HIF-ALNAP
WORKING PAPER

Evaluating Humanitarian Innovation

Written by Alice Obrecht

With contributions from
Alexandra Warner and Neil Dillon

HIF Humanitarian innovation fund | elrha

ALNAP

Elrha's **Humanitarian Innovation Fund (HIF)** supports organisations and individuals to identify, nurture, and share innovative and scalable solutions to the challenges facing effective humanitarian assistance. The HIF is an Elrha programme and is supported by funding from the UK Department for International Development (DFID), the Swedish International Development Cooperation Agency (SIDA) and the Netherlands Ministry of Foreign Affairs (MFA). **www.elrha.org/hif**

ALNAP is a unique system-wide network dedicated to improving humanitarian performance through increased learning and accountability. **www.alnap.org**

About the authors

Alice Obrecht is a Research Fellow at ALNAP

Alexandra Warner is a Research Officer at ALNAP

Neil Dillon is a Research Fellow at ALNAP

Suggested citation

Obrecht, A. with Warner, A. and Dillon, N. (2017) 'Working paper: Evaluating humanitarian innovation' HIF/ALNAP Working Paper. London: ODI/ALNAP.

Design and typesetting by Alex Glynn

Copyediting by Roo Griffiths

Contents

About HIF-ALNAP's research on innovation in humanitarian action

Over 2015-6 ALNAP - in partnership with Elrha's Humanitarian Innovation Fund (HIF) - looked at how to define innovation in humanitarian action, and what successful innovation looks like. 15 innovations in the humanitarian sector were chosen as case studies to provide an empirical evidence base for the final research study **'More than just luck: innovation in humanitarian action'**.

Following on from these foundations, ALNAP and the HIF developed the implications of their research findings for monitoring and evaluating humanitarian innovation processes, producing two working papers on these topics.

The ultimate aim of the research is to improve humanitarian actors' understanding of how to undertake and support innovative programming in practice. This research partnership builds on ALNAP's long-running work on innovation in the humanitarian system, beginning with its 2009 study, Innovations in International Humanitarian Action, and draws on the experience of the HIF grantees, which offer a realistic picture of how innovation actually happens in humanitarian settings.

About the case studies

15 case studies, were undertaken by ALNAP in partnership with Elrha's Humanitarian Innovation Fund (HIF), exploring the dynamics of successful innovation processes in humanitarian action. They examine what good practice in humanitarian innovation looks like, what approaches and tools organisations have used to innovate in the humanitarian system, what the barriers to innovation are for individual organisations, and how they can be overcome.

The case study subjects were chosen to reflect innovation practice in the humanitarian system. They covered information communication technology (ICT) innovations and non-ICT innovations, and offered a balance between innovations that have reached a diffusion stage and those that had not. They also reflected the wide geographic range of the areas where innovations are being trialled and implemented.

About 'More than just luck: innovation in humanitarian action' research paper

'More than just luck: innovation in humanitarian action' presented the synthesised findings from the 15 case studies, focusing on three main questions:

- What is innovation in humanitarian action?
- What does success in humanitarian innovation look like?
- What can humanitarians do to achieve success?

Acknowledgements

The authors wish to acknowledge their debt to the peer reviewers and commentators, whose input to this paper greatly improved its quality: Josiah Kaplan (Elrha), Claire Dusonchet (Elrha), Paul Knox-Clarke (ALNAP), Robert McCouch (United Nations OIOS-IED), Marpe Tanaka (MSF Sweden Innovation Fund), Josephine Tsui (RAPID, ODI). All errors or inaccuracies are the author's.

1. Introduction

For funders, evaluating humanitarian innovation is critical, given the higher risks involved in innovation projects; for others, including innovators themselves, evaluating innovation can appear to be an oxymoron. Evaluations are the systematic and objective examination of an activity, policy or programme to determine its worth or significance. They often rely on a fixed set of criteria and a clear focus of 'what' is being evaluated. In contrast, innovation processes are characterised primarily by their open and iterative structure: innovation teams are expected to learn and make significant changes to their intended intervention or product along the way, rather than stick to an initial set of aims or criteria.

The core challenge for evaluators of humanitarian innovation (HI) is to correctly recognise and incorporate, rather than ignore or suppress, the value of iteration when carrying out their formal evaluations. Currently, humanitarian evaluations rarely tackle innovation explicitly. Where they do, they often translate an innovation process into the concepts and language of a standard humanitarian project. As a result, evaluations of innovation rarely distinguish between innovation processes and their outputs, and they overlook many of the features most relevant to assessing the worth of innovation, such as iterative learning, impact, being problem-/user-driven and the promise of a 'step change' or comparative improvement to current practice.

Humanitarian evaluators will need to consider how to select from the toolbox of evaluation when doing an evaluation of innovation and adapt some of their thinking around theories of change and evaluative criteria in order to offer a meaningful assessment of the worth or value of an innovation process.

This relationship also needs to work in the opposite direction, with humanitarian innovators becoming more familiar with and open to evaluative tools and thinking. Evaluation practice provides key tools for learning and communicating the value of a project or process to external audiences; humanitarian innovators could capitalise much better on these benefits.

This paper aims to make advances in both areas, addressing the perspective of both the evaluator and the innovator. Sections 2 and 3 offer an introduction to the range of evaluative practice that may be relevant to innovation processes and that humanitarian innovators can use, as an overview of how they can better utilise evaluative inquiry for their own purposes. The remaining Sections 4–7 primarily address evaluators of humanitarian action, providing an in-depth look at the concepts and evaluative criteria that are relevant to carrying out a summative evaluation of HI – that is, an evaluation that judges the merit or worth of a programme at its conclusion (EHA, 2016: 63).

The Annexes contain more content on developmental evaluation (Annex I), on assessing the 'innovativeness' of an innovation (Annex II) and on assessing the scalability of an innovation (Annex III).

There has been very little evaluation of HI to date, and therefore this paper does not draw heavily on previous evaluations or evaluative practices. Instead, the concepts and analysis draw on the 15 case studies on innovation management in the humanitarian system and synthesis report (**More Than Just Luck**) produced through ALNAP's research partnership with Elrha's Humanitarian Innovation Fund (HIF). To write this paper, ALNAP reviewed key learning on how these 15 cases of HI used evaluative inquiry and formal evaluations, identified key strengths and gaps and reflected on the adaptations to evaluative practice that might facilitate better-quality evaluations of HI.

The content of the paper is therefore suggestive: it proposes a way of framing and conceptualising the evaluation of HI but does not offer tested guidance; it should therefore be read as a discussion paper, in complement to ALNAP's Evaluation of Humanitarian Action (EHA) Guide.

The paper has two intended end users:

1. **Evaluators and evaluation staff, including staff at donor institutions, who are seeking to evaluate an innovation project or programme:** The purpose of the paper is to suggest an overarching framework for HI evaluation for evaluators to consider in their practices.
2. **Managers of an HI:** The paper also introduces managers of an HI to the opportunities for incorporating evaluative inquiry into their own practices. It can also be used as a tool for communicating with evaluators hired to assess their innovation. This working paper focuses only on high-level issues and should be read in companionship with ALNAP's EHA Guide.

2. Evaluating humanitarian innovation: Challenges and opportunities

2.1. Defining humanitarian innovation

The term 'innovation' is used to refer to a variety of practices. We define HI as:

An iterative process that identifies, adjusts and diffuses ideas for improving humanitarian action.

Evaluators may be asked to assess this process, or to assess the primary outputs of an innovation process: the innovation itself, typically represented by a prototype of a product, process, position or paradigm (see Box 1), or the consolidated learning and evidence produced by the innovation process. Evaluators may also be asked to carry out an evaluation of the impact of an innovation after it has been diffused. Section 4 describes these three distinct focus areas in greater detail.

The difference between humanitarian innovation and standard programming

Understanding how to evaluate HI requires an understanding of how innovation as an activity differs from standard projects and programming. In standard programming, there is expected to be a robust understanding of the expected causal pathway for the improvement a programme is going to bring about. This can be supported by prior evaluations or experience. Previous applications of the intervention or tool in the same or similar context can be used to construct theories of change for how the interventions are expected to bring about the desired results. For example, an organisation that begins to use cash-based programming in Lebanon can draw on previous applications of cash-based programming in other humanitarian settings, as well as previous cash-based programmes other agencies have potentially used within the Lebanon context. Even if a humanitarian programme lacks an explicit theory of change, this is the informal logic that supports most standard programming in humanitarian contexts. In standard programming, managers implement an intervention because it was implemented before, and because it is expected to 'work'.

In contrast, in an innovation process, the causal pathway for change is unknown. Innovation projects can construct a general theory of change but the assumptions are more conjectural, making the theory much more

Box 1: Speaking the language of innovation: Key innovation terms and concepts

The '4 Ps' of Innovation: Innovations are typically organised into four types, depending on where the proposed improvement or change is occurring (Tidd and Bessant, 2005)*:*

- **Product innovation**: Changes in the things (products/services) an organisation offers
- **Process innovation:** Changes in the ways products and services are created or delivered
- **Position innovation:** Changes in the context in which the products/services are framed and communicated
- **Paradigm innovation:** Changes in the underlying mental models that shape what the organisation does

Value proposition: The value proposition is a short description of an innovation. It is a clear statement that describes the innovation and its proposed value. Value propositions can vary in their level of detail, but should include at least five components: 1) what the innovation is; 2) its characteristics; 3) the change it is expected to bring; 4) the performance area or problem it is addressing; and 5) why that performance area or problem matters (Kimball, 2014: 124).

Proof of concept: A proof of concept is an early test applied to a single idea or key assumption of the overall innovation, to understand whether the innovation as a whole might be physically or conceptually possible to achieve. Doing a good proof of concept is more of an art than a science and requires a careful selection of the most questionable assumptions on which an innovation's value proposition rests.

Prototype: A prototype is a working model of the innovation that can be implemented in a pilot. It is used for a 'real-world' test of the innovation. While prototypes are often associated with tangible 'products', they also feature in process, position and paradigm innovations.

Pilot: A pilot is an initial implementation of an innovation in conditions that closely mimic, or instantiate, the conditions in which the innovation is expected to be regularly used.

Diffusion/scaling: The final stage, or set of activities, that occurs within an innovation process. Diffusion and scaling are concerned with generating wider adoption of an innovation. Through these activities, innovating teams build on demonstrated successes to ensure solutions reach their maximum potential, have the greatest possible impact and lead to widespread change.

like a hypothesis. The innovation manager may not know if a new water treatment system will work or if an approach to disaster risk reduction will be successful, despite inference from the evidence base, simply because there is a good chance that no one has ever tried these interventions in a humanitarian context. Innovation relies on an explicit emphasis on learning and readjustment because so little is known about whether, how and why an idea for improvement might work.

That said, evaluators may actually find innovators are more knowledgeable of their theory of change and the evidence base for their work than managers of standard humanitarian programmes. In some cases, theories of change used in standard programming are the result of a generic 'cut and paste' approach that does not think through the context or the evidence base for the effectiveness of a particular type of intervention. Innovators, in contrast, are often forced to think more critically about their approach, the problem or opportunity that it is

responding to, and the views of end users. To that degree, evaluating innovation can offer richer opportunities for engaging with a theory of change and the evidence base supporting it.

Challenges and opportunities

There are four main challenges to evaluating humanitarian innovation. Several of these challenges are also true for standard evaluation, but their difficulties become even more pronounced in the context of innovation.

1. **Complexity:** Innovations are complex processes. They seek to address problems and exploit value in new ways, often working on issues that lack a clear structure or common understanding. To succeed, innovations must shift the behaviours of many different actors, all of which interact and influence one another in different ways. Within the area of practice addressed by an innovation process, the patterns and relationships of cause and effect are unpredictable and potentially nonlinear, and are being discovered as the innovating team progresses.

2. **Iteration:** Perhaps the most practically difficult challenge to evaluating innovation relates to finding the appropriate basis for carrying out a summative evaluation. Innovation processes rarely unfold according to expectations; this unpredictability is not only intrinsic to innovation, it is also considered part of its *raison d'être*. Learning, adapting and taking advantage of unexpected opportunities are all part of a successful innovation process. Yet this can undermine the accountability function of evaluations: innovating teams can be tempted to shift the goals of the innovation to meet the reality of what they were able to achieve, rather than acknowledging the challenges they faced in delivering the original value proposition. There are different reasons for changing the original direction of an innovation. Some have to do with the natural learning process of innovation, others less so. These differences can be difficult to capture in an evaluation.

3. **Timing:** In addition to being unpredictable, innovation processes can often take years to reach their fullest impact, or even immediate outcomes. This can make it difficult to fully assess innovation activities before they are at least one year into their diffusion or scaling stage. Claims of longer-term impact play a particularly strong role in the argument for investing in humanitarian innovation. Yet impact measurement remains a problematic area for humanitarian actors (SOHS, 2015) and can be especially difficult for evaluations of innovation given the complex factors shaping successful uptake and use.

4. **The innovative personality:** Those who tend to excel as innovation leaders and thinkers may also be inclined to rebel against standard humanitarian practices around accountability, including evaluations (Sandvik, 2016). This reaction may be stoked by the current tendency to apply standard programme evaluation thinking to evaluating innovations. This points to the need for concrete ways of collaborating between evaluators and innovators and the identification of win-wins, such as the higher credibility and recognition of HI projects delivered through evaluation and the value of evaluative methods for improved innovation processes and learning (as outlined further below). Without clarity on these win-wins there can be challenges in getting innovation managers to use evaluations to learn about their process and inform future innovation activities.

While these challenges are considerable, it should also be noted that HI also presents important **opportunities** for HI evaluation and evaluative work:

1. **Focus on learning:** Innovation is essentially one overarching process of active learning, and adaptation in response to that learning. Many of the practices that are supportive of successful innovation are themselves

processes of evaluative inquiry. There is therefore already a tilt within good humanitarian innovation practice towards learning and seeking out the kind of analysis that evaluations can offer.

2. **Support for baseline measurement:** Innovation processes seek to bring about improvements to the way humanitarian assistance is currently structured and delivered. This aspiration for improvement necessarily invites comparative thinking: in order to know whether a new approach to water treatment in emergency settings is an improvement, we need to compare it with current practice. Innovation processes therefore draw immediate attention to the importance of baseline measurements, and innovation managers can be more likely than humanitarian programme leads to think about and collect baseline data at the outset of their project, for use in an evaluation. Indeed, innovation managers have been found to contribute to an understanding of how current humanitarian approaches and tools are performing through their pursuit of baseline data and protocols, thereby inadvertently contributing to the data needed to evaluate both standard humanitarian approaches and their own innovative approaches (Obrecht and Warner, 2016).

3. **The innovative personality:** There are also aspects of the innovative personality that are supportive for evaluation. Innovators tend to seek continual improvement, and therefore will be more likely to look for feedback to guide them. Evaluations can provide a key channel for this feedback. Innovators may also value evaluations as an advocacy tool to externally demonstrate that their innovation meets desired standards or criteria for quality and effectiveness.

3. Why evaluate? The purposes of evaluation for humanitarian innovation

All forms of evaluative inquiry and assessment are considered to serve two overarching **purposes** in humanitarian action: **learning** and **accountability** (EHA, 2016: Section 2.3). This is also the case for evaluations of HI. Evaluative inquiry and assessment, especially those focused on process, can support an innovating team or unit to reflect on their innovation practices; examine what went well, what decisions were taken at critical moments and why; and consider how they might do things differently (learning). Evaluative assessments can also be used to demonstrate that an innovation process has made good use out of its funding (accountability). As Section 2 mentioned, the core challenge for evaluators of innovations is to correctly recognise and incorporate, rather than ignore or suppress, the value of iteration when carrying out evaluation for the purpose of accountability.

Beyond learning and accountability, evaluations can serve a third purpose for humanitarian innovation: **uptake**. Evaluations focused on the innovation itself can support wider uptake of an innovation if they demonstrate to existing and potential new stakeholders that it offers a clear improvement in terms of quality or effectiveness over current practices. When used for this purpose, evaluations become a source of evidence that humanitarian innovators can then use in their diffusion activities to encourage other users to adopt the innovation. Evaluations focused on the innovation process that seek to develop and improve the innovation *process* as it unfolds can also support greater uptake by identifying and addressing weaknesses in how the innovating team engages with potential end users.

While uptake should never be the *only* aim for an evaluation, innovating teams and evaluators of innovations may want to consider this as a potential use at the outset of an evaluation. This would mean thinking about the

Box 2: Setting the scope for an evaluation of humanitarian innovation

HI can be engaged with at different levels:

- **Project-level innovation** relates to one-off innovation processes that are treated and managed as single projects. This remains the most common form of engagement by humanitarian actors in innovation. Many innovation projects are carried out as part of broader sectoral programmes, or as single information and communication technology (ICT)-driven projects that seek to capitalise on the opportunities new technologies offer. For example, evaluations of individual grantees of the HIF would be project-level evaluations.
- **Programme-level innovation** relates to programmes or organisational units that encompass multiple innovation processes. Increasingly, multi-project 'programmes' of innovation are run by 'hubs' or 'units' within organisations that work across different sectors or programme areas, overseeing or supporting multiple distinct innovation processes. These are still rare in the sector, but are growing in popularity. For example, an evaluation of the UN Refugee Agency (UNHCR) Innovation Unit would be a programme-level innovation.
- **Portfolio innovation** looks at multiple separate innovation projects and programmes. Donors interested in funding and supporting innovative practice will often look at innovation from a portfolio perspective. Doing this enables them to fund projects with a range of risk levels and assess them at a collective rather than an individual level, thereby mirroring the approach to investment in innovation in the private sector. For example, an evaluation of the UK Department for International Development (DFID) Innovation Programme, or of the HIF, would be a portfolio innovation.
- **Systems innovation** looks at innovation within a system comprising of institutions, organisations and their interactions. It is concerned with how each part of an innovation process – the identification, adjustment and diffusion of an idea for improvement – is enhanced or inhibited by rules, norms and incentives within a given arena. For example, the Centre for Research in Innovation Management (CENTRIM) research on evaluating different systems for innovation would be an example of a systems-level scope for assessing and understanding innovation practice (Ramalingam et al 2015).

Much of this paper will focus on how to evaluate innovation processes, outputs or impact at the individual **project** level, though many of the issues raised here will remain relevant to evaluations of innovation programmes, funding portfolios and 'systems'.

targeted end users of the innovation and seeing these parties also as potential end users for the evaluation. The purpose of the evaluation is to offer 'proof' of the value proposition for the innovation.

It is critical to understand who will be using evaluative inquiry and for what purpose. Table 1 lists the typical end users of HI evaluation, the scope they are typically interested in and the purpose for which evaluations are used.

In practice, many evaluations will serve two or all three of the purposes of learning, accountability and uptake. However, it is important to consider which of these purposes is most dominant, in order to inform the design of the evaluation (EHA, 2016: Section 2.3).

Table 1: End users of HI evaluation

Potential end users of HI evaluation	Tend to be interested in the following scope	Tend to have the following purposes for using evaluation
Donor or funder of an innovation project or programme	• Portfolio • Project • Programme • Possibly systems, if interested in shaping broader policy environment for innovation	• Accountability • Learning
Senior policy staff	• Systems • Portfolio • Programme	• Learning
Senior HQ management in an organisation	• Programme	• Accountability • Learning • Uptake
Innovation managers and teams	• Project • Programme	• Learning • Uptake
Technical or programme staff working in the sector relevant to the innovation	• Project	• Learning • Uptake

Box 3: Using evaluations to trigger innovation

In *More than just luck*, ALNAP found that innovation teams could generate the first spark for their innovation idea through evaluations of previous programmes or emergency responses. The use of evaluation findings and results in humanitarian action is a key driver for innovation, as it enables organisations to identify problems and opportunities for improvement. This can be especially powerful when evaluations involve the participation of affected people and reflect their views on humanitarian programming. Participatory evaluations with affected people can generate new ideas, which, if acted on quickly, can have a highly unique impact on humanitarian practice. See, for example the case study on International Federation of Red Cross and Red Crescent Societies (IFRC) Menstrual Hygiene Management Kits.

4. Applying evaluative practices to humanitarian innovation

There are different entry points for applying evaluative concepts and practices to HI. To understand these, it is helpful to consider four components: the three main **types of evaluative inquiry** and **how innovators use them**; three **perspectives** for the evaluation of innovation; and the four **focus areas of HI evaluation**. These are summarised in Figure 1; Sub-sections 4.1, 4.2 and 4.3 describe each level of Figure 1 in detail.

4.1. Types of evaluative inquiry and their uses by humanitarian innovators

There are three different types of evaluative inquiry that provide the main orientation for an evaluation. Familiar to evaluators, each of these types reflects a different use, summarised for innovators by the three concepts of 'develop', 'improve' and 'prove.'

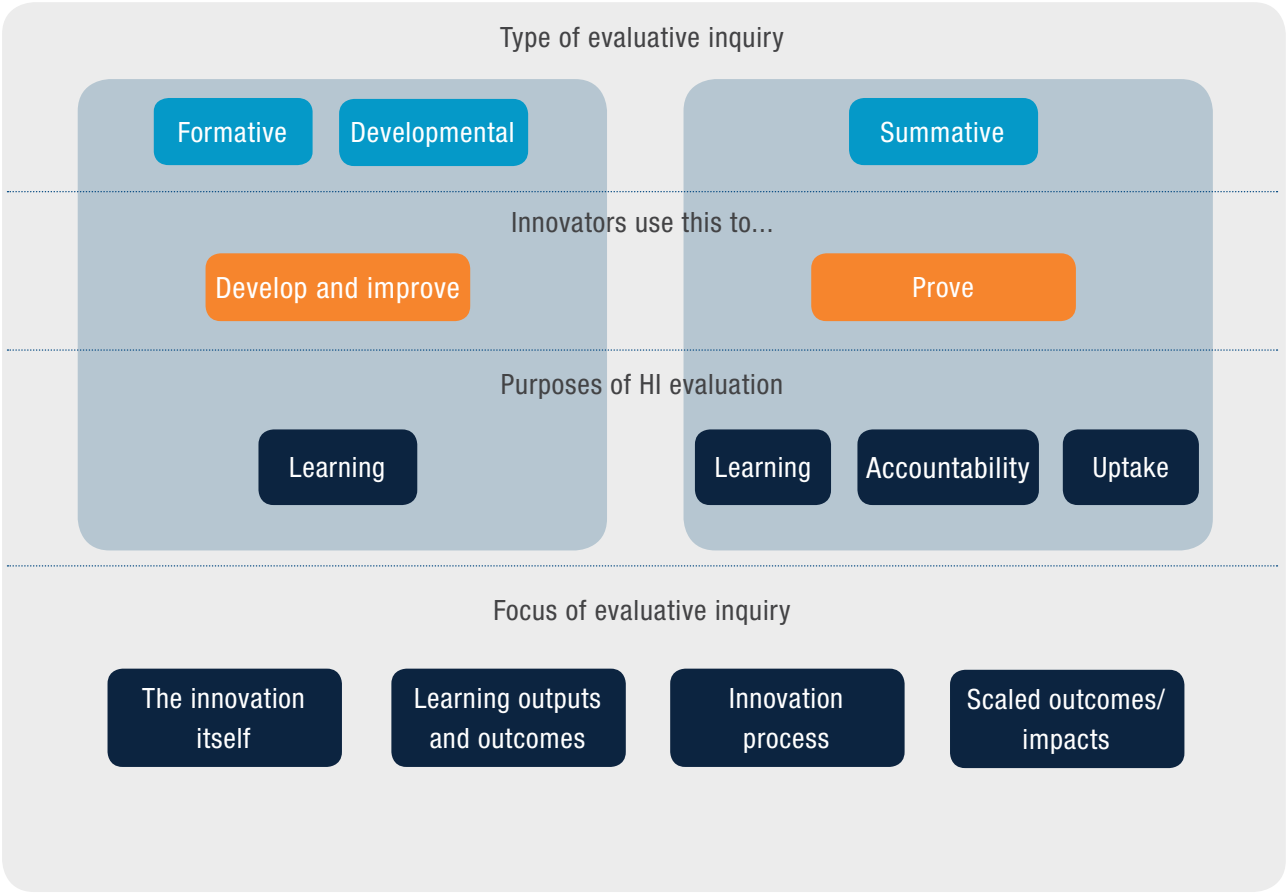
Formative and **developmental** evaluation can be used to develop and improve any of the four focus areas: the innovation itself; the consolidated learning and evidence; the innovation process; or the scaled outcomes and impact of an innovation. Formative evaluations assess a programme in order to identify areas for improvement and inform design (EHA, 2016: 63). Developmental evaluation works within a design framework to develop a programme or process based on judgements of how well it is doing. As suggested by Patton (2010), who introduced developmental evaluation, the primary difference between formative and developmental evaluation is that formative evaluation assumes there is a pre-existing model, and seeks to identify tweaks or improvements to this, whereas developmental evaluation maintains a broader perspective, making it possible to question the original assumptions and ideas behind a programme and change these where necessary. Developmental evaluation is therefore very similar to the actual design processes of ideation, development and implementation in an innovation process, and can easily be incorporated into the innovation activities in order to strengthen an innovation team's learning and agility. Formative evaluation focused on identifying improvements to the innovation itself may be useful later in the innovation process, when an innovating team is finalising a prototype.

When applied to the innovation *process* as the focus, formative and developmental evaluation may be used to identify strengths and weaknesses in an innovation unit or hub, or in an organisation's broader approach to innovation; it can also be used within an individual project to assess how well the innovation process is doing with respect to certain success factors, such as engagement of end users, internal learning processes or monitoring of risk.

When applied to *scaled impact* as the focus of evaluation, formative and developmental evaluation will seek to shape the scaling or diffusion activities of an innovation by regularly assessing the impact of an innovation and seeking adjustments in order to calibrate a desired path for impact. For example, if positive impact achieved with one end user group is threatening positive impact with a different end user group, an innovating team may seek adjustments to its diffusion plan in order to achieve a more even and comprehensive impact across multiple end user groups.

Summative evaluations judge the merit or worth of a programme at its conclusion (EHA, 2016: 63) and can be used as evidence to 'prove' (or disprove) the value of an innovation or innovation process or the scaled impact of a widely diffused innovation. Innovating teams can use summative evaluations of the innovation to demonstrate the proven value of their prototype for the purpose of *uptake*, while evaluations of the innovation process can be used to *learn* about what went well and what contributes to successful innovation. Donors can use summative evaluations to hold an innovating team accountable by examining the proof for their value proposition.

Figure 1: The purposes of humanitarian evaluation practice



Entry points for evaluation and humanitarian innovation: Evaluation has two main entry points in HI. The first is concerned with how innovators can actually use evaluative practice and evaluative thinking in *their own work*, to help design and improve their innovation or their process. This is the realm of formative and developmental evaluation. Box 4 (page 14) describes some of the key benefits of developmental evaluation in particular for innovators. While most of this paper concerns summative evaluation, Annex I provides further introduction to developmental evaluation for innovation managers and outlines the potential benefits for innovation managers in working with a developmental evaluator throughout their innovation process.

The second, and perhaps more familiar, entry point for evaluation in HI is through summative evaluation, which entails evaluating an innovation process, its outputs or its impact to assess and prove their worth. Summative evaluation comprises most humanitarian evaluation activity and is typically what people refer to when thinking about evaluative practice in humanitarian action. While innovators can also use summative evaluation to learn about their process and their innovation, or to drive uptake, one of the main uses for summative evaluation is accountability for how resources were used. Here, a primary question is how to adapt current approaches to summative evaluation in humanitarian action in light of the distinguishing characteristics of innovation as described above in Section 2. This question will be the focus of Sections 5–7.

Box 4: Considerations when using developmental evaluation in humanitarian innovation

Key considerations for evaluators doing developmental evaluation for innovators:

- Rather than focus on a theory of change, consider the different inquiry frameworks that can be used in developmental evaluation and review with the innovating team which frameworks might be most appropriate for their situation. Inquiry frameworks are sets of questions that help critically analyse and understand the situation in which innovators are acting, in order to inform thinking and decision-making (Patton 2010: 100).
- Evaluation designs will need to change as the innovation progresses and you will need to adapt data collection methods and focus as the programme adapts and changes. How can innovators help make this happen? How will you plan for continual evaluation framework reviews?
- Unintended consequences are as important as intended ones. So evaluation needs to properly integrate a search for these; consider how you will approach open-ended field inquiry.
- As the design of the prototype progresses, keep an eye out for when it may be appropriate to decide on a finalised set of design criteria and transition to formative evaluation to hone and improve the prototype for piloting and diffusion.

Key points for innovators to address when working with evaluators:

- Work alongside an evaluator during the innovation process and ensure adequate resources to support this. Innovating teams should consider budgeting for such support when applying for funding to support an innovation process.
- Set up a clear communication and management structure to ensure the evaluator is kept up to speed with changes in the project during the innovation process.
- Remember developmental evaluation is primarily there to serve your design needs; if it is not working for you, it should be adapted or changed.

References for developmental evaluation:

- Patton, Michael Quinn (2010) *Developmental evaluation: Applying complexity concepts to enhance innovation and use*.
- Better Evaluation: Developmental evaluation toolkit: **http://betterevaluation.org/resources/toolkit/developmental_evaluation_toolkit**

With contributions from Neil Dillon, Research Fellow for Evaluation, ALNAP

4.2. The perspective of an evaluation

The 'perspective of an evaluation' in humanitarian innovation refers to the point at which the evaluation takes place in the innovation process. Summative evaluations tend to take place at the 'end' of a project. Yet innovation processes are best understood as ongoing and open-ended, often defying a clearly defined end date. However, evaluators should be aware that there are three moments in an innovation process where evaluations tend to be commissioned; each of these implies a different perspective for the four focus areas of an innovation.

While this rarely happens in practice, evaluators could be commissioned at the outset of an innovation process in order to work alongside the team to help develop and improve the innovation. Such evaluative practices take a **design perspective**, entering the innovation process at the early stages where ideas are generated and worked out and different designs are trialled. This involves formative and developmental forms of evaluative inquiry, which, as described in greater detail in Section 4.1 and Annex I, can offer great value to innovation managers.

Many evaluations of humanitarian innovations are commissioned at the end of a pilot phase, where a working prototype has been implemented and tested. These evaluations run the risk of treating the innovation project as a standard project. Evaluations conducted at this stage offer a **pilot perspective** on the innovation process, the learning generated by the innovating team and the outcomes achieved.

Many innovation activities in the humanitarian sector end with a pilot, leading to what has been described as 'pilot-itis' (McClure and Gray 2014: 4): the proliferation of pilots with insufficient follow-up to see these innovations adopted widely in the sector. Increasingly, donors and innovation practitioners are investing more resources in diffusing and scaling proven innovations in order to see them achieve positive outcomes at a wider scale. Evaluators may therefore be commissioned to look at how innovations are being scaled, and what value these innovations offer once broadly diffused. This is the **scaling perspective** onto innovation.

4.3. The focus areas of HI evaluative practice

Evaluative inquiry in HI is also shaped significantly by the **focus** of evaluation, as outlined in Table 2.

There are many issues that may be of interest to end users of an evaluation of HI. As with all evaluations, the particular focus and evaluative questions will always be based on the particular project, and will need to be defined and negotiated based on the intended purpose and end users of the evaluation. This section is not intended as an exhaustive list of the possible activities, outputs or outcomes of interest for HI evaluation (for more on how the results chain intersects with the focus areas of HI evaluation, see Box 5 below). Rather, it describes the common focus areas for evaluating HI, asking: If we want to know the value of an innovation project or programme, what might we be interested in looking at?

In the HIF-ALNAP case studies on innovation, the research team needed to understand what successful innovation meant, in order to be able to explore what factors support or inhibit success. Based on a review of the literature, ALNAP identified three success criteria for humanitarian innovation:

1. **Adoption:** The innovation is taken to scale and used by others to improve humanitarian performance.
2. **Improved solution:** The innovation offers a measurable, comparative improvement in effectiveness, quality or efficiency over current approaches to the problem addressed by the innovation.
3. **Consolidated learning and evidence:** New knowledge generated or the evidence base enhanced around the area the innovation is intended to address or performance of the innovation itself.

Based on these success criteria, we suggest there are three potential focus areas for an evaluation of HI: the innovation itself (which can be assessed to understand whether it offers a comparative improvement); the information and evidence produced by an innovation process (which can be assessed to understand whether it offers consolidated learning and evidence to the sector); and scaled outcomes/impact (which can be looked at to understand whether a broader improvement was brought about to the sector through wide adoption). In addition to these three, we also add the innovation process as a focus area, as end users of HI evaluations may be keen to learn about the innovation inputs and activities and how well these worked.

Focus area 1: The innovation itself

For some, evaluating innovation is about evaluating the outputs and outcomes of the pilot. One of the main outputs of interest is the **innovation itself** that has been, or is being, developed, tested, implemented.

Innovations can be 'things' (products), new ways of doing things (processes), new ways of framing how things are done (positions) or radical changes to how humanitarians think, relate and work (paradigms). One of the key areas of interest for end users of an HI evaluation is: How good is this product, process, position or paradigm? For example, users of HI evaluations may want to assess the newly developed piece of software, a new approach to disaster risk reduction programming, an alternative way of conceptualising and measuring malnutrition programming, or a kit for supporting menstrual hygiene for women and girls in crisis.

In carrying out developmental evaluations to support the further refinement of a prototype, evaluators might work closely with innovating teams to explore what intended end users and gatekeepers think about the prototype and where they might want to see adjustments. When doing summative evaluations of the innovation at the pilot stage, evaluators will need to assess quality, by looking at quality indicators either at the output level (i.e. looking at how the innovation performs against relevant standards or design protocol) or at the outcome level (i.e. the quality of the outcomes it produces). In both cases, it is recommended that evaluators use an evaluative criterion of Comparative Improvement, described in Section 7.

Focus area 2: Consolidated learning and evidence

Some pilots will result in 'failed' prototypes: innovations that do not offer an improvement over current practices, or that do not appear to be effective at addressing the main problem identified by the innovation process. Users of HI evaluation may therefore be interested in evaluating the learning outputs produced by the innovating team, and the learning outcomes achieved through these outputs. These can be a useful way of demonstrating the value of an innovation process when the original idea for innovation has not proven to work. Evaluators may ask: What were the learning outcomes around disaster risk reduction programming or menstrual hygiene management produced through the development and testing of this innovation? It is critical to look at the learning outcomes of every innovation process, as this can mark the difference between a 'bad fail' (a poorly managed process) and a 'good fail' (an innovation process that was wrapped up because the team had strong learning processes in place to recognise progress was not being made, and where learning from the project was shared externally to support future innovation efforts).

Focus area 3: Process

Other users of HI evaluations may be interested in evaluating the processes taken to develop these tools, interventions and ways of working, perhaps in order to understand how learning was captured or how efficiently the innovating team used its resources. Here, the focus is on **evaluating the innovation process**

itself. Such learning can be used to make changes to the innovation process as it is underway, inform future innovation processes or improve innovation practice more generally within the organisation.

Monitoring data will be very important for a process evaluation of humanitarian innovation. The ALNAP Working Paper on *Monitoring humanitarian innovation* (Warner 2017) identifies a set of milestones – standard outputs that every innovation process produces – and indicators of quality that can be used to monitor the progress of an innovation process. Evaluators should refer to this paper and consider these outputs when carrying out a summative evaluation of an innovation process.

Focus area 4: Scaled outcomes/impact

A final focus area for HI evaluation, but one not currently practised widely in the sector, is **the scaled outcomes or impact of an innovation's outputs**. This refers to the broader effects of an innovation or its consolidated learning once a wide range of end users have taken it up. Evaluating scaled impact of an innovation is distinct from assessing the scalability of an innovation. Increasingly, donors of humanitarian innovation are interested in supporting the diffusion of high-performing prototypes or new practices so that they achieve a wider impact on the sector. This points to the need to identify good 'scalable' prototypes: prototypes that have a likelihood of success if they are taken to scale. Such assessments of an innovation's scalability are not strictly evaluations, yet evaluators may be asked to participate in or advise these. Annex III offers some considerations on scalability to support evaluators in these situations.

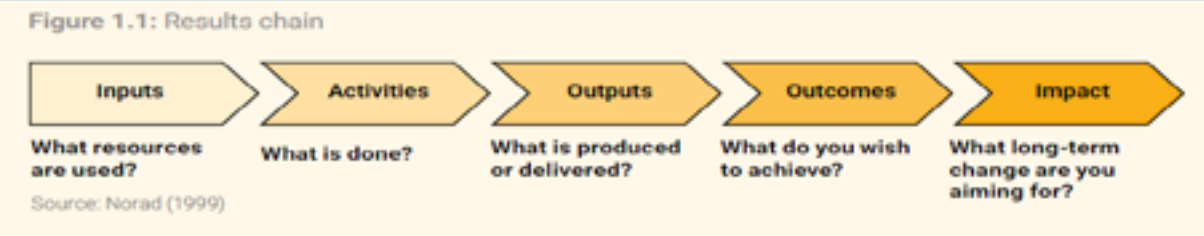
Table 2: Four potential focus areas for an HI evaluation

<table>
<tr><th colspan="2">1. The innovation itself</th></tr>
<tr><td>Product innovation
Changes in the things (products/services) an organisation offers</td><td>Process innovation
Changes in the ways products and services are created or delivered</td></tr>
<tr><td>Position innovation
Changes in the context in which the products/services are framed and communicated</td><td>Paradigm innovation
Changes in the underlying mental models that shape what the organisation does</td></tr>
<tr><td colspan="2">Source: Tidd and Bessant (2005); Obrecht and Warner (2016: 15)</td></tr>
<tr><th colspan="2">2. Consolidated learning and evidence generated by the innovating team</th></tr>
<tr><td colspan="2">• Policy briefs or learning papers publically shared
• Blogs
• Focused and targeted contributions to relevant consortia and learning platforms or networks
• Case studies publically shared</td></tr>
<tr><th colspan="2">3. The innovation process</th></tr>
<tr><td colspan="2">The process by which an idea for innovation was conceived, developed, tested, and diffused</td></tr>
<tr><th colspan="2">4. Scaled outcomes/impact</th></tr>
<tr><td colspan="2">The broader effects of an innovation or its consolidated learning achieved through the diffusion and uptake of the innovation or this learning by many actors.</td></tr>
</table>

Box 5: The results chain and its relationship to humanitarian innovation

Some evaluators may seek to apply the standard results chain to humanitarian innovation processes.

Figure 1. Results chain



Source: Norad (1999), in EHA (2016: 28).

Application of the results chain to an innovation project will depend on the focus and perspective of the evaluation. When applying these categories to an innovation, it is important to remember the following key points:

- In a typical project, outputs and outcomes are fixed, as the relationship between the two (outputs produce outcomes) is expected to be grounded in previous experience or evidence. In an innovation project, there may be some existing evidence to suggest a relationship between outputs and outcomes. However, the outputs can change, and, importantly, so too can the desired outcomes, as an innovating team learns more about the problem it is trying to address.
- Results chains can be used to categorise the different areas of interest in an innovation process into activities, outputs, outcomes and impacts, but the relationship between these is unlikely to be linear or clearly defined.
- Timing matters: be clear on what perspective your evaluation is adopting (design, pilot, scaling), as this will affect what you identify as outputs, outcomes and impact.
- Innovators may be using a different language to talk about the results chain, but these problems are not insurmountable. Particularly when it comes to outputs, outcomes and impact, innovators can and should be able to explain what these are in the context of their innovation. Understanding the innovation itself (output), the benefits it is expected to offer (outcomes) and the broader change it will bring about, especially if widely adopted (impact), is a critical part of forming a strong value proposition.

5. Selecting humanitarian innovation evaluation questions for summative evaluation

Selecting evaluative questions is one of the most important steps in an evaluation. In order to ensure utilisation, it is recommended that evaluators develop the questions based on the knowledge needs of primary intended users,[1] so detailed discussion with the users of an evaluation and their purposes for the evaluation (described above) is important (ALNAP, 2016).

Once users' needs are clear, the evaluation questions can be drafted. Based on these needs, the evaluator should clarify the focus area for the evaluation. While questions will vary depending on the sector and level at which the innovation practice is being evaluated, Table 3 provides some examples of high-level evaluative questions for innovation for each focus area. The evaluator should refer to the specific aspects of the innovation in order to develop more concrete questions for the evaluation.[2]

The evaluation team should break down high-level evaluation questions into more focused questions. As such, to avoid a 'laundry list' of evaluation questions, it is suggested to limit the number of high-level evaluation questions to two or three. Any more than this will reduce how much time and evidence can be gathered to answer each question, thus possibly compromising the quality of the evaluation.

Since innovation is still a relatively new area of formal practice for the humanitarian sector, there may be a tendency to attempt to apply standard evaluative questions to innovation processes. Evaluators must be aware of this and should aim to engage evaluation commissioners and intended end users in a discussion over how they see innovation as distinct from standard programming, and what this means for the evaluation of such activities.

Table 3: Sample evaluative questions for each focus area of HI evaluation

Focus of evaluation	Sample evaluative questions
The innovation itself	• Does the innovation offer an improvement in quality, efficiency or effectiveness compared to current approaches? • Did the innovation respond to a clear need?
Consolidated learning and evidence	• What is the quality (accessibility, relevance) of the learning outcomes?
The innovation process	• What factors shaped the quality of this innovation process? • How efficient was the innovation process? • How effective was the innovation process in terms of achieving the objectives of learning, comparative improvement and wide adoption? • How successful was the innovating team at learning and adapting? • How useful was the structure and management of the innovation process? • Was the pilot implemented effectively?
Scaled outcomes/ impact	• Has there been successful uptake of the innovation and what are the factors that have supported or impeded this? • What have been the effects of the wider uptake of the innovation? • Has there been successful uptake of the consolidated learning and evidence from the innovation process and what are the factors that have supported or impeded this? • What have been the effects of the uptake of learning and evidence from the innovation process? • What have been the results – intended or unintended, positive and negative – of the diffusion activities undertaken by the innovating team? • Has there been a measurable improvement in system performance in the area addressed by the innovation?

6. From 'intervention logic' to 'innovation logic': Working with theory of change in an humanitarian innovation summative evaluation

A central challenge in carrying out a summative evaluation of HI lies in identifying an appropriate 'fixed point'. In standard EHA practice, this fixed point is provided by an 'intervention logic' – typically represented in a logframe – which outlines the intended objectives of a project or programme. Evaluators use the intervention logic to explore how what happened in the programme measures up against the desired aims.

As Section 2 described, innovation processes differ significantly from standard programming in their intervention logic, or theory of change. A higher degree of novelty in an innovation process leads to a higher degree of uncertainty as to its results and how these will come about. The function of a theory of change in an innovation process is to identify a loose set of issues believed to be important for addressing the core problem or performance area. Innovation then proceeds more as a process of **discovery** than one of implementation, in which the innovating team collects further information in order to iteratively develop the solution and learn more about the problem or performance area as they progress. This can be an important part of monitoring an innovation process (see Warner, 2016) as it allows innovating teams to 'fail small' and 'fail early', making necessary adjustments or course corrections early on, or supports a decision to call an end to a design process and walk away before over-investing resources or taking on additional risk.

Evaluators must use methods and evaluation designs that allow them to acknowledge this iteration while also supporting accountability. Changes in the design of the innovation could be legitimate, based on discoveries about the problem being addressed by the innovation, or they could be driven by an innovating team feeling under pressure to cut corners and downsize ambitions to meet deadlines. To illustrate, take two examples, one of an opportunity-driven innovation (innovation that seizes an opportunity for improved performance); and one of a problem-driven innovation (innovation that seeks to solve a specific problem).

Problem-driven innovation: Example A: The Shelter Pod

A humanitarian agency is grappling with the challenge of serving large populations of internally displaced persons (IDPs) and refugees outside of camp settings, over longer periods of time than is typical for them. This is creating higher costs, particularly in the provision of emergency shelter, and there are many reported problems with the shelter units coming from the IDPs and refugees who live in them. An innovating team within the agency receives £250,000 to design a new type of emergency shelter that is more cost-efficient to build, and provides more comfortable shelter over a longer duration. The key to achieving these criteria is the use of a 'pod' design in which the shelter is shaped like a large egg. At the close of the project, however, the shelter pod does not have a 'pod' design and instead looks much more like standard transitional shelter, though it has a few interior design features that are different from current forms of transitional shelter. For short-term use, it is more costly than other designs because of the reliance on more expensive materials.

Opportunity-driven innovation: Example B: The Humanitarian Bluetooth Network

An innovating team receives £150,000 to carry out early development of a Bluetooth network to enable highly secure communications with affected people. The network operates via a set of drones that provide Bluetooth wiring across mobile phones in an African sub-region with high mobile use. This is expected to help improve communications between humanitarian agencies and affected people, providing agencies with information for needs assessments as well as feedback on quality of the aid provided. At the close of the project, the Bluetooth network has been set up and piloted in one area of the Central African Republic. The humanitarian agency that participated in the pilot has not been able to use the system to improve its needs assessments or gather feedback on its operations. However, there are many reports that affected people are using the network's secure lines to communicate with each other and the Bluetooth network is a highly popular alternative to the local mobile network provider.

In both examples, an evaluator is faced with innovations that have shifted from their original expectations. There is a key question in both cases of whether or not the innovation addressed the original problem or opportunity area that was identified. However, simply assessing both projects at the end of their funding period against the original plan would be problematic, for two reasons. First, it undermines the very purpose of innovation as a more open-ended approach to humanitarian action that seeks new ideas or improvements rather than repetitions of known models. Second, evaluating against the original 'plan' (to the degree that there was one) misses the opportunity to understand what was learnt about the problem or the performance area that these innovations were targeting.

Yet, at the same time, innovation cannot be entirely open-ended: if the shelter pod group had instead set up a troupe of clowns to entertain IDPs and refugees, or if the Humanitarian Bluetooth Network had instead delivered empowerment workshops, this would seem to fall 'out of scope' of the original value proposition. How can donors, and evaluators, determine what lies 'within scope' and 'out of scope' in an innovation process?

The answer to this question lies in understanding the 'innovation logic' as opposed to the 'intervention logic' or logframe. The **innovation logic** is the pathway of decisions and discoveries that occurs throughout the lifespan of the innovation process. This logic unfolds and shifts as the innovation process proceeds and can be captured in full only at the end. To grasp the innovation logic, evaluators will need to interrogate the *reasons* behind the key decisions that were taken to adapt the project. Evaluating innovation therefore places a much greater emphasis on understanding the *decision-making and learning processes* that have taken place in the lifespan of the project. The primary task of an evaluator in an HI evaluation is to 'think along' with the innovating team in order to describe and understand the innovation logic.

Here, we suggest three approaches evaluators can use to engage with the innovation logic. **Evaluators do not need to choose among these approaches**: all of them can be used in combination, or as alternatives, to support the evaluation of an HI. The first two approaches apply mainly to evaluating the innovation itself (Focus area 1); the third approach can be used to evaluate the innovation process (Focus area 3).

This is not tested guidance, and should be read as suggestions to help frame how an evaluator approaches and thinks about their HI evaluation.

6.1. Preliminary step: Use the value proposition to create an evaluation matrix

For all three approaches outlined below, it will be essential to have an evaluation matrix based on the value proposition of the innovation. Value propositions can change over time. Initially, it is useful to start with the most current statement of the value proposition – typically the one provided at the end of the innovation's funding period. The value proposition should be provided or developed in consultation with by the innovating team, rather than drawn from original project proposals (as these would reflect an earlier version of the value proposition, before the team embarked on its learning process in the innovation design and development). The value proposition lays out the stated problem, or area of opportunity, that the innovation is addressing, and describes the solution or value it offers.

The innovating team's understanding of the problem or opportunity context directly informs its selection of design criteria and objectives for its innovation. The understanding of the problem or opportunity context is the most important column in the matrix, as it will be used to anchor the mapping of the innovation logic. Evaluators should seek to develop a list of the design or performance criteria implied by the value proposition for the innovation, and work with the innovating team to understand how these performance criteria are ranked in order of importance.

Table 4: Template for an HI evaluation matrix

Characteristics of the problem or opportunity for improvement	Sources of verification (for the characteristics of the problem or opportunity)	Design/ performance criteria	Rank	Indicators	Comparative practice/tools	Methods	Sources of verification (For the assessment of the innovation and its comparative practice)

In order to assist the evaluator in understanding the value proposition, innovating teams may try to share early proposals or project documents to inform the evaluation matrix. If an innovating team claims that little in its value proposition has changed from the original proposal documents, there should be strong evidence that the team did a thorough job of understanding user needs at the outset of the innovation process. If this is not the case, and the idea for the innovation comes primarily from within the innovating team, the lack of any major change from the original proposal **may be a red flag that the innovation process was not sufficiently adaptive and did not produce an innovation that responds to end user needs.**

To illustrate how the value proposition can shift from the beginning to the end of the project, consider the example case of **the shelter pod.** Suppose that the innovating team made two discoveries during the early development of its pod-shaped shelter: first, that IDPs and refugees found the egg-shaped design of the shelter silly and insulting to live in; second, that IDPs and refugees had their own coping mechanisms for regulating temperatures within the shelter structures and there were other design features more relevant to their comfort, as identified through focus group discussions. Based on this, the team decided to eliminate the pod design, include the interior design features suggested by affected people, and focus on ways to extend the durability of the shelter in order to increase cost-effectiveness over time.

Value proposition at the **start** of the funding period:

> **'Humanitarian agencies in region X are having to serve IDPs and refugees outside of camp settings for much longer periods of time. Current forms of emergency shelter make this longer-term service provision more costly and also less comfortable for the IDPs and refugees. The shelter pod** *uses higher-quality building materials more cost-effectively through its pod design and also provides better cooling and heating, allowing it to be used year round and for longer durations than current emergency shelter.'*

Value proposition at the **end** of the funding period:

> **'Humanitarian agencies in region X are having to serve IDPs and refugees outside of camp settings for much longer periods of time. Current forms of emergency shelter make this longer-term service provision more costly and also less comfortable for the IDPs and refugees. The shelter pod** *uses higher-quality building materials and includes interior design features that fit the expressed needs of IDPs and refugees, allowing it to be used for longer durations than current emergency shelter'* **[change in value proposition in italics]**

Initially focusing on the value proposition at the end of the funding period, an evaluator of the Shelter Pod might fill in the evaluation matrix as follows:

Table 5: Example of a filled in evaluation matrix for the Shelter Pod case

Characteristics of the problem or opportunity for improvement	Sources used for verification	Design/ performance criteria	Rank
Current emergency shelter is not cost-effective for longer durations	Agency evaluation of response 2015	Durability; Cost-efficiency of materials	2
Current emergency shelter is not comfortable due to interior design issues.	Focus group discussions held during design process	Comfort/ Appropriateness	1

6.2. Approach 1: Verify the value proposition.

An initial approach to summative evaluation of an innovation itself is to carry out a basic assessment of whether the innovating team's claims about the innovation are true. This is called verifying the value proposition. Since innovations are intended to offer an improvement over current practice, evaluators can fully **verify the value proposition** only by comparing the innovation's performance against a baseline of a relevant current approach or standard practice. In order to support a comparison across current practices (see below, Section 7, under evaluative criterion Comparative Improvement), evaluators should identify relevant existing approaches to the problem and collect measures on the same indicators for these. At the outset of an evaluation of an innovation prototype, evaluators should work with the innovating team as well as interview independent experts outside the innovating team to identify a relevant comparator. Table 6 provides the additional columns for the evaluation matrix to undertake a verification of the value proposition.

Table 6: Evaluation matrix with added columns for verifying the value proposition

Characteristics of the problem or opportunity for improvement	Sources used for verification	Design/ performance criteria	Rank	Indicators	Methods and sources for assessing performance and comparison	Comparative practice/ tools	
						Tents	T-shelter
Current emergency shelter is not cost-effective for longer durations	Agency evaluation of response 2015	Durability; Cost-efficiency of materials	2	Cost of shelter materials over average duration of usability	Cost-Benefits Analysis; Procurement documents, project monitoring documents	Equal to innovation	Worse than innovation
Current emergency shelter is not comfortable due to interior design issues	Focus group discussions held during design process	Comfort/ Appropriateness	1	Average ratings of satisfaction with shelter	Focus group discussion, survey	Worse than innovation	Equal to innovation

6.3. Approach 2: Assess the relevance of the problem or opportunity.

A second approach to evaluating an innovation output is assessing whether the stated problem or opportunity area is relevant to anyone outside the innovating team. To do this, evaluators should work with each of the stated characteristics of the problem or opportunity, asking, 'For whom might this be a problem or opportunity?', and developing a set of methods to assess whether this problem or opportunity is indeed recognised by the stated end users and primary beneficiaries of the innovation. For example, in the case of the Humanitarian Bluetooth Network, was the lack of communication between agencies and affected people recognised by affected people themselves? Or was it recognised only by humanitarian agencies? Using this approach in an evaluation can bring the focus back to those meant to benefit from an innovation, whether they are affected people or humanitarian aid workers or policy-makers.

It should be noted that the end users of an innovation are not always affected people: a software-based innovation may target humanitarian field staff as end users and seek to achieve performance efficiencies that provide overall improvement to a humanitarian response. In these cases, evaluators should note that affected people are not the targeted end user and ensure that the evaluation acknowledges and includes the appropriate stakeholders – namely, end users and gatekeepers (Box 6 on "Box 6: Three types of stakeholder in HI" on page 27).

Table 7: Evaluation matrix with added columns for evaluating for whom the value proposition is relevant

Characteristics of the problem or opportunity for improvement	Sources used for verification	Design/ performance criteria	Rank	Who is this a problem/ opportunity for?	Methods and sources for assessing that this is a relevant problem/ opportunity
Current emergency shelter is not cost-effective for longer durations	Agency evaluation of response 2015	Durability; Cost-efficiency of materials	2	Humanitarian agencies	Project evaluations, key informant interviews with field staff
Current emergency shelter is not comfortable due to interior design issues.	Focus group discussions held during design process	Comfort/ Appropriateness	1	IDPs and refugees	Focus group discussion, surveys, interviews with field staff

6.4. Approach 3: Reconstruct the innovation logic

In order to reconstruct the innovation logic, evaluators will need to revisit original proposal documents or interviews with early staff in order to understand what was originally planned at the outset of the innovation process. This, along with the evaluation matrix depicting the value proposition at the end of the innovation process, will enable the evaluator to track the shifts and changes in the innovation. Many of these will have come out of difficult choices in weighing design priorities or addressing different aspects of the overall problem. The point here is to focus on the larger 'course corrections' and decision points rather than on mapping each detailed meeting or small decision.

It is also useful for the evaluator to consider what the innovation team used as sources for new information to reprioritise design or performance criteria. Did the team organise consultations with potential end-users? Did they present their idea at a conference? Did they use field-testing? Identifying what information was used can also prove valuable in assessing the innovation team's adaptability and learning processes.

The most challenging and most important aspect of reconstructing the innovation logic lies in identifying the **types of reason** given for different decisions made by innovators. In many cases, the rationalisation of the innovation process may not map fully onto reality: a shift in the innovation may have owed to many small shifts in the mood of an innovating team over time rather than to a formal decision. However, rationalising this process to some degree is important as it helps make the implicit explicit: it brings to light the underlying prioritisations and sometimes unconscious decisions made in the process of an innovation; this helps the innovating team better understand their journey and assess the value of the innovation process or the innovation that it has produced. Quality or Evaluation criteria (Section 7) can be used to assess the strength of these reasons and identify where the key decisions might have moved the innovation further towards, or away from, these criteria.

Table 8 (on page 30) presents a matrix for reconstructing and applying evaluative or other quality criteria to an innovation logic. The right five columns can be replicated multiple instances for each decision point. Alternative visualisations, such as decision modelling, can also be used.

Box 6: Three types of stakeholder in HI

Primary beneficiaries are those who benefit directly from an innovation. This is not the same as affected people; in many cases, humanitarian staff are the primary beneficiaries of an innovation. Primary beneficiaries can be identified by answering the following question, 'If the innovation works, who would see the most obvious and immediate benefit?'

End users are those who interface directly with the innovation (and whose behaviour must change in order for the innovation to deliver its value). End users are those who must 'use' the innovation in order for it to work. They are not always the primary beneficiaries of an innovation. For example, the Humanitarian eXchange Language – which aimed to resolve the lack of a common operational picture of humanitarian crises – involved at least two main types of actor: information management officers (IMOs) and data entry specialists. While IMOs are the primary beneficiaries for addressing this problem (they are the primary users of a common operational picture), the innovation required behaviour changes from data entry specialists who would be the end users of the new technology. As end users, and not beneficiaries, therefore they were not initially incentivised to support the innovation. End users can be identified by answering the question, 'Who needs to interact directly with the innovation in order for it to work?'

Gatekeepers are those who can significantly influence uptake because of their control over the behaviours of primary beneficiaries and end users. The humanitarian system is not a free market: gatekeepers are the actors whose choices construct the environment of services, products and paradigms. For example, international NGOs (INGOs) are often gatekeepers for innovations in which the end users are affected people. Donors and governments can be gatekeepers for innovations in which the end users are INGO staff. Gatekeepers can be identified by answering, 'Who determines the range of choice for the innovation's end users and primary beneficiaries?' **(Obrecht and Warner, 2016)**.

Many evaluations of HI may find very few problems or opportunities arise from the perspectives of affected people themselves. This is part of a broader problem in humanitarian action, whereby affected people are afforded very few opportunities for real decision-making or influence in humanitarian assistance. Therefore, this approach can be applied to affected people in particular, in order to encourage the humanitarian sector and humanitarian innovators to consider the views of affected people in problem definition and in the search and identification of opportunities for innovation.

As outlined in ***More than just Luck*****.**

In the case of the shelter pod, the original design feature that made the innovation seem 'innovative' was removed from the final prototype. The evaluation matrix enables the evaluator to map out the initial understanding of the problem and how the design of the shelter pod shifted in response to changes in the team's understanding of the problem as experienced by IDPs and refugees. In this case, innovators learned about what was actually relevant to affected people for their comfort in a shelter unit and shifted their design criteria in light of this learning. This shows how the team was responsive to newly identified needs and also provides some learning to the organisation on its own internal knowledge management practices. For example, the source for establishing the original design/performance criteria was an agency evaluation in 2015: this might indicate

to the innovating organisation that their evaluations have not been picking up the views and preferences of affected people in a way that allows them to accurately inform the design of future projects.

In an innovation process the design criteria, or original aims, should be adaptable. Yet evaluators need an anchor to understand whether these criteria have been shifted in a positive or negative way. The 'anchor' for these changes should be 1) the understanding of the problem or performance area to which an innovation is responding and 2) the quality or evaluative criteria applied to the innovation. Innovation designs can change dramatically, so long as they are always meeting a relevant problem or opportunity area, and are doing this in a way that reflects the desired quality criteria. This is what allows for evaluations of HI to be sufficiently open-ended without losing the fixed assessment points necessary for an accountable evaluation.

Working out these fixed assessment points will always be open to interpretation and will need to be discussed for each evaluation. Returning to the example of the Humanitarian Bluetooth Network, the evaluator has found that the network is very popular among affected people but is not being used for its original intended purpose of enhancing communications with humanitarian agencies. Is this a problem? It depends on the fixed assessment points agreed with those commissioning the evaluation. The donor or innovating team may have set out with the goal of improving accountability to affected people, or improving the accuracy and timeliness of needs assessments. Instead, it has ended up with an innovation that scores highly on the quality criterion of relevance (as it is highly desired by affected people) and may serve protection goals (if it is used by affected people to share information that increases their safety). The wider the range of quality criteria considered acceptable for an innovation to meet, the more open-ended the innovation can be. A continuing question for those who fund and engage in innovation is the degree to which they are willing to let go of predetermined goals, objectives and values, and instead identify a range of different values or criteria that an innovation could potentially meet in order to be deemed of worth. These are not discussions that evaluators should be expected to have an answer for, but evaluators will need to be prepared to engage in these discussions should they arise.

Table 8: Evaluation matrix with added columns for mapping the innovation logic – the series of revisions, decisions and reasons that mark the changes of an innovation

Original characteristics of the problem or opportunity for improvement	Design/ performance criteria	Source	Rank	Final characteristic of problem or opportunity	Sources used for verification	Revised design/ performance criteria	Rank	Reason	Evaluative/ quality criterion which this reason reflects
Current emergency shelter is not cost-effective for longer durations	Durability; Cost-efficiency of materials	Agency evaluation of response 2015	2	**Unrevised**	Agency evaluation of response 2015	Durability	2	Materials less cost-efficient with revised non-'Pod' design, but the inclusion of new interior design features may lead to longer term use, and thereby increase cost-effectiveness over time	Efficiency
Current emergency shelter is not comfortable for longer durations because of heating/ cooling issues	Heating/ Cooling capacity	Agency evaluation of response 2015	1	**Removed: shown to be false**	Focus group discussions held during design process	N/A	3/ removed as design criterion	Affected people have own coping mechanisms	Relevance
				Added: Current emergency shelter is not comfortable owing to interior design issues	Focus group discussions held during design process	Interior design features	1	Identified by affected people as relevant to their comfort	Relevance, Effectiveness

7. Evaluation criteria for humanitarian innovation

The most commonly used evaluation criteria in humanitarian assistance are the OECD-DAC criteria (ALNAP, 2006). However, not all of these are appropriate for evaluating innovation processes and outputs, and their use can over-complicate an HI evaluation. For instance, connectedness and coherence typically have less relevance for assessing the value of an innovation. Other criteria, such as impact/sustainability or effectiveness, have a specific orientation when applied to an innovation process, output or scaled impact.

In this section, we outline the evaluative criteria relevant to evaluations of innovation processes, outputs and scaled impact. Before moving to the evaluative criteria, a few words should be said on a shadow criterion that often haunts evaluations of HI. Many donors are concerned with directing their innovation funds towards projects that are 'truly' innovations. This raises the question of whether evaluators assessing an 'innovation' project should assess the degree to which the project is indeed innovative. Evaluative questions that pertain to an innovation's degree of 'innovativeness' are questions about uniqueness. Uniqueness reflects the original contribution an innovation offers to the sector through its proposed improvement, or deviation, from standard practices. Uniqueness is often what donors and senior HQ managers have in mind when they use 'innovativeness' as a criterion for deciding which projects to fund in an innovation-specific programme or portfolio.

However, uniqueness is ill-suited as an evaluative criterion because novelty or uniqueness is never a valued characteristic of an innovation in itself: innovators are not engaged in innovation 'for the sake of innovation' but for the sake of achieving broader aims. While the feature of uniqueness is helpful for evaluators to bear in mind when understanding the degree of innovation the project they are assessing offers, it should not be used as an evaluative criterion. Annex III discusses the criterion of uniqueness in more detail and provides a graphic for evaluators to consider in assessing the uniqueness of an innovation.

The remainder of this section is divided into two parts. The first describes six evaluative criteria for the first two focus areas of HI evaluation: the innovation process and the innovation outputs. We then look at the third focus area of HI evaluation: the impact of the innovation and/or its learning through diffusion activities.

Below we describe in detail the evaluative criteria, providing:

- A definition of each evaluative criterion
- A brief explanation of what the criterion means in practice
- Considerations for evaluators regarding the evaluative criterion

7.1. Evaluating innovation processes and innovation outputs

Do No Harm

Definition
Taking precaution in order to minimise the potential harm caused to end users and primary beneficiaries by an innovation process

There remains little guidance for humanitarian practitioners on how to apply Do No Harm in practice, outside of the area of conflict sensitivity (CDA, 2016; Christopolos and Bonino, 2016). The need for such guidance has

been amplified by the rise of innovative practice in the humanitarian sector. While Do No Harm has been used to mean avoiding or minimising any adverse effects of an intervention on the affected population (Christopolos and Bonino, 2016), it is unclear how this principle should be weighed when trialling new approaches or interventions. On the one hand, there is potential harm that could occur; on the other hand, there are potential increased benefits.

To address how Do No Harm might be applied in an HI, some have focused on the overlap between innovation processes and research, seeking to apply a framework of research ethics to innovation processes (Jobanputra et al., 2016). Others tend to treat innovation as similar to standard humanitarian programming, and apply the same models of risk management (Betts et al., 2015). In other sectors, the precautionary principle has been discussed as a potential alternative to cost–benefit analysis as an approach to weighing harms and benefits under uncertainty (Steele, 2006; Meyerson, 2013). In lieu of more specific guidance, humanitarian innovators must often find their own way in balancing the potential harms and benefits of an innovation, the potential harms and benefits of not innovating and the uncertainty of these benefits and harms. Regardless, all innovation activity inherently bears some degree of risk owing to the uncertainty that characterises it: evaluators of HI should be conscious of this, seek to identify who bears which risks and build this line of inquiry into every evaluation.

Do No Harm: Considerations for the evaluator

Innovation outputs

- Were pilot participants or immediate contacts of pilot participants harmed by the innovation?

Innovation processes

- Did the innovating team consider potential harms and benefits of the innovation, and the probability of these occurring, at the outset of the innovation? Was this used to inform the design of the pilot?
- Think about how the innovating team framed expectations with pilot participants, particularly when participants were from affected communities. Did the innovation process raise expectations of future assistance or benefit that will not be delivered?
- Was it clearly communicated to participants in the innovation process, including pilot participants, that this was an innovation, i.e. not standard programming?
- How did this pilot or innovation process approach the dignity of participants and what steps were taken to uphold/respect this?
- What are the potential harms in scaling the innovation more widely?
- Was it clearly communicated to participants how data of their involvement were being captured and used, and did they provide informed consent for this use?
- What approaches did the innovating team take to collecting, storing and analysing data responsibly? Did they have a policy or framework in place for this?
- Was there harm caused in the process of the pilot?

Relevance and appropriateness

Definition
The extent to which the innovation responds to a recognised problem or meets end user needs and priorities

In standard EHA, the criteria of relevance and appropriateness are used to assess whether humanitarian programming is sufficiently demand-driven. In an HI, these criteria can be used similarly to determine whether the innovation is responding to a real need or problem, or whether it is primarily the perceptions and interests of the innovators that are driving it.

There are nuances required in how the relevance and appropriateness criteria are applied. The rise of innovation in humanitarian action has led to a proliferation of projects with some type of 'new', typically technology-focused, component. Relevance and appropriateness are particularly important criteria, as they highlight the desire for innovations to be not simply 'new' but also demand-driven, and to originate out of end users' needs and perspectives. This is particularly important for innovations that seek to bring about improvement for affected people as their end users. For each innovation process in which relevance and appropriateness are used as criteria, the evaluator should seek to answer: What problem was the innovation trying to address? Who was this a problem for? Did the demand for the innovation come from them, or from the innovating team? And how did the innovation process incorporate these stakeholders' needs and preferences?

In some cases it can be difficult to apply the relevance and appropriateness criteria. Some innovations – particularly those characterised as more 'radical' innovations bringing significant shifts in technologies or practices – offer an improvement precisely because they do *not* respond to a recognised problem. Instead, these radical innovations are responding to a problem many people do not yet see. Radical innovations can shift how end users think about their own priorities, fulfilling needs they do not yet recognise prior to the innovation being created. In these cases, an innovation may be **appropriate** but not perceived as **relevant**. However, in order for such an innovation to become widely accepted and successfully diffused, the innovation must at some point come to be considered relevant by a core set of gate keepers and end users. It is essential that teams engaged in radical innovation have taken steps to understand the wider humanitarian sector and how they will build recognition for the problem or performance area they are addressing. Even the most 'disruptive' innovations ultimately rely on their solution being relevant and appropriate to end users in order to succeed. Also, it is important to remember that the resources dedicated to HI could have been used for direct lifesaving activities instead, and therefore evaluations of radical innovations should examine closely how the innovating team has made headway in generating wider recognition and change in humanitarian actors around the problem area they are seeking to address.

Relevance and Appropriateness: Considerations for the evaluator

Innovation outputs

- Does the innovation address a clear need?
- To what extent did the innovation meet the needs of its intended end users and primary beneficiaries and how do we know?
- To what extent was the innovation accepted by end users as meeting their needs?
- Did the demand for the innovation come from primary end users, or from the innovating team?
- Was the consolidated learning and evidence produced and presented in a way that was relevant to ongoing evidence generation and research on this issue?

Innovation processes

- Map the feedback loops that the innovating team established between themselves and pilot participants or other end users. How did the innovation team incorporate end users' needs and preferences into the development and design?'
- If the innovation was opportunity-driven, what did the innovating team do, in terms of information collection or research, to determine that this opportunity was addressing a clear need and would be valued by potential end users?

Efficiency

Definition
The rate at which inputs are converted into valued outputs and outcomes

Efficiency is concerned with the amount of valued outputs or outcomes achieved by the inputs of an intervention or activity.[3] Evaluators will apply this criterion to HI differently depending on whether the focus of an evaluation is on the innovation itself or the innovation process.

When evaluating the innovation itself, such as an SMS-based communication tool or a new approach to psychosocial care for victims of gender-based violence, the concern is how the SMS tool or psychosocial approach improves the efficiency of humanitarian interventions. Assessing the efficiency of an innovation relies on a comparison with current practices and approaches, and it is the task of the evaluator to work with an innovating team to identify the appropriate comparator for evaluating efficiency (see more below under comparative improvement). Ideally, an innovating team will already have some form of comparative assessment tool in place, as such tools play an important role in guiding the testing and refinement of a prototype.

In some cases, evaluators may be asked to consider the scope for efficiency more widely, taking into account the efficiency savings the innovation would provide if it were to be widely diffused. An innovation process may produce a useful piece of software that offers advantages in efficiency over a humanitarian agency's current systems. However, if the process taken to develop the software lasts many years and runs significantly over budget, senior HQ managers and donors may question whether it was actually an efficient and good use of

resources. The iterative nature of innovation creates the risk that the outputs of an innovation process will not produce enough benefit relative to the inputs that went into the process and thereby waste funds that could go towards humanitarian assistance. A further issue complicating this is that most humanitarian agencies do not collect the kind of performance data they need in order to make informed decisions on how much is a reasonable amount to allocate to innovation and how much is too much. Many humanitarian agencies are unable to assess whether innovation offers higher efficiencies because they do not have accurate efficiency measures for their current interventions.

To assess the full cost of an innovation output, the efficiency gains of the innovation should first be calculated against some comparator, as one would do when evaluating the efficiency of the innovation itself. For example, new software that enables SMS-surveys with affected populations should be compared with the costs of face-to-face surveys, or with other SMS-based surveys.

Using the roll-out or scaling plan, or an evaluation of how successfully the innovation has been diffused (depending on the timing of the evaluation), the evaluator can then calculate an estimate of the efficiency gains through the innovation over time – for example multiply the efficiency savings in a single use of the software by the number of surveys an organisation expects to carry out over the next two years. Finally, the cost of the innovation process (including staff time) can be subtracted from these savings in order to reach an understanding of the 'net' efficiency gains.

Finally, regarding the outputs of an innovation process, it is important to remember that, while not all processes will produce an effective or workable prototype, they should all produce some form of consolidated and usable learning. Considering the value of this learning is important for getting an accurate picture of an innovation's outputs and outcomes, and therefore its efficiency.

More and more, evaluation stakeholders within humanitarian agencies may be interested in evaluating the efficiency of the innovation process. This is a complicated question. Looking at the efficiency of innovation processes requires an awareness of how innovating teams convert inputs into outputs and in which areas innovating teams can exercise some control over the amount of resources expended. Based on the 15 case studies carried out by ALNAP from 2014-16, evaluators could probe the following aspects of an innovation process to assess its efficiency:

Moving from brainstorm to decision: Many innovating teams report that, during the ideation activities – in which teams brainstorm many potential approaches and design features for the innovation – there is the sense that this creative sharing of ideas could go on *ad infinitum*. Efficient innovation processes will find a way to guide this brainstorming to a definitive decision and the creation of a basic plan for the development of the prototype fairly early; while teams can always go back to brainstorming as they learn more about the innovation idea in development activities, it is critical to have a timely transition from ideation to the deeper learning processes that occur through development.

Strategic collaboration: While relationships are critical to the success of an innovation, they can also pose huge resource drains. Efficient innovating teams will be strategic about who they engage with and when, potentially working with a wide number of stakeholders at the outset of an innovation to generate ideas and then narrowing their engagement to select advisors or end users. Evaluators can look at how relationships were approached and managed for different stages and purposes across the innovation process in order to assess efficiency.

Managing and integrating feedback: Similar to relationships, feedback loops can serve different functions and therefore should be designed and managed differently in order to use staff resources most effectively. For some innovating teams, efficiency savings were gained by creating a division of labour between different pilot sites. For others, feedback was initially gathered internally or with end users who were 'friends' of the innovating team on early prototypes.

Capturing internal learning for external purposes: Innovating teams that have strong capture processes for their internal learning can more efficiently convert this into materials appropriate for external audiences, such as blogs, case studies, conference presentations or pamphlets.

Linking the innovation to other organisational processes: Bringing about radical changes through a major paradigm innovation can be very expensive for an organisation, but if it is seen as part of their core mission these costs could be viewed as acceptable and appropriate. Evaluators should use triangulation and cross-organisational surveys or interviews to seek an understanding of how the innovating team linked the innovation to core organisational costs or drew on broader organisational resources outside the project budget to support the innovation process.

Bootstrapping: In innovation processes that involve multiple components, it is possible for innovating teams to enhance their efficiency by 'bootstrapping' – that is, adapting and incorporating existing products and approaches into parts of the innovation in order to cut down on the costs of the design process. For example, a technology-driven innovation that seeks to bring about a change in practice might use existing software rather than invest in developing a new software from scratch, in order to reduce costs and focus resources around the behaviour change they are seeking to achieve (Obrecht, 2016).

Efficiency: Considerations for the evaluator

Innovation outputs

- Consider how to set the frame for understanding the efficiency of the innovation or innovation process. This typically requires identifying the input–output/outcome ratio for comparable current practices and approaches.
- Consider how scaling affects the assessment of the efficiency of the innovation: the wider the innovation is planned to be used, or is being used, the more efficient it is.

Innovation processes

- Why and how did innovating teams manage their resources, including staff time, throughout the innovation process? How did these choices affect the cost of the overall process?
- Did the team learn from expensive mistakes or did they repeat these mistakes?

Effectiveness: Learning; comparative improvement

Learning

Definition
The degree to which the innovation process generates new knowledge or evidence

The generic OECD–DAC criterion of effectiveness can be specified for the evaluation of humanitarian innovation by focusing on two key outcomes for any innovation process. Learning is the first of these desired outcomes and the easiest one for innovating teams to achieve, as it rests primarily within their own control.

Learning is an important component in any innovation process. Evaluators should be aware, however, that innovating teams may be strong on internal learning but weak on consolidating and sharing such information externally with others. Depending on the focus of the evaluation, evaluators may find it useful to identify a set of 'experts' in the area the innovation addresses and work with them as key informants to understand how relevant and important the learning from the innovation process is to their work.

Learning: Considerations for the evaluator

Innovation outputs

- Has evidence in the humanitarian sector around this issue been strengthened?
- Has there been an effort to report on the findings and key lessons from the innovation process and disseminate this widely to external experts in the relevant area addressed by the innovation?

Innovation processes

- What learning processes or practices did the innovating team use and how effective were these?
- Did the organisation/programme team learn from the process in a way that will affect future programming or attempts at innovation?

Comparative improvement

Definition
The measurable, comparative improvement in effectiveness, quality or efficiency the innovation offers over current practices.

This criterion applies specifically to evaluations focused on the innovation itself and is the criterion most closely related to the effectiveness criterion in the standard OECD–DAC framework.

Innovation is meant to offer a significant improvement over how humanitarian assistance is currently structured and provided, which serves as the rationale for tolerating the higher degree of uncertainty and costliness that can come with innovation activities. This means that, for HI, 'effectiveness' means something quite specific:

it is not concerned merely with whether an innovation has achieved its stated objectives; as discussed above, in some cases stated objectives have less relevance at the end of an innovation process, as they shift in light of new knowledge generated about the problem area. For innovations to be effective they must offer **measurable comparative improvements** over current practices and approaches. Innovations that offer a change to humanitarian action but no discernible improvement in performance are likely not only to be highly inefficient but will also fail to deliver the specific type of effectiveness expected of innovation activities. Comparative analysis should therefore be at the centre of an evaluative assessment of a prototype or workable innovation.

The primary challenge for evaluators in applying this evaluative criterion will be the lack of adequate performance data on current humanitarian practices and approaches. Ideally, resources should be allocated within an HI evaluation to assess the innovation against its design criteria or agreed standards, and also to analyse relevant comparable practices or tools against the same evaluative framework. Evaluators of HI should prepare to devote some considered attention to what would serve as an appropriate and relevant comparator for the innovation they are evaluating.

Comparative improvement: Considerations for the evaluator

Innovation outputs (the innovation itself)

- Try to use key informant interviews with sector experts who are not involved with the innovation project in order to gain some objectivity on the most relevant comparators for the innovation and the relevant metrics for measuring improvement.
- Consider the performance measurement practices around current humanitarian action in this area and how you might be able to draw on existing performance data to draw a comparison.
- Consider relevant technical standards and how these might be used as a metric for the comparison.
- Does the innovation offer a comparative improvement in the coverage, timeliness, relevance, connectedness, coherence, effectiveness and/or impact of humanitarian assistance?
- Does the innovation offer a better solution to the problem it seeks to address compared to current approaches?
- Does the innovation out-perform comparative approaches in meeting or exceeding relevant technical standards?
- In cases where innovations are seeking to address unacknowledged problems or issues in the sector, there may not be an obvious practice to compare. Consider the characterisation of the problem and what an alternative solution (other than the innovation developed) would have been. This can be used as an analytical device to think about the choices made by the innovating team.

7.2. Scaled outcomes/impact

As described in Section 4.3, the fourth focus of an evaluation is the outcomes or impact of an innovation when it is taken to scale and diffused widely across multiple end users. While evaluators may be able to look at issues such as uptake, sustainability or impact at the pilot level, typically these become more relevant once an innovation has progressed in diffusion or scaling activities. Through these activities, innovating teams seek to achieve their value proposition at scale, and they – as well as those who fund their diffusion activities – will be keen to understand what impact their innovation is having at a broader level.

It is also important not to lose sight of the consolidated learning and evidence produced by the innovation process when thinking about uptake and impact. Innovating teams should hopefully include room in their diffusion plans for diffusing learning as well as the innovation itself. Innovating teams that experience an unsuccessful pilot might still engage in diffusion activities but focus these activities around uptake of the learning from their experience, in order to influence and inform future attempts to address the targeted performance issue.

Uptake or sustainability

Definition
The extent to which others adopt and use the innovation to improve humanitarian performance

Uptake and sustainability are two related criteria concerned with the lifespan of the innovation. Uptake refers to the ownership and adoption of the innovation by targeted end users. Sustainability broadly refers to the ability of end users to use and implement the innovation without continued support from the innovating team. Many donors of humanitarian innovation tend to think of sustainability more narrowly and in financial terms, seeking to support innovations that can eventually be scaled through business models that do not rely on continuous grant funding.

Evaluating uptake can consist primarily in understanding how many external actors have adopted or committed to adopting the innovation. As with all aspects of HI evaluation, asking 'how' and 'why' questions are important. Understanding why external actors adopted the innovation and how they found a way to integrate it with their existing approaches can be critical to generating learning on what kinds of scaling and diffusion strategies work best. This is important, as evidence surrounding effective scaling models in humanitarian contexts is currently weak.

Timing is a major factor that will shape an evaluation's assessment of uptake. **Realistically, an assessment of uptake cannot be meaningfully carried out until at least one year after organised diffusion activities have begun.** Even then, for many innovations this will offer only a partial picture of uptake; any evaluation should acknowledge these limitations.

Sustainability will be a relevant criterion for evaluations focusing on the scalability of the innovation. This is discussed further in the Annex III. For evaluations focused on the innovation **process**, assessing uptake will involve a focus on the diffusion strategy or business model used by the innovating team.

In evaluating sustainability, several different business models have been used to diffuse humanitarian innovations (see Box 7); evaluators may want to be aware of these when discussing the diffusion strategy with

the innovating team. There has been little work thus far to understand what business models may be most effective in securing sustainability for innovations in the humanitarian sector. Innovation researchers are currently exploring how to adapt this model for a non-profit environment (McClure and Gray 2015; Tidd and Bessant, 2016) and this may serve as a fruitful area for guidance for humanitarian innovators on sustainably resourcing the diffusion of their innovations.

Box 7. Business models for scaling humanitarian innovations.

Below are four business models identified in the 15 HIF-ALNAP case studies, with examples from the case studies identified in parenthesis.

- Partner with a private sector organisation that will produce the innovation at scale and market it (e.g. WFP's mVAM, Improving Water Quality in Emergencies).
- Partner with a government that will take on innovation and implement it (community-based financing for disaster risk reduction, Linking Communities to Mine Action).
- Provide the service for free, or sell it directly to humanitarian agencies (Motivation's appropriate and affordable wheelchairs, Words of Relief, The CMAM Report, The Humanitarian Lessons-learned Genome Project, Mapping a Response, Humanitarian eXchange Language).
- Roll out internally (Improving Menstrual Hygiene Management in Emergencies, The CMAM Report, WFP's mVAM, SMS Feedback in Somalia, Gaza Risk Reduction and Mitigation).

Uptake and Sustainability: Considerations for the evaluator

Innovation outputs

- How did organisations hear about the innovation and why did they feel compelled to adopt it and use it themselves? Or why did they not use it?
- What did the innovating team do to influence diffusion and what external factors, or changes in the broader ecosystem, influenced diffusion?

Innovation processes

- What is the scale of uptake? How many external organisations or actors are supporting or using the innovation, and in how many countries?
- Is the innovation sustainable over time? Consider what sustainable funding means for the users of your HI evaluation. Does it mean a specific business model, e.g. that the innovation is funded publically through state budgets, or taken on and funded by local and national actors, or self-funded through a for-profit model? Or does it simply mean the innovation does not require external grant funding?

Impact

Definition

Looks at the effects of the diffused/scaled innovation or the diffused consolidated learning produced by the innovation process. These effects 'can be social, economic, technical and environmental, and occur in individuals, gender, age-groups, communities and institutions. Impacts can be intended and unintended, positive and negative, macro (sector) and micro (household, individual), short or long term' (**EHA 2016**).

Proper impact evaluations are extremely important for evaluations of mature innovations. Uptake and sustainability are at best only proxy measures for impact: an innovation could be widely adopted but fail to effect a significant change in practice or performance.

Impact assessments are in general challenging to undertake, as they require strong research methods to establish causal relations between the activity/intervention of interest and the longer-term and broader effects. Impact assessments are particularly challenging for HI because of the timing factor mentioned earlier. Simply measuring uptake may require a significant time lag from the end of the innovation activities. Measuring the effects of this uptake, which would constitute the innovation's impact, requires an even further time lag, possibly five or more years from the start of the innovation process.

A number of methods for assessing impact can be considered when looking at an innovation, many of which are similar to the methods used for standard impact assessment. One method that may be particularly useful for evaluating humanitarian innovation is **Outcome Mapping** or Outcome Harvesting. Outcome Mapping, developed by the International Development and Relief Council in 2002 to support the assessment of the impact of complex development processes, is 'a methodology for planning and assessing programming through the gathering of information on the outcomes, defined as behavioural changes, of a change process.' (IDRC, 2012).

Outcome Mapping is not strictly an evaluation method, but is used also for the monitoring and even design of programming, particularly programming that is oriented towards change and social transformation. Its use in the design of advocacy strategies implies it may also be highly relevant to humanitarian innovators in their diffusion activities, as there is significant overlap between advocacy and the types of outreach needed to successfully diffuse an innovation. If an innovating team has used methods or techniques from Outcome Mapping in its diffusion activities, this can provide a useful entry point for evaluating the impact of the diffused innovation. Outcome Mapping is described in more detail in the companion paper to this working paper, '*Monitoring humanitarian innovation*.'

Impact: Considerations for the evaluator

- Will the evaluation seek to establish contribution or attribution, and with what methods?
- Are there system- or sector-wide performance measures that can be used to determine if the scaled innovation has contributed to a broader performance change?
- What are the observable behaviour changes (outcomes) that would be expected if the diffused innovation had its intended impact?
- Consider building in questions that probe the unintended consequences of the diffused innovation.
- Consider how you will harvest quality baseline data; for innovation managers at the start of an innovation, consider how to collect baseline data during the innovation process for use in a later impact study, based on expected or desired outcomes from the innovation.

8. Conclusion

The purpose behind the ALNAP working papers, *Evaluating Humanitarian Innovation* and *Monitoring Progress in Humanitarian Innovation Processes* (forthcoming in February 2017) is to explore how we might approach monitoring and evaluation practices for iterative processes like innovation. There are other early attempts to explore these themes (Valters et al. 2016; ODI 2016), however further thinking and trialling of the monitoring and evaluation of iterative, learning-intensive processes is needed. This is important not only for the expanding arena of humanitarian innovation, but also for so-called 'standard humanitarian programming', where concepts from innovation practice, such as iterative learning, adaptiveness and 'testing' of ideas, are beginning to influence the way humanitarians think about programme management (Mercy Corps and IRC 2016).

Annex I: Developmental evaluation and its value for innovators

The iterative nature of innovation creates demand for strong learning practices and processes. Evaluative inquiry is integral to learning, and is therefore also one of the essential tools in the humanitarian innovator's toolkit. There are many forms of evaluative inquiry, but the most relevant for HI is likely to be **developmental evaluation,** a type of evaluation that arose specifically to meet the evaluative and learning needs of social innovations.

Developmental evaluation was created by Michael Quinn Patton in the 1990s as an alternative to formative and summative evaluation. It arose out of a demand for his evaluation services in a social innovation seeking to improve the leadership skills of at-risk youth. After engaging in formative evaluation to improve the programme model several times, Patton found that his client wanted to continue using evaluation in this way – tweaking and changing the programme – rather than finalising the model and carrying out a summative evaluation: 'What was judged to be working was not a standardized and routinized model, but rather the ongoing development of leadership programming in response to changing conditions, lessons learned, and the emergent needs of different kinds of participants as the program expanded its outreach' (2010: 4).

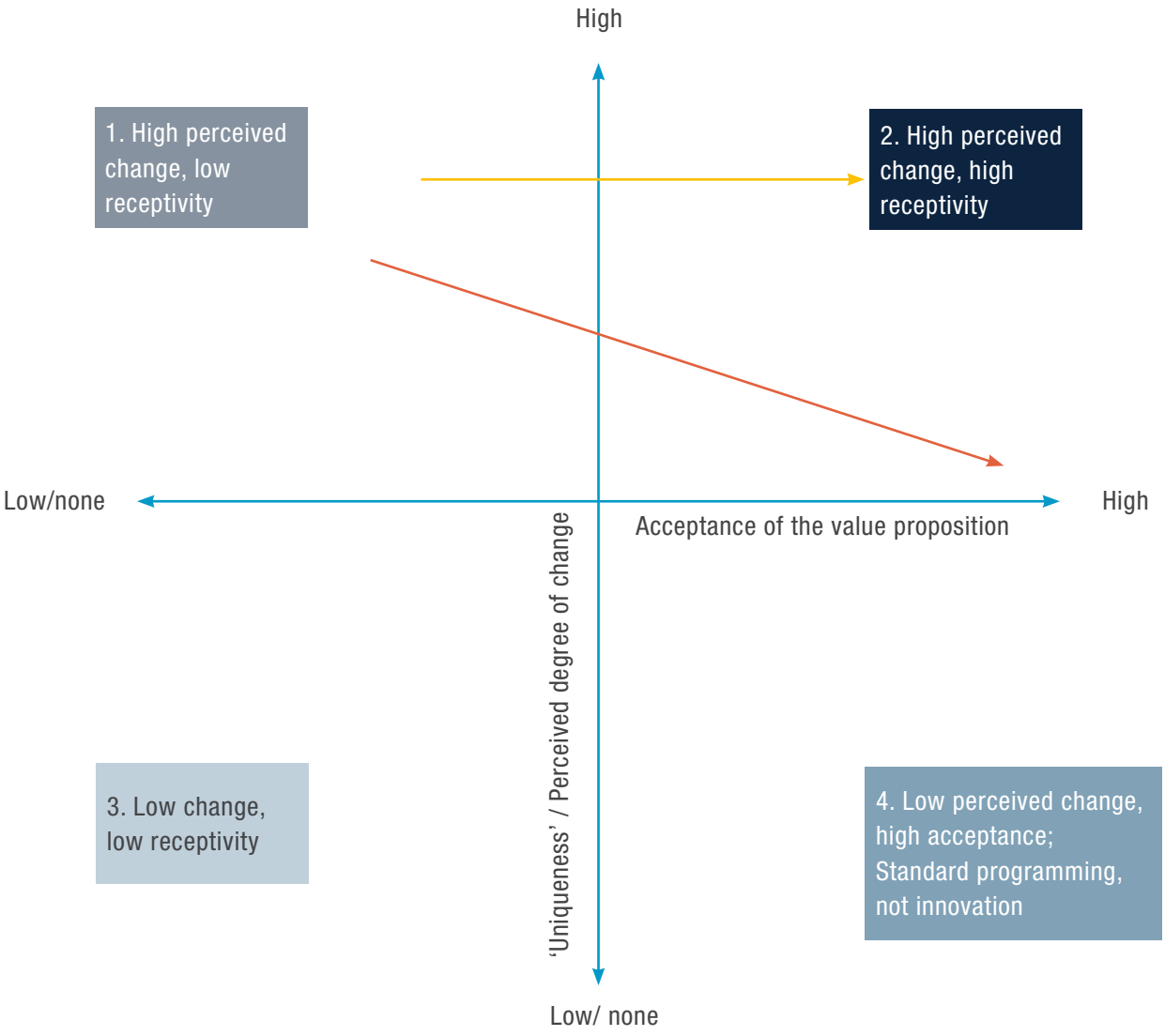
From this experience, Patton began to develop evaluative approaches that could be used to support continuous iteration, particularly in programming contexts featuring high degrees of uncertainty and complexity. The result is 'developmental evaluation', evaluative inquiry used to 'support the exploration and conceptualisation of an innovative idea and help innovators clarify, focus and articulate what they are trying to do as they do it' (Patton, 2010: 39). Developmental evaluation is essentially evaluation put to the purpose of developing an innovation, from early exploration of ideas to design and testing. Developmental evaluation structures and provides systematic feedback to an innovating team to help it reflect on, assess and then pivot its design or approach. In contrast to formative evaluation, which seeks to identify improvements for a fairly fixed model, developmental evaluation embraces the core principles of innovation, forgoing any approach that treats the intervention as a 'fixed' model, and instead facilitating an ongoing inquiry that calls into question what the model should be and what are the desired outcomes.

If using developmental evaluation, innovators would work alongside, or in consultation with, an evaluator throughout the ideation, development and implementation activities of the innovation process. While this would require extra resourcing, it could be extremely advantageous for humanitarian innovators. A developmental evaluator can provide a valued independent perspective to identify blind spots and unexploited opportunities in the design process. Using developmental evaluation also helps document and rationalise the decision-making process, which can be critical for later summative evaluations of the innovation (see Section 6). It can also help build a clearer narrative around the innovation itself, helping an innovating team understand its own innovation better, and communicate it clearly to external stakeholders to achieve wide uptake. Finally, developmental evaluation provides a structure that can enable better use and uptake of the consolidated learning generated through the innovation process, informing future innovation processes.

Annex II: Assessing the 'innovativeness' of an innovation

Uniqueness is a challenging criterion to work with as an evaluator, as it is greatly affected by the wider environment and often changes as an innovation process progresses. Paradoxically, in order for innovations to scale successfully, they often end up losing their uniqueness. The figure below depicts the four quadrants that a project or innovation can occupy depending on the deviation from standard practice offered by the idea and the degree of acceptance of the idea.

Figure 2. Mapping the uniqueness and acceptance of an innovation's value proposition



Projects that occupy position 4 are not truly innovations: these projects are very similar to current practice and are widely accepted. Many innovations start at position 1. They then can move directly across to position 2 (yellow line), achieving increasing buy-in while retaining their uniqueness. These are more radical innovations, which offer significant changes and are able to generate high buy-in without compromising on the degree of change offered, and without broader practices shifting to make them less innovative in comparison as they are diffused. However, these cases are quite rare.

More commonly, innovations have a downward trajectory (red line) as they move towards greater acceptance and uptake. This downward trajectory in uniqueness can be caused by innovating teams making compromises to the design in order to accommodate current practice, or, importantly, by the surrounding environment shifting and changing to become more like the innovation, thereby reducing its originality. As wider acceptance for the innovation is gained across end users, copycats and similar initiatives are likely to emerge to meet the demand, thereby lowering the uniqueness of the innovation.

For example, as cash-based programming became more widely used, it becomes less and less innovative, transitioning into a form of standard programming (position 4). Or new technologies offering a radical shift in how humanitarians operate in crisis tend to offer too radical a change and must be modified by innovators, reducing the degree of change they offer in order to obtain wider acceptance and uptake (move towards position 2, then rightward towards position 4). Evaluators may therefore want to consider how the innovation process worked to shift the innovation idea across this spectrum over time, by assessing the environment around the innovation at the time when the idea was first identified (where it is most likely to be at position 1) and comparing to where it is at the point of evaluation (where it may be closer to 2, 3, or 4).

Uniqueness: Considerations for the evaluator

- How have the existing and prior attempts at addressing the problem area been evidenced by the innovating team? How does the innovation's value proposition differ from these?
- What is the scope of the innovation's novelty? Is it: new to the organisation (but practised elsewhere); new to the humanitarian sub-sector (but used in other humanitarian sub-sectors); new to the humanitarian sector as a whole (but used in development practice or private business); or new to the world (has not been used before in any recognisable format)?

Annex III: Assessing scalability

Scalability refers to the viability of the innovating team's plans for diffusing an innovation in order to achieve impact through wider use. When practitioners speak about successful innovations, they often have in mind an innovation that not only improves humanitarian action but also is successfully adopted by a wide range of actors. Scaling, or diffusion, is typically the hardest part of innovation practice. It can also be extremely costly. For this reason, some stakeholders in humanitarian innovation, particularly donors and senior HQ management positions, may want to see some initial evidence that a piloted prototype has the potential for scaling before they provide financial support to diffusion activities. Some funders of HI may even wish to evaluate the potential for scaling before funding the development and piloting of an innovation idea. This may be an increasingly popular approach as donors and agencies attempt to subvert the trend of investing more resources in piloting new ideas than in actually diffusing them to achieve wide impact.

While not strictly a focus of evaluation, scaling assessments will involve evaluative criteria and forms of evaluative inquiry. Scalability requires its own separate criteria, as assessments of scalability are typically formative and directed towards understanding whether further resources should be directed towards an innovation that has passed the milestone of a successful pilot. Based on the ALNAP case studies, as well as other studies on the factors that contribute to successful scaling (World Bank, 2012), potential evaluative criteria for assessing the scalability of a prototype include:

- **Proven value proposition:** Evaluations of scalability should consider whether there has been a summative evaluation or other evidence generated to demonstrate that the prototype has met the evaluative criterion of comparative improvement.
- **Relevance:** While having proof for an innovation's value proposition can be useful for generating wider support during scaling, a good degree of recognition and acceptance of the value proposition among potential end users is also very helpful for scalability. Relevance in this context refers to the scope of acceptance of the innovation's value proposition: both the problem it sets out to address and its proposed solution.
- **Deviation from current practice:** Even innovations that meet the needs of end users may not be effectively and efficiently scalable if they require too significant a deviation from current practices. Radical innovations can appear to have high costs to potential end users, given the high degree of change they involve. Innovations that offer a significant departure from current practice may need to rate highly on the other criteria in order to be considered worthy of scaling.
- **Credibility of innovating organisation:** The presence of a strong advocate is a key factor for success throughout an innovation process. Once an innovation reaches the diffusion stage, it is not only the presence of this advocate that matters for success but also the organisation for which they work and the position and credibility that this organisation has in the broader network of actors influencing the uptake of the innovation. Here, it is particularly important to look beyond direct end users of an innovation to identify the gate keepers: those who do not directly use an innovation but who influence the incentives and rules that can keep an innovation from being adopted, or can support its wider use.

Bibliography

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

Betts, A. et al. (2015) *Principles for Ethical Humanitarian Innovation*. PHAP Online Consultation. Retrieved 15 September 2016, from http://www.elrha.org/wp-content/uploads/2015/08/150708-Principles-for-Ethical-Humanitarian-Innovation-event-report.pdf

Buchanan-Smith, M., Cosgrave, J., & Warner, A. (2016). *The Evaluation of Humanitarian Action Guide*. London: ALNAP/ODI.

Christoplos, I. & Bonino, F. (2016) *Evaluating protection in humanitarian action: Focus on decision-making processes and options to address common issues and challenges*. ALNAP Pilot Guide. London: ALNAP/ODI.

Sheather J, Jobanputra K, Schopper D, Pringle J, Venis S, Wong S, et al. (2016) A Médecins Sans Frontières Ethics Framework for Humanitarian Innovation. PLoS Med 13(9): e1002111. doi:10.1371/journal.pmed.1002111

Kimball, L. (2015). *The Service Innovation Handbook: Action-Oriented Creative Thinking Tool Kit for Service Organizations*. Oxford: BIS Publishers.

McClure, D. and Gray, I. (2015) 'Engineering scale up in humanitarian innovations missing middle', in *IEEE 2015 Global Humanitarian Technology Conference*. 114–122.

McClure, D. and Gray, I. (2014) *Scaling: innovation's missing middle*. Oxford: Humanitarian Innovation Programme Conference 2014.

Meyerson, D. (2013). Innovative Surgery and the Precautionary Principle. *Journal of Medicine and Philosophy*, 38: 605–624.

Obrecht, A. (2016) 'Linking communities to mine action: Innovating in active conflict' HIF/ALNAP Case Study. London: ALNAP/ODI.

Obrecht, A., & Warner, A. (2016) *More than just luck: Innovation in Humanitarian Action*. London: ALNAP/ODI.

Patton, M. Q. (2010) *Developmental evaluation: Applying complexity concepts to enhance innovation and use*. London: Guilford Press.

Ramalingam, B., Rush, H., Bessant, J., Marshall, N., Gray, B., Hoffman, K., Bayley, S., Gray, I. and Warren, K. (2015) *Strengthening the humanitarian innovation ecosystem*. Brighton: CENTRIM.

Renard, R. and Lister, S. (2013) *Technical Note on Efficiency Analysis*. Geneva: World Food Programme.

Sandvik, K. (2016) Now Is the Time to Deliver': Understanding Humanitarian Innovation as Buzzword. Presentation at the World Conference on Humanitarian Studies, Addis Ababa, Ethiopia, 5-8 March 2016.

Steele, K. (2006) The precautionary principle: A new approach to public decision-making? *Law, Probability and Risk* 5:19–31.

Stoddard, A., et al (2015). *The 2015 State of the Humanitarian System Report*. London: ALNAP/ODI.

Tidd, J. and Bessant, J. (2005). Managing Innovation: Integrating Technological, Market and Organizational Change. New York: Wiley.

Tidd, J. and Bessant, J. (2016) *The entrepreneur's journey*. Chichester: John Wiley & Sons.

Wallace, M. (2015). From Principle to Practice: A User's Guide to Do No Harm. Cambridge: CDA Collaborative Learning Projects. Retrieved 02 September, 2016, from http://cdacollaborative.org/wordpress/wp-content/uploads/2016/02/From-Principle-to-Practice-A-Users-Guide-to-Do-No-Harm.pdf

Footnotes

1. For tips on identifying your primary intended user, see EHA (2016: Section 3).
2. For more guidance on developing an evaluative question see EHA (2016: Section 6).
3. Classically, efficiency refers to input–output ratios and cost-effectiveness refers to input–outcome ratios; however, there is some argument for treating cost-effectiveness as the most important, or 'priority' way of assessing overall efficiency (Renard and Lister 2013).

Other case studies from Elrha and ALNAP on innovation

Research papers

More than just luck: innovation in humanitarian action

Case studies and working papers

Monitoring humanitarian innovation

Mapping a response:
Using satellite images to aid humanitarian action

Improving menstrual hygiene in emergencies:
IFRC's MHM Kit

Understanding the performance of emergency feeding programmes

Using mobile voice technology to improve the collection of food security data

Improving water quality and quantity in emergencies:
The Inclined Plate Settler water treatment system

A community financing mechanism for disaster risk reduction: The Bio-rights approach

Words of Relief: Translators without Borders' local language translation for emergencies

Supporting disabled people in emergencies: Motivation's appropriate and affordable wheelchairs

Innovations in international humanitarian action:
ALNAP's 8th Review of Humanitarian Action

www.elrha.org/hif

www.alnap.org/innovation

Humanitarian Innovation Fund
Elrha
C/o Save the Children
1 St John's Lane
London EC1M 4AR

ALNAP
Overseas Development Institute
203 Blackfriars Road
London SE1 8NJ
United Kingdom
Email: alnap@alnap.org