

## WHAT IMPACT?

On whether evidence is really needed and what drives us to get it

By Fiona Kelling with input from Simone van Dijk



Research as a buzzword: are current studies seeking information that can influence our decisions or just following global trends?

Having researched and reflected on the (lack of) evidence on the impacts of shelter and settlements assistance and what should be done about it, I have come to the realization that, in the current conditions within the humanitarian sector, measuring impact is unnecessary. Simply put, if continued funding or programmatic decisions really depended on it, we would have done more of it by now.

I am, of course, being intentionally provocative. By claiming that impact evaluation is ‘unnecessary’, I mean to say that shelter and settlements practitioners are not using or relying on the evidence that emerges from impact studies in order to keep on doing what they are doing. I am not saying that more impact evaluation is not needed, or that we shouldn’t be seeking better evidence to improve what we do. What I am saying is that this need and utility will only be realized when we start to address why we haven’t.

### THE EVIDENCE GAP: NEED VERSUS VALUE

Increasing accountability, coordination and standards have undoubtedly improved both the process and provision of aid and its subsequent documentation and evaluation. Yet of the 3,512 evaluations across all sectors in ALNAP’s HELP library,<sup>1</sup> only 61 of them are classified as impact evaluations – and a mere three of them tagged with shelter.

The need to improve our evidence base was highlighted to me through the research study I carried out for InterAction exploring the wider impacts of shelter and settlements assistance.<sup>2</sup> The report provides a variety of illustrations of the ways in which shelter can have impacts on a wide range of other sectors, including physical and mental health, education, livelihoods, food security, DRR and gender. However, the research also showed the weakness of the reliability of the evidence through over 190 relevant reports and evaluations from across the humanitarian, development and housing sectors. If you are fortunate enough to have the time and tenacity to read the

1 ALNAP, Help Library, <https://www.alnap.org/help-library>

2 InterAction (2020), [More Than Four Walls and a Roof](#)



Power and accountability: whose interests does it serve to better understand the impacts of assistance?

methodology and data analysis, you would find that more than 60 per cent of the included documents did not use any kind of quasi-experimental<sup>3</sup> or controlled study, that is, which identified impact by comparing results against a counterfactual, by which we are able to identify what difference the intervention has made. What this translates to is a disturbing lack of ability to demonstrate robust results at the project or program outcome level. It is not that there aren’t any good examples, but rather that on the whole, the quality of our evaluation and reporting is woefully lacking.

This is far from the first report to conclude this.<sup>4</sup> Furthermore there is no shortage of information and advice on why generating better evidence through outcome and impact evaluation is becoming more urgent, nor guidance on what needs to happen to accomplish this – some of it over twenty years old. Alongside a serious lack of independent peer review in the sector, the availability of numerous resources<sup>5</sup> has done little to change the way in which the majority of evaluations are done. Notwithstanding our aversion to reading long reports, the bigger issue is the substantial gap between the rhetoric and reality regarding the claimed imperative to demonstrate impact.

And yet despite the apparently tolerable apathy to doing so, we would struggle to argue that knowing our impact would not be of value. So why is it that we have not done more of it? Is it not actually that useful? Are we too scared of what we might find out?

3 Quasi-experimental research design attempts to establish a cause-and-effect relationship, but where the comparative groups are not randomly assigned.

4 To choose but a few, [Guerrero et. al., 2013](#); [Watson, 2008](#); [Hoffman, 2004](#)

5 These include recognition of the challenges facing humanitarian action in particular, guidance on choosing appropriate methodologies and practical approaches, tools and quality assurance checklists. See: [Dillon, 2019](#); [ELHRA, 2019](#); [WFP, 2018](#); [Puri et. al., 2015](#); [ODI, 2010](#); [OECD, 2010](#); [Vaessen, 2010](#); [Proudlock et. al., 2009](#); [Oxfam, 2007](#); [Roche, 1999](#) amongst many others

### ACCOUNTING FOR THE LACK OF EVIDENCE

Perhaps there is just little incentive, because when it comes down to it, the point of humanitarian assistance is to deliver, not to evaluate. And impact evaluation – even good outcome evaluation – is complicated. It requires specific skills and resources and time and early consideration when there are numerous other pressing concerns and funding priorities in an emergency response. There are trade-offs to be made between getting the mostly right assistance there in time, or getting it better, but too late.

The reality is that some learning may be better placed to happen through experience and sharing rather than through formal evaluations. Evaluations can be expensive – although arguably this should encourage us to gain as much from them as possible when they are carried out. But with a broad range of factors to be considered, there are limitations on how much can be addressed.

However, there is increasing recognition that organizations should create space to think about and invest in certain practices more and earlier. As a response goes on, it becomes more feasible to invest more time in data collection and analysis to be able to assess an intervention’s effects – whether that is by establishing a baseline or evaluating the context to inform a solid Theory of Change.

### ACCOUNTABILITY AS A DRIVING FORCE

Without implying that humanitarian organizations do not care about what they are providing (or who they are providing for), it is apparent that the aid system operates without natural feedback loops, as exist in most other client-producer relationships. Consumers simply don’t buy a product if they do not feel a company is providing quality goods or services. In this context the end-user makes their own decision – as opposed to with humanitarian assistance – where the end-user is ill-placed to reject help offered, even if it is not the most efficient or effective.

Even though many steps have been taken to increase downward accountability – for example, through increased participation – without consumer checks in place the onus remains on the decision-making ‘producer’ (be they donors or implementers) to ensure the responsible use of power. Measuring the real impacts of assistance, intended



Closing the feedback loop: how many organisations ask the population for verifications of findings or to explore the actions that should be taken to address recommendations?

and unintended, is part of this. This process and power-imbalance underlies why measuring impact has sadly been dispensable.

### BEING HELD TO ACCOUNT

The intrinsic nature of humanitarian assistance is such that its purpose is rarely questioned. Who can argue that having a roof over your head is not important? Nevertheless, having more reliable results could contribute to sectoral advocacy and more effective decision-making in a resource-scarce environment. Yet often we are only required to report on whether we did what we said we would, rather than on what it accomplished or how. A number of evaluations were excluded from the Wider Impacts research, as although they provided detailed descriptions of what was carried out, they did not include any analysis of what difference it made. Given the commitments made in the Grand Bargain, the strong upward accountability and influence donors have, and fundamentally, the financial resources that would enable implementation, there is a question as to why donors have not wielded their power more strongly to create incentives for humanitarian organizations to do better.

Contextual and cultural factors hinder or enhance the accomplishment of intended change, as much in the shelter sector as in any of our project evaluations (see Figure 1). Looking at the drivers and inhibitors of change in the humanitarian system<sup>6</sup> and the prospects for progress emphasizes the need to understand the motivations and incentives that might contribute towards or deter any process of change. In this regard, donors have their own ‘context’; limitations and incentives which do not always align with professed priorities.

This brings us back to both accountability and necessity. Does the fact that donors are not requiring us to improve our evidence base mean that we shouldn’t? Or does it mean that we should be requesting donors to also improve? Perhaps it is time that donors are the ones being held to account on their commitments towards the affected population.

6 GPPi (2016) *Drivers and Inhibitors of Change in the Humanitarian System*; ODI (2010), *The Humanitarian’s Dilemma: collective action or inaction in international relief?*

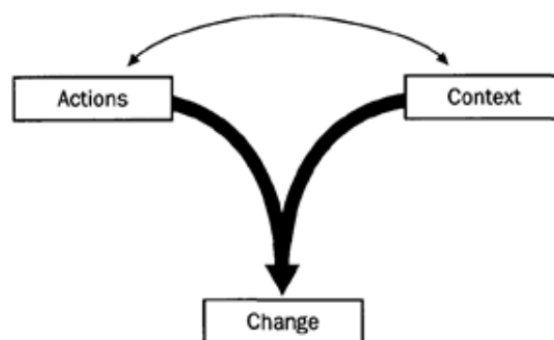


Figure 1: Actions and context combine to produce [or hinder] change (Roche, 1999: 24)

### BRIDGING THE GAP

Huge investment and improvements in M&E have been made by organizations, even if they are largely compliance-oriented rather than results-focused. For many years, M&E has been added-on rather than built-in to programming, resulting in a separation of functions where programs ‘do’ and M&E ‘measures’. As a result, technical teams can feel criticized or defensive when told they are ‘not meeting their indicators’ and ‘need to do better’. Likewise, the shelter sector can feel self-protective if research implies their assistance is not helping because robust evidence is lacking.

Instead of pitting one discipline against the other, reframing the role that M&E plays within a program or organization would recognize and result in greater inter-reliance and mutual benefit. For program staff to realize the value of M&E to their work, a people-centered approach is required that puts utility for programming at the center, instead of reporting. A rise in adaptive programming and settlement-based approaches may be heralding a change, where M&E specialists act as facilitators to conversations on what needs to be altered and why in light of ongoing data collection and analysis. Rather than adding more work or doing someone else’s job, it becomes about technical staff thinking about and being involved in data collection that is useful for them, supported by M&E specialists.

### CONCLUSION: MAKING IT EVIDENT

What is clear is that having the tools or technical guidance is not enough to transform our performance. Information needs to be relevant, timely, and succinct.<sup>7</sup> It needs to take account of and appeal to implementer’s motives and desires, whether that is improved quality, a more efficient response, or more funding.

The InterAction report warns that ‘to accept the examples collated in this report without engaging with the need to generate better evidence runs the risk of perpetuating the lack of information available and delaying the required investment in generating better data.’ Assuming we do really want to know where to best direct funding, how to maximize inter sectoral linkages, or what conditions might need to be in place for an intervention to achieve certain outcomes, we need to address the evidence gap. Investment is essential at all levels, but first we need to honestly reflect on our priorities and motivations. Only then will we have the drive to do what we need to fill it.

7 UCL (2013), [Data, decision-making and disasters](#)



*Making an impact: better evidence can support advocacy, inform decision making and ultimately improve response.*