/help-library/new-financing-partnerships-for-humanitarian-impact).

Wilson-Grau, R. and Britt, H. (2012) *Outcome harvesting*. New York: Ford Foundation. (www.alnap.org/help-library/outcome-harvesting).

Young, J., Shaxson, L., Jones, H., Hearn, S. Datta, A. and Cassidy, C. (2018) *Rapid outcome mapping approach: A guide to policy engagement and influence*. London: ODI. (www.alnap.org/help-library/rapid-outcome-mapping-approach).

Related ALNAP publications

- Breaking the Mould: Alternative approaches to monitoring and evaluation
- Beyond the Numbers: How qualitative approaches can improve monitoring for humanitarian action
- Evaluation of Protection Guide
- What is Monitoring in Humanitarian Action?
- Evaluation of Humanitarian Action Guide



ALNAP

Overseas Development Institute
203 Blackfriars Road
London SE1 8NJ
United Kingdom

alnap@alnap.org