

From Barriers to Breakthroughs: Co-Creating Impact Evaluations in Humanitarian Aid



Photo Credit: Fabeha Monir/Save The Children

Humanitarian Assistance Evidence Cycle

There is an unsettling reality in the humanitarian space that while humanitarian need continues to grow, funding to meet that need is increasingly insufficient. This underscores the importance of ensuring current humanitarian assistance programming is as impactful and cost-effective as possible. Determining the viability of any humanitarian assistance program requires an understanding of the program's efficacy both in terms of the outcomes achieved and its cost-effectiveness. Impact evaluations can provide insights into both questions.

However, there is a dearth in the availability of rigorous, high-quality evidence to inform humanitarian response limited by a number of barriers. There are wide-ranging misconceptions around impact evaluations, including what they can answer and what sets them apart from other evaluation methods. Implementers perceive high risk in conducting impact evaluations, worried they will call their credibility and future funding into question. Misaligned research partnerships can reduce the value of learning products for implementers, further diminishing the demand. Humanitarian contexts also present unique research implementation challenges for design or data collection. These and other constraints reduce the demand for impact evaluations and constrain their use in humanitarian contexts.

To address these challenges, the [Humanitarian Assistance Evidence Cycle](#) (HAEC) encourages implementing partners to adopt a 'co-creation' process when engaging researchers to conduct impact evaluations. This approach facilitates collaboration from an early stage in humanitarian program implementation, ensuring that research questions are aligned with the program's needs and that both implementers and researchers are fully invested in the evaluation's success. Through this process, HAEC learned that demand for impact evaluations does exist. However, converting this into successful research studies hinges on establishing an effective process early in research's inception. By fostering collaborative partnerships, the research co-creation process helped to overcome existing barriers and enhance the effectiveness of humanitarian impact evaluations.

What is an Impact Evaluation?

Impact evaluations are a systematic approach used to ascertain if a causal relationship exists between a program or intervention and the observed changes in outcomes. In essence, they seek to answer the fundamental question: ***“Would outcomes have turned out differently if the intervention had never been implemented in the first place?”***



USAID
FROM THE AMERICAN PEOPLE



Background

Conflict, crisis, and disasters threaten the lives, rights, and security of millions around the world. [In 2023](#), the United Nations Office for the Coordination of Humanitarian Affairs (OCHA) estimated that 339 million people required humanitarian assistance and protection, an increase of 65 million people compared to the previous year. However, the current funding for humanitarian responses worldwide struggles to keep pace in the face of this persistently expanding need. ALNAP’s [2022 State of the Humanitarian System Report](#) highlights that between 40-50% of UN-coordinated appeals, which are the humanitarian system’s best collective estimate of needs and costs, were unmet in the last decade.

The growing need and limited funding for humanitarian assistance reinforces the need to optimize humanitarian assistance programming to ensure it is as impactful and cost-effective as possible. Impact evaluations provide insights into the efficacy of humanitarian programs in terms of outcomes achieved and cost effectiveness. This provides decision makers with important information to inform their allocation of limited humanitarian programming resources. However, compared to development sector programs, there is far less rigorous, high-quality evidence to inform humanitarian response, as show in HAEC’s [Evidence Gap Map](#). There are numerous barriers that limit the utilization of impact evaluations in humanitarian contexts, as outlined in the report [Navigating Constraints to Impact Evaluations in Humanitarian Settings](#).

Increasing Demand for Impact Evaluations: Lessons from Research Co-Creation

The HAEC Activity works to increase the utilization of impact evaluations to improve the effectiveness and cost-effectiveness of emergency food security activities funded by USAID’s Bureau for Humanitarian Assistance (BHA). During 2023, HAEC funded a portfolio of impact evaluations of BHA’s emergency food security activities through a funding application process to accommodate and work within common constraints. HAEC received 29 expressions of interest for the funding. These eventually converted into 15 applicants beginning the research co-creation process and six activities ultimately receiving funding to implement an impact evaluation (*see Figure 1*).

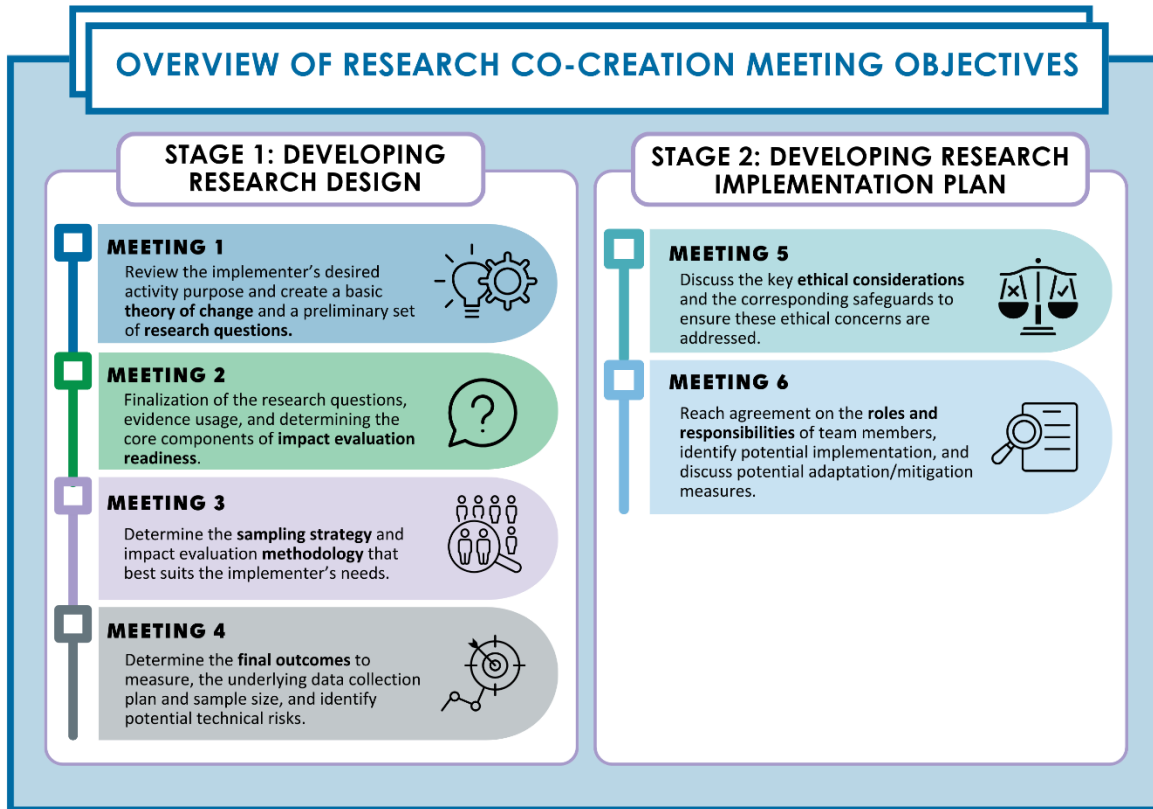
There were three phases during the application process. During the first phase applicants submitted their expression of interest indicating information about their activity, their research question(s) of interest, and capacity for conducting impact evaluations. HAEC then invited applicants to the second phase, which it termed research co-creation, during which IPs and researchers collaborated with HAEC to create a full impact evaluation design for their funding application. During this process, HAEC supported applicants to develop research question(s), identify potential evaluation methods and research partners, and develop a management/implementation plan. HAEC facilitated six one-hour meetings between the research teams

Figure 1: Application Overview



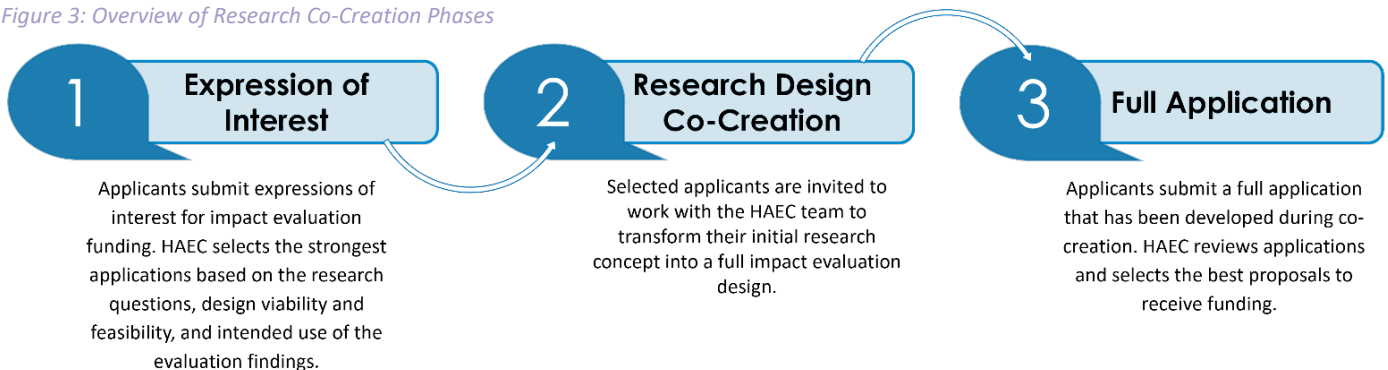
(representatives from the implementing organization and research partners). Each meeting focused on specific research design objectives and the research teams were responsible for completing designated preparatory work to ensure that meetings were as productive as possible (see Figure 2). The research co-creation was a critical aspect of the application process to collaboratively transform initial research concepts into refined impact evaluation research designs.

Figure 2: Overview of Research Co-Creation Meeting Objectives



In the third and final phase, applicants who completed the research co-creation phase were invited to submit a full application for research funding. The inputs for this application were entirely developed during the research co-creation phase. In particular, the research co-creation process enabled implementers to overcome four major constraints that typically limit utilization of impact evaluations in humanitarian contexts. These include reputational risks for implementers, limited proficiency in impact evaluations, ineffective research partnerships, and research implementation challenges. The four constraints and how the research co-creation process was designed to navigate these are discussed below.

Figure 3: Overview of Research Co-Creation Phases



REPUTATIONAL RISKS FOR IMPLEMENTERS



CONSTRAINT #1

Constraint #1: Reputational Risk for Implementers

One of the major constraints for implementers wanting to conduct an impact evaluation is the perceived reputational risk. Implementers fear that poor evaluation results could imply poor performance, and this could undermine their authority and call their previous work into question. This perception of reduced credibility may jeopardize their future funding prospects. **From the implementer's perspective, impact evaluations are often viewed as an accountability mechanism to gauge their performance. This masks the most significant contributions that impact evaluations can make, which is the opportunity for implementers to identify which program modalities have yielded the greatest beneficial impacts.**

Addressing this constraint requires that implementers view impact evaluations as tools that generate learning to best optimize programs to achieve the greatest impact. One way that HAEC has shifted this mindset was through continued advocacy on the use of impact evaluations to answer targeted operational research questions. Operational research questions are typically the questions that provide the most learning value for implementers. These questions often include analysis of the effectiveness or cost-effectiveness of different program modalities (e.g., different combinations of activities, different delivery mechanisms, different timing of assistance).

To uncover what research questions implementers were truly grappling with, HAEC designed the initial sessions of the research co-creation process to understand the implementers' key learning objectives and whether they could be achieved through an impact evaluation. A key part of these discussions was thought exercises facilitated by HAEC on how evidence from the research question(s) would be utilized with a particular focus on the application of learnings for the activity, country office, organization, and/or BHA. If there was no clear learning application, HAEC would encourage the teams to consider alternative research questions or discuss internally if an impact evaluation was the right tool.

During this early stage, HAEC underscored the potential for using A/B testing approaches - or comparing different versions of humanitarian assistance programs against each other – to answer questions that were most valuable to implementers. Reframing impact evaluations in this way increased implementer buy-in for impact evaluations and ensured that the evidence generated could be directly applied to future humanitarian activities (see Box 1).

Evaluation Spotlight: Mercy Corps BEGE

HAEC supported Mercy Corps to design an impact evaluation for its Broadening Emergency and Graduation Effort (BEGE) project in Nigeria's Borno State. This activity delivers lifesaving and sustainable interventions with the goal of enhancing the adaptability of displaced, returned, and host households to shocks or systemic constraints under the highly uncertain conditions in Borno State. Households that are part of the program receive a combination of food assistance, cash assistance, technical training on agricultural livestock and production practices, animal feed transfers, and other interventions such as supplementary nutritional assistance or clean and fuel-efficient stoves.

Continue on Next Page

Evaluation Spotlight: Mercy Corps BEGE Continued

Mercy Corps wanted to understand how to most effectively transition households from humanitarian assistance into early recovery. Based on experience delivering similar programming in Iraq and Niger, they hypothesized that incorporating an additional “life skills” training that teaches participants financial health best practices would lead to improvements in individuals’ economic and psychosocial wellbeing outcomes. Specifically, they wanted to measure changes in their self-confidence, locus of control, and confidence in the future, all of which have been found to improve households’ food security and resilience in times of crises. With their research partner, Causal Design, Mercy Corps designed an impact evaluation to answer whether they should incorporate the life skills training in future programming, including the next phase of the BEGE program which plans to provide 1,830 households with food aid, before transitioning them to early recovery assistance.

The team utilized an A/B testing approach to answer whether the life skills training component would improve participant outcomes relative to the package of interventions without it. The team designed a clustered randomized control trial where the intervention package with the life skills training component was randomly assigned to a subset of BEGE-supported communities. The research measures outcomes of participants in both sets of communities to assess the marginal impact of the life skills training. An important aspect of the decision to scale this intervention package is assessing the life skills training’s cost-effectiveness; therefore, the study also included a cost effectiveness analysis to compare the relative effects of the two program packages in relation to their relative costs.

Throughout the research co-creation process, HAEC encouraged Mercy Corps to think about how they would use the evidence and to shape the impact evaluation to ensure that the results provided value to Mercy Corps. Through these conversations, the team identified specific avenues for evidence utilization that would be useful for them including humanitarian assistance and development practitioners and government agencies within Nigeria and the west Africa region more broadly. Despite the broad scope of the BEGE project, the Mercy Corps research co-creation case demonstrates that tailoring the impact evaluation towards answering important and relevant operational research questions provides important information for humanitarian program design and implementer learning.



Photo Credit: Etinosa Yvonne/Save The Children

LIMITED PROFICIENCY IN IMPACT EVALUATIONS



CONSTRAINT #2

Constraint #2: Limited Proficiency in Impact Evaluations

Wide-ranging misconceptions around what an impact evaluation is and how it differs from other evaluation methodologies also diminishes demand for impact evaluations. In particular, many implementers and donors incorrectly define impact evaluations based on the types of outcomes measured, not by the type of methodology used. Commonly, performance evaluations (specifically, outcome evaluations which measure changes in outcomes through a baseline and endline comparison) are mistakenly interpreted to measure impact. Since performance evaluations do not involve comparison against a counterfactual, they do not provide sufficient evidence to demonstrate impact, and consequently can lead to incorrect conclusions around the effectiveness of humanitarian programming.

Building proficiency of what an impact evaluation is and the specific questions it can answer are key inputs in helping implementers determine if it is the right evaluation tool for them. The research co-creation process provided capacity strengthening around impact evaluations for implementers through resource sharing and open dialogue. For example, HAEC asked implementers to review key resources ahead of research co-creation meetings to guide their internal discussion as they designed their evaluations. For instance, HAEC published and shared this [cheat sheet](#) that outlines common threats to impact evaluation design with examples of how these can creep into designs and introduce bias. This resource helped prepare implementers to identify possible risks in their contexts. Complementing these resources, HAEC also provided consultative advice during the research co-creation sessions around impact evaluation design options and provided technical guidance and feedback on the rigor of different methods (see Box 2).

Evaluation Spotlight: ADRA Integrated Actions for Sustainable Food Security (TRANSFER) Project

HAEC supported the Adventist Development and Relief Agency (ADRA) in designing an impact evaluation for its Integrated Actions for Sustainable Food Security (TRANSFER) activity in Honduras. This activity aims to improve the food security of vulnerable households that reside in the southern portion of the country's dry corridor which has a high susceptibility to shocks arising from adverse weather events including hurricanes and drought. The activity delivers agricultural interventions (i.e., seed transfers) in conjunction with unconditional cash transfers.

The primary activity examined under the impact evaluation is the reintroduction of sorghum cultivation, a crop that possesses greater drought resilience, with the goal of providing a basic grain for consumption even in the face of drought. ADRA was interested in understanding how reintroducing sorghum would affect annual household sorghum production and human consumption practices. The challenge with this question was how to find a rigorous comparison group given ADRA's activity context. Given that sorghum seeds were distributed based on specific eligibility criteria such as land ownership and land area size and past food security, simply comparing those that did not versus did receive sorghum would generate a biased estimate of the impact of introducing sorghum.

Continue on Next Page

Evaluation Spotlight: ADRA Integrated Actions for Sustainable Food Security (TRANSFER) Project Continued

Through the research co-creation process, HAEC, together with ADRA and their research partner i-APS, identified potential ways to construct a rigorous comparison group to answer this question. In one session, ADRA shared that not all households who were eligible to receive sorghum ended up with seeds because there was a fixed amount of seeds to distribute. This illuminated an opportunity for a potential comparison group of households, although comparing eligible households that received to eligible households that did not receive does not take into account that some households turned down the sorghum.

Through the research co-creation process, the HAEC team highlighted the importance of identifying households that would have been interested in sorghum had it been available. Specifically, differentiating this group from farmers that did not receive sorghum because they chose to not receive sorghum was important because outcomes from these two groups would likely be different, independent of the sorghum. As such, ADRA and i-APS developed a plan for identifying these households in the endline survey to generate a comparison group that excluded these households. They then planned to use statistical matching techniques to ascertain whether the sorghum introduction leads to better food security outcomes.



Photo Credit: Adriana Loureiro Fernandez/Save The Children

INEFFECTIVE RESEARCH PARTNERSHIPS



CONSTRAINT #3

Constraint # 3 – Ineffective Research Partnerships

Another constraint to the uptake of impact evaluations, raised by implementers, was ineffective research partnerships. In other words, implementers did not find that the final research products generated relevant learning for them and their questions. There were several drivers of this constraint. First, implementers routinely underscored the need to remain flexible with methodologies, such as utilizing quasi-experimental methods in contexts especially where randomization is more challenging. Additionally, there can be misalignment between which research questions to prioritize. Implementers usually prioritize highly specified research questions that inform their programmatic decisions. However, these questions may not be as compelling for research partners whose primary focus is publication. Finally, not all research partners can adapt and pivot quickly in the face of changing circumstances (which are particularly common in humanitarian contexts) such as implementation delays or unexpected shocks that may necessitate changes in the research design or timeline. These misaligned incentives can undermine research collaborations and lead to research products that are not useful for implementers.

To navigate this constraint, the HAEC team facilitated early and proactive conversations between research partners and implementers. Creating a space for open and honest dialogue early was essential to establishing successful partnerships. These initial discussions between implementers and researchers were valuable opportunities to build rapport and establish a deep understanding of the motivations driving both parties to carry out an impact evaluation, as well as the context in which the evaluation would be situated. Having these discussions at the earliest stages allows researchers to understand the core operational research questions that implementers have. Furthermore, researchers can develop a better understanding of the specific context under which the impact evaluation would be carried out, providing an awareness of operational or resource constraints to both more effectively design an impact evaluation and collaborate during the partnership (see Box 3).

Evaluation Spotlight: Blumont's Closing Gaps Project

HAEC supported Blumont in the creation of an impact evaluation design for its Closing Gaps activity in Colombia. This activity provides support to recently displaced households through unconditional food assistance, WASH and shelter non-food items, psychosocial support and connections to government social protection programs.

From a practical perspective, the initial interactions brought about by the research co-creation process established points of contact early on between Blumont and their research partner, Causal Design, streamlining the sharing of information, establishing norms over interactions, and improving the overall collaboration process. The initial contact facilitated by the research co-creation process allowed for the research team to develop a deep understanding of the programmatic context that was shaping the desired set of objectives that Blumont wanted to achieve with the impact evaluation.

Continue on Next Page

Evaluation Spotlight: Blumont's Closing Gaps Project

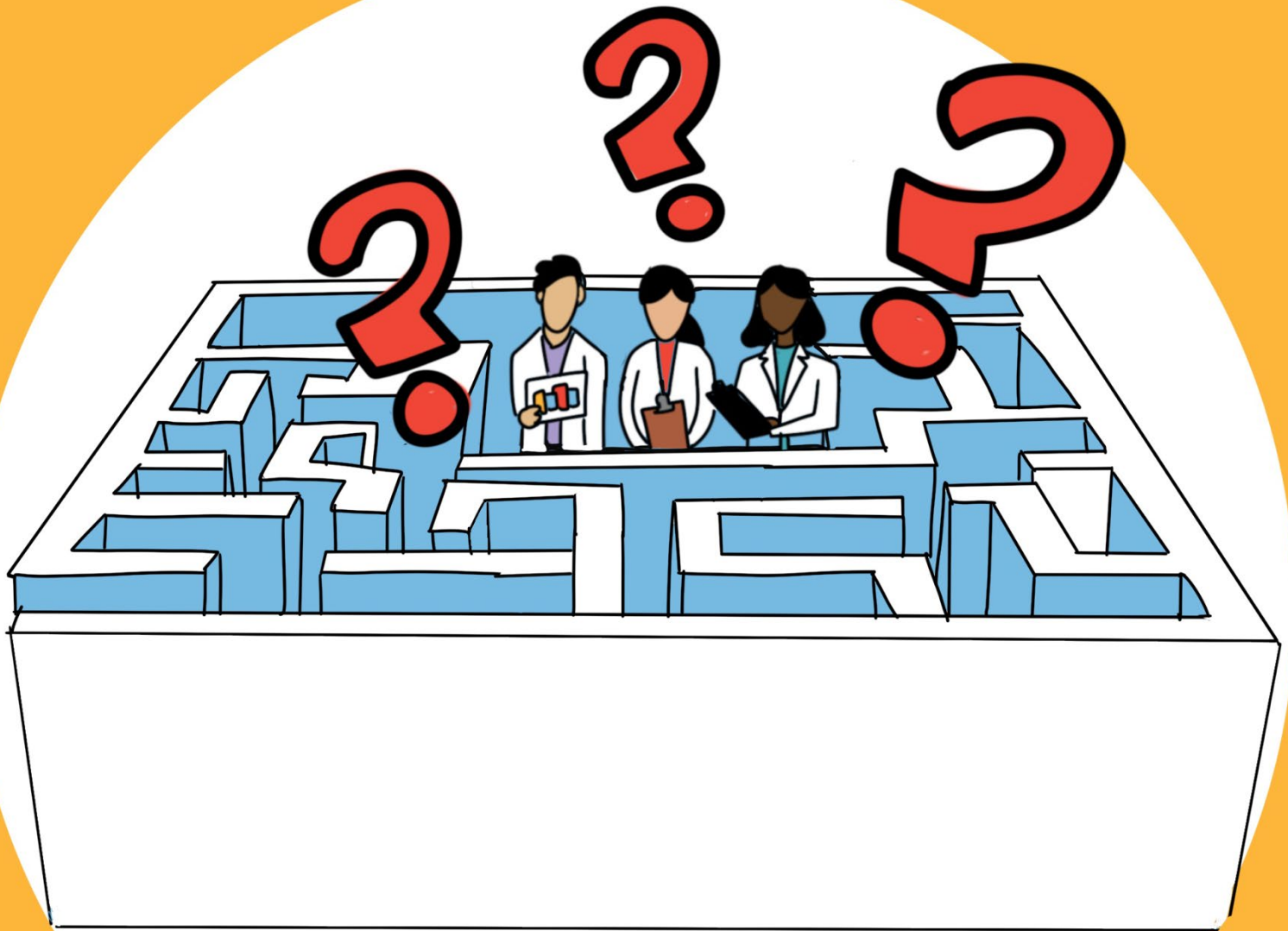
Moving past the initial interaction stages, the involvement the research co-creation required of both parties created synergies that have continued to be beneficial into the implementation of the impact evaluation. The programmatic and contextual knowledge the research team gained during the research co-creation of the impact evaluation design informed their own resourcing decisions, allowing them to tailor their interactions with Blumont towards the more difficult aspects in the implementation of the evaluation design.

The experience with Blumont highlights the important role that the research co-creation process can play as a platform to facilitate the research partner's understanding of the implementer's specific needs, constraints, and operational context. The organization and structure of the research co-creation process are conducive to productive discussions between researchers and implementers. This fosters a mutual understanding between both parties. Throughout the research co-creation process, the research partner builds an intimate understanding of the underlying motivations, needs, and constraints of the implementer.



Photo Credit: Nadège Mazars/Save The Children

RESEARCH IMPLEMENTATION CHALLENGES



CONSTRAINT #4

Constraint #4: Research Implementation Challenges

Conducting an impact evaluation under humanitarian contexts presents unique challenges for research design and implementation. The dynamics that make humanitarian contexts difficult to work in are the same dynamics that make research difficult to conduct. For example, the security concerns that often arise in humanitarian settings create challenging or even dangerous environments for enumerators, making data collection difficult or impossible. Population movement is also common in humanitarian contexts as conflict or environmental crises often lead to population displacement. These challenges can threaten the integrity of research design or shut down an impact evaluation entirely.

While these types of research implementation risks may be unavoidable, the research co-creation process provided a valuable opportunity for both implementer and research partner teams to identify and plan for potential risks before the impact evaluation began. Once the research design and subsequent implementation plan were developed, the latter stages of the research co-creation process were structured around the management of potential research risks. This process allowed researchers and implementers to identify a preliminary set of risks, outline the implications for the research, develop contingency plans and formalize multiple decision points such that the decision to pursue, delay or terminate the evaluation was based upon the latest on the ground developments (see Box 4).

Evaluation Spotlight: Danish Refugee Council in Niger

In October 2022, Danish Refugee Council (DRC) began implementing a two-year project aimed at providing emergency and early recovery response services for host protracted displaced populations in the Diffa and Tillaberi regions in Niger. The activity provides multi-sectoral emergency services and bolsters household early recovery response capacity through emergency food security assistance during the lean seasons via the distribution of food vouchers. The initial research question of interest for DRC was focused on identifying whether there are additional benefits to their packaged food assistance by adding a shelter protection program compared to just a food assistance program. This research design involved comparing outcomes of households receiving food assistance delivered by the World Food Programme (WFP) to outcomes of households receiving food assistance and shelter delivered by DRC. However, shortly after the research design was finalized, the WFP announced that its food assistance would be delayed.

DRC and their research partner, Dr. Christine Moser from Western Michigan University, worked with HAEC to preemptively lay out a decision timeline. The team identified milestones for when food assistance would need to resume to continue with the research. It eventually became clear that the delay would not be resolved quickly, and the initial research question would no longer be viable because the timing of the assistance would not be the same for the two groups. Luckily, the research team recognized that the delay itself was an important policy question that could be studied in light of these new circumstances. As such, the team pivoted the research design to study the impact of delayed food assistance relative to on-time food assistance and was still able to leverage the baseline targeting data that DRC had previously collected. The regular touchpoints provided through the research co-creation process enabled HAEC (the funder), DRC (the IP) and Dr. Moser (the research partner) to stay informed with developments on the ground. This enabled effective scenario planning that resulted in the adaptation of the research design to answer a key learning question for DRC and the broader humanitarian community.

Recommendations for Moving Forward

Rising humanitarian needs and persistent funding gaps highlight the importance of ensuring that humanitarian response is as impactful and cost-effective as possible. Yet, impact evaluations remain an underutilized evaluation tool in the humanitarian assistance sector. However, HAEC's own experience procuring research highlights that the demand for impact evaluations is there, but how impact evaluations are designed and research partnerships are formed really matters for generating successful collaborations and valuable learning products for implementers. Below we highlight how these principles can be integrated into existing procurement processes to better set up research teams for success and increase demand for impact evaluations in humanitarian settings.

1. **Build research co-creation into a Statement of Work (SOW)** One way to emulate HAEC's research co-creation process is to incorporate a similar activity in a Statement of Work, which is issued by implementers to procure research. For example, an implementer could procure research in phases where the first phase is a research co-creation phase. In this phase, the implementer and research partner could work together to refine the research questions to ensure they generate evidence that is useful and actionable for the implementer and develop a viable research design. In this process, the implementer and research partner could also collectively identify research design and implementation risks and develop a mitigation strategy. If later it was determined that the research design was not viable or a challenge arose that prohibited the research implementation, the second phase of executing the research would not have to be funded. However, this phasing within an SOW may not always be possible, so the subsequent recommendations outline lighter touch approaches for integrating the lessons learned from HAEC's co-creation process.
2. **Reflect on evidence utilization when developing research questions.** Impact evaluations provide the most value to implementers when they generate evidence that can inform specific programmatic decisions. To identify the most useful questions, *implementers should reflect on what concrete steps they would take with the findings.* The experience from the research co-creation process has demonstrated that facilitating discussions or thought exercises considering how evidence would be used can effectively guide implementers towards the right questions. This may take the form of brainstorm sessions with various levels of implementer staff including leadership down to program teams. In particular, HAEC advocates to reflect on the programming questions that "keep program teams up at night" – these questions often shed light on the most pressing operational queries that might be best answered by causal research. These reflections can surface the most useful questions for implementers. Desired evidence may also address diverse needs – some more targeted and some broader in scope – such as more operational adaptive management for ongoing programming, informing the design of a future iteration of a program, or contributing to broader donor policy.
3. **Seek out research partners that have aligned learning goals.** When implementers and research partners have different priorities, this often will not generate valuable learning products for implementers. It is important to identify research partners that are implementer-focused when it comes to their learning priorities and have a clear willingness to be flexible with the research design and questions as research and programming begin. This alignment check could happen during an initial

research co-creation process but could also happen during early conversations in the research procurement stage. This enables the impact evaluation design to be tailored specifically towards answering the research questions that matter most to the implementers.

4. ***Involve research teams early during the impact evaluation design phase.*** Bringing research teams in during the preliminary design stages creates the opportunity for implementers and researcher entities to build rapport and an understanding of their organizational systems and constraints. Bringing in research teams early in the design phase allows them to understand the program implementation and contextual contexts that will shape the impact evaluation methodology and execution. This also opens the opportunity for them to be involved in refining research questions as research partners are best equipped to assess viability of generating causal evidence and whether an approach can generate the insights implementers are looking for. The synergies created by these early interactions apply not only during the design stages, but the implementation of the impact evaluation as well. Early involvement can allow research teams to anticipate which phases of the impact evaluation require intensive effort, allowing them to better balance their commitments.

HAEC's experience highlights that the design of impact evaluations and the formation of research partnerships are critical factors in their success. While implementing HAEC's same research co-creation process may not be feasible, incorporating these lessons learned can help alleviate some of the major constraints to impact evaluation utilization in the humanitarian sector. By more intentionally selecting research questions and forming research partnerships, impact evaluations will more likely provide value to implementers in addressing these barriers. Enhancing the production and utilization of rigorous evidence can improve the effectiveness and efficiency of humanitarian assistance, to better support those affected by conflict and crises.

