START NETWORK

How to assess the impact of a Drought Risk Financing facility: *a guide*

SIMON LEVINE AND BILL GRAY

OCTOBER 2017

HIF humanitarian innovation fund

Contents

Summary

This guide outlines some of the conceptual questions, and practical tools, that can be used to evaluate drought risk financing (DRF) initiatives. It offers a framework for thinking about the impact of drought risk financing, along with risk financing more broadly and the wider general issue of the added value of earlier humanitarian response.

Our approach is different from other commonly used methodologies, because it looks to go beyond benefit-cost ratios and instead takes a programme theory approach that tests each assumption and each causal link individually.

We hope this will result in a richer picture of the impact of DRF, which can be used to inform learning about what works and what doesn't. The guide is directed at the practitioners who will be delivering the interventions: this will allow the collection of evidence in real time, enabling shorter feedback loops.

This guide was written for the Start Network by external consultants Simon Levine and Bill Gray, and reflects their own personal experience and opinions as experts in this field. We hope that this guide can inform thinking beyond the Start Network, and be a useful tool for the establishment of wider risk financing initiatives.

1 *In indemnity insurance, the payout is triggered after the event by the actual loss, so that the crashed car is replaced or the damaged home is repaired back to its prior state. In parametric (also called index-based) insurance, the amount of the payout is dictated by an ex ante objective measure of the causal event (eg, the levels of soil moisture indicating an emerging drought).*

Introduction

The Start Network is developing and testing new drought risk financing initiatives that aim to deploy funds to humanitarian responders (NGOs) to protect communities at risk of major droughts in a faster, more efficient way.

These initiatives are underpinned by a disaster risk financing approach that involves scientific modelling of drought risks, focused scenario-based contingency planning and pre-positioned financing (including parametric insurance[1]).

The vision for the first initiative, the **Drought Financing Facility (DFF)**, is of an NGO-led network of interconnected national facilities that enable faster and more coordinated humanitarian response to major droughts as they emerge. It is being designed in Pakistan and Zimbabwe, where groups of NGOs are working with the Start Team to tailor a funding mechanism which meets their needs.[2]

The second is the **African Risk Capacity (ARC) Replica** programme, in which the Start Network is replicating the drought insurance policies of African Union governments and expanding their coverage.[3]

Because disaster risk financing is a new and experimental approach to humanitarian action, it needs strong monitoring and feedback. It is not enough simply to assess whether it fulfils its function (pays out on time, delivers response plans etc). We also need to measure its impact for the most vulnerable communities compared to a 'traditional' reactive humanitarian response.

The approach

Our approach is different from other commonly used methodologies for quantifying impact and for calculating benefit-cost ratios (BCR). One of the main differences is the audience we have developed this framework for.

Analysis of impact and Value For Money is usually undertaken by expert consultants with specialist knowledge of the methodologies for assessing impact and calculating BCR. This guide is instead written for people who are experts in the systems and projects being assessed. These are principally the staff of the agencies who will pilot DRF initiatives, along with other interested parties such as monitoring and evaluation staff.

There are several reasons for the approach that we have taken.

Firstly, it is very hard to draw general lessons from humanitarian crises, because every crisis is different. The impacts on each individual or household are also very different: the same crisis will affect different people in very different ways, and projects play out very differently even in neighbouring villages. Methodologies which would require us to remove the context before we can learn general lessons are the least likely to generate useful learning for improving humanitarian response.

Secondly, it is very difficult for methodological experts who are not also experts in the local context to interpret the data they collect or the testimony they hear, or even to know which lessons are the most important to draw. Improving humanitarian response is too important a task for lesson learning to be subcontracted out.

2 *See https://startnetwork.org/start-labs/drought-risk-financing#DFF*
3 *See https://startnetwork.org/start-labs/drought-risk-financing#ARC*

The starting point for this guide is that all staff involved in humanitarian response can contribute to learning, and need to benefit from the lessons of experience. The role of methodological experts is to support them in this learning process, by helping ensure that their methodologies are sound and that they reach the correct conclusions: those that genuinely follow from the evidence as it was collected.

Programme theory evaluation

This guide does not provide a tool for quantifying the benefit-cost ratio of DRF. These types of calculation arrive at a price per unit measure which, while useful for advocacy purposes, cannot capture impact or help us with learning. Instead, we need to think intelligently about how DRF contributes to impact: we need to find ways to do so that avoid many of the common traps and biases and we need to ensure that whatever conclusions we draw are rigorous and credible.

One way to achieve this is to use an approach called programme theory evaluation or theory-based impact assessment.[4] This approach starts from the programme theory, or theory of change, behind any humanitarian project. Rather than making a single connection between a project and its benefits, it takes a long causal chain and breaks it down into individual links. Each of these can be tested much more simply.

THERE ARE MANY ADVANTAGES TO THIS APPROACH:

1. **Assessment is much simpler, because we are always asking ourselves smaller questions with clearer answers.**

2. **It allows us to use a variety of methodologies and tools, since each small question lends itself to being answered in very different ways.**

3. **It provides much more learning, since we will discover not only whether or not a project achieved impact, but also how it did or where the chain broke down. If agencies connected with DRF were only given an overall judgement on how DRF helped to mitigate the crisis, they would miss the vast majority of the lessons which can be drawn from experience.**

4. **It allows us to learn more about a range of areas: about coordination; about the role of risk analysis in response design; about the contribution of contingency planning to quicker response; about the different speeds at which different financing mechanisms were able to respond; about the impact of responding earlier, etc.**

5. **It focuses more on how to learn about these many practical questions than on how to scientifically measure the combined impact of different factors and to establish statistically significant conclusions about their causes.**

4 *See Funnell and Rogers, 2011 and White, 2009.*

BOX 1:
PROGRAMME THEORY EVALUATION – AN EXAMPLE

One way of rigorously testing whether or not a project had any impact is to look for evidence of causal links in the chain from planning to outcome. In this example, a project is teaching farmers how to make compost. Farmers report that as a result of the training their yields have doubled. Could this be true?

WE FIRST HYPOTHESISE THE CAUSAL CHAIN FROM THE TRAINING TO THE YIELD, WHICH MIGHT BE THE FOLLOWING:

TRAINING → INCREASED KNOWLEDGE ABOUT COMPOST PRODUCTION → INCREASED PRODUCTION OF COMPOST → INCREASED APPLICATION OF COMPOST → IMPROVEMENT IN SOIL FERTILITY → INCREASED YIELD.

WE MIGHT THEN LOOK FOR EVIDENCE, ASKING THE FOLLOWING QUESTIONS:

- *How many farmers knew how to make compost before the training?*
- *How many farmers had ever made compost before the training?*
- *How many farmers made compost after the training?*
- *How much compost did they make?*
- *Over how great an area did they apply it?*
- *On which crops did they apply it?*
- *What was the fertility of the soil where they applied it?*
- *Using an estimation of the application rates on the soil, what is the scientific evidence for the likely yield increase of the crop in question? (This stage is critical. It is very common to see highly implausible yield increases attributed to changes in agricultural practices.)*
- *Is there independent evidence of the increased yield of the crop on which they applied the compost (eg, evidence from traders of an increased supply leaving the area)?*
- *What were the yields of other crops grown by the same farmers in that year? (In other words, is there another explanation for the change in yield, such as good rainfall that year?)*
- *Were there any other changes that might have resulted in a change in yield? For example, were there changes in the use of fertilisers or other inputs? Did they use different seed that year?*
- *Were there similar changes in yield for farmers who did not receive the same training? These questions can be answered in different ways: some by interviewing project beneficiaries, others by interviewing non-beneficiary farmers; some by speaking to soil scientists, looking at market data, observation in farmers' fields, etc.*

The enquiry can stop as soon as the causal chain breaks down. For example, if we find that no more farmers made compost after the training than before, then we already know that the training project could not have caused a change in yield, and there is no need to continue any further. If we can follow the chain (or 'the project theory') all the way to impact, and if we have good evidence – beyond the simple testimony of project beneficiaries – for each and every step in the chain, then our conclusion that the project was responsible for any change in yield is rigorous.

BOX 2:
THE PRINCIPLES

- **LEARNING CANNOT BE CONTRACTED OUT.**
  - *People engaged with the response or DRF are the ones who need the understanding.*
  - *People engaged with the response or DRF have the knowledge needed for the analysis.*
- **THE LEARNING PRIORITY IS FOR CONTEXT-SPECIFIC UNDERSTANDING THAT IS OF IMMEDIATE, PRACTICAL CONSEQUENCE.**
- **THE MORE 'BIG' QUESTIONS ARE BROKEN DOWN INTO THEIR COMPONENT PARTS, THE EASIER THEY ARE TO ANSWER.**
  - *Breaking down these big questions is a practical learning experience in itself, making explicit many of the hidden assumptions that are often made.*
  - *Finding the small questions helps to demystify research, making it possible for everyone to see where they can contribute – and where they can learn.*
- **THE OBJECTIVE IS NOT TO 'FIND THE ANSWER', BUT THE INCREMENTAL GAIN OF UNDERSTANDING.**
- **EVERY POSSIBLE ASSUMPTION SHOULD BE TESTED, NEVER ACCEPTED.**
  - *Revealing assumptions can be the most important part of learning. The more assumptions that can be laid bare, the better.*
  - *Even the most common-sense and intuitive assumptions may turn out not (always) to be true.*
  - *It is rarely useful to ask 'is this assumption true?'; instead, ask 'in which circumstances, and for which people, is it true?'*
- **THE HUMANITARIAN SECTOR NEEDS TO IMPROVE THE RIGOUR OF ITS EVIDENCE GATHERING AND OF THE WAYS IN WHICH IT DRAWS CONCLUSIONS FROM THE EVIDENCE.**
  - *Rigour depends more on attitude than on methodology.*
  - *Science works by trying to disprove ideas that it hopes to prove. Aid agencies tend to do the opposite. It is essential to search for every possible alternative explanation for any positive change, or any complimentary testimonies, before accepting that it might be due to your own intervention. (Always assume that you and your agency got everything wrong, until you can prove otherwise! This is not about being critical, but about scientific rigour.)*
  - *Pro-project bias is inherent in almost everything we do and everything we hear as aid agencies. It cannot be avoided, but it can be recognised and thus mitigated.*
  - *At least two different sources of evidence should be found for any assumptions. These sources must not be susceptible to the same biases. (Speaking to two project beneficiaries is NOT rigorous enough.)*
- **METHOD FOLLOWS CONTENT: THE CHOICE OF METHODOLOGIES OR TOOLS FOR INFORMATION GATHERING DEPENDS ON WHAT YOU WANT TO FIND OUT.**
  - *Where possible, mixed-methods approaches should be used, even for answering each 'small' question.*
  - *Common sense can sometimes be misleading, but can sometimes be the most powerful tool. In order to answer questions, it is often better to think like a detective or investigative reporter rather than a researcher. Ask yourself: what tools or methodologies do they use?*
- **RIGOUR AND LEARNING DEMAND TRANSPARENCY.**
  - *All data should be shared as widely as possible with others who could help the analysis.*
  - *Challenging conclusions is a way of helping, not of obstructing: this is an accepted way of working in science and in most sectors, but rarely occurs in the aid sector.*
  - *If you are nervous about sharing all of your raw data, then think why – and whether you are still confident in offering your conclusions.*
  - *All agencies and their staff should be willing to help the collaborative learning by helping – including by challenging – the work of others. Learning how to do this in a constructive way should be a core competency.*

Measuring impact

It is tempting to look for a tool that prescribes clearly which data we should collect and how to plug it into a calculation that will then generate for us a simple answer or number. However, there are both practical and theoretical reasons[5] why these kinds of figures would not be valid, let alone useful, in considering a humanitarian response. We hope that this guide will give agencies and their staff the confidence to proceed in a different direction, using evidence and analysis which is rigorous to produce understanding which is useful.

We face some limitations when measuring impact. We can only measure the overall impact of responses facilitated through DRF after the event. This means that, because DRF is being piloted in only a few countries (initially two to three), and because it will not be triggered every year, there will be very few opportunities to assess its impact even in a time frame of three to five years. One or two uses of DRF are not enough to draw any meaningful general conclusions.

However, we believe that we all need to gain much more experience and at a much faster pace. We propose an approach to learning about the impact of DRF by breaking it down into its component parts: risk pooling, insurance, funding mechanisms, risk profiling, contingency planning, early response, etc. This way, we can learn many lessons about DRF's potential contribution to impact and value for money, even in years or countries where it was not actually triggered.

The importance of collaboration

To build up a large enough body of evidence in a short time, we will need the many agencies linked to DRF to collaborate. The framework proposed in this guide is designed to facilitate this collaborative learning. It offers a way for different agencies (and different staff members within any one agency) to come to a shared understanding on how lessons can be drawn from experience. This is based on a shared understanding of how information is collected as well as a shared understanding of how that information is to be analysed, ie, what it all means.

The overall goal – understanding how DRF contributes both to mitigating suffering and to value for money – is probably too great for any one individual or agency to achieve on their own. So we also provide a framework into which different individuals and agencies can contribute smaller pieces of learning: knowledge about individual links in the long and branching causal chains between DRF and reduced suffering. This framework ensures that over time these individual pieces can be combined to reveal a larger picture.

5 *It is beyond the scope of this document to make these arguments.*

The framework for measuring impact

Because there are different DRF models in use, we use the example of the Start Network's Drought Financing Facility to illustrate the process. Some modifications may be needed when examining different forms of risk financing.

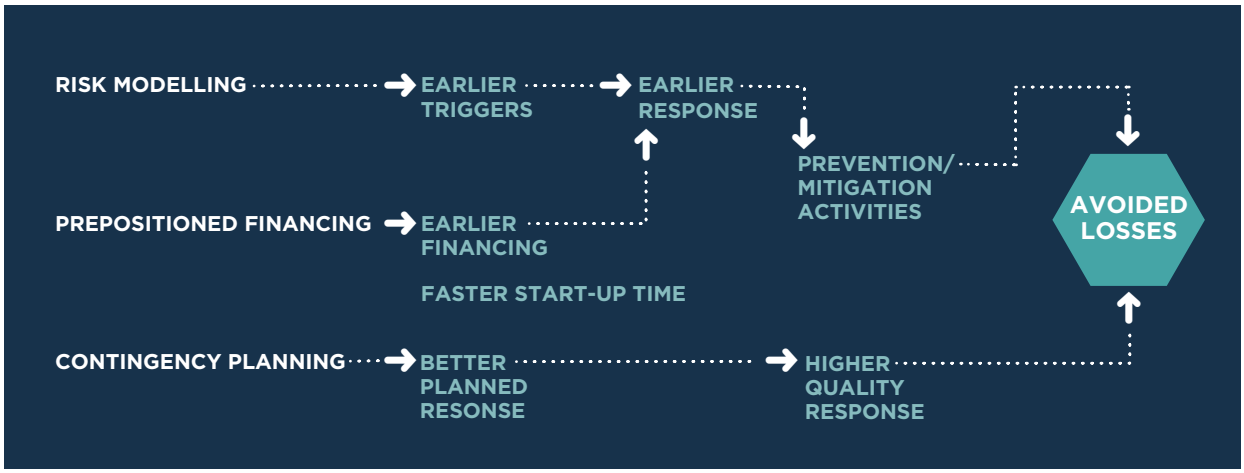
It is beyond the scope of this guide to analyse in detail how to investigate all the links that lead from risk modelling and pre-positioned financing to the outcomes of projects financed by DRF. However, we have mapped out the overall conceptual framework, showing how it can be broken down into component parts that need to be tested.

THE MEASUREMENT FRAMEWORK IS BASED ON THE FOLLOWING QUESTIONS:

1. How does predictable funding influence financial flows and response?
2. Was the science-based risk modelling accurate?
3. Did DRF lead to earlier response?
4. Did DRF facilitate more effective response?
5. Did DRF result in losses being avoided?
6. Was DRF cost-effective?

These question are considered in detail in the following sections.

FIGURE 1: THEORY BASED IMPACT ASSESSMENT APPROACH



Question 1:

How does predictable funding influence financial flows and response?

The DRF model can be thought of in two parts. Firstly, having a predictable mechanism for financing response means that agencies and communities know they can invest in planning. This planning process itself improves understanding of how to respond when needed and allows a faster response.

Secondly, DRF achieves this predictable financing by creating a specific structure for generating funds. In the present example, this is based on several national contingency funds coming together to pool risk, and using insurance to provide additional protection if the scale of response needed goes beyond their joint financing capability.

Although they function together, these two parts are in many ways separate. Assessing the funding side of DRF is a very different kind of learning exercise from looking at the response side, and requires a different set of skills and tools.

There are two broad sets of questions: is the mechanism well designed? And did it function effectively? Assessing how well the bureaucracy functioned in any one country is fairly straightforward, provided that it is very clear just how it was supposed to function. Questions about whether the funds were released in the amount that was anticipated and to the planned timetable are included in the section *Did DRF lead to earlier response?*

Questions about the design of the funding mechanism require an understanding of economics and probability.[6] As this is a technically specialist area, it is beyond the scope of this guide to cover and will probably be beyond the competence of most NGO country teams to assess. In this case, it is likely that the Start Network would need to hire an expert consultant.

6 *Anyone interested in this area should read Clark and Hill, 2013.*

Question 2:

Was the science-based risk modelling accurate?

DRF is based on an underlying assumption that scientific modelling can identify triggers that accurately predict crises. How far is this faith justified? Few households rely on household level subsistence food production (most depending to a significant degree on the market and cash income) so when households suffer acute difficulties in accessing food, how well do parametric insurance triggers capture the causes?

TO UNDERSTAND BETTER THE BENEFITS AND LIMITATIONS OF SCIENCE-BASED MODELLING, WE MUST ASK THE FOLLOWING QUESTIONS:

- **Were there crises which did not result in DRF triggers firing?**
- **Did the triggers fire and yet humanitarian crises not result?**
- **How accurate were the predictions about who would be in crisis (geographically, demographically, economically)?**
- **Are there any broader consequences of using science-based risk modelling to identify where people are vulnerable? Does this approach aid objectivity – or does it depoliticise crises, taking attention away from the structural causes of vulnerability by focusing on natural hazards?**

Question 3:

Did DRF lead to earlier response?

This is one of the main objectives of DRF and something that Start Network members will want to focus on. This isn't just to prove that DRF was quicker. It is also to understand what helped make things quicker and what could be modified to make things even faster.

THERE ARE TWO QUESTIONS.

- **Did DRF result in *earlier* response (ie, earlier than it would have been without it)?**
- **Did DRF deliver *early enough* response?**

There is an important difference between these two questions, which is not always appreciated.

The two questions should be approached quite differently. The first question is answered by looking internally at how DRF and the partner agencies worked; the second question is answered by looking at how well DRF functioned in matching the developing crisis on the ground.

Did DRF result in earlier response? (the internal question)

At its simplest, asking whether DRF brought about an earlier response is a straightforward quantitative question. We look at the date at which projects funded by DRF reached the ground, and we compare this to the date at which aid from non-DRF funded projects did so. This gives us a simple answer in days, weeks or months.

However, we can't take this approach too simply: aid from different projects reaches the ground at many different times, and there are many reasons (including some good ones) why some aid can be given earlier or later. Besides, a simple answer from counting the days between two responses won't tell us anything very useful.

The difference in delivery dates between DRF aid and non-DRF aid may not necessarily be due to the funding mechanism. Even if it is, nothing will be learned about why DRF was quicker (or slower) just by counting days. We want to learn in which circumstances we can expect DRF to be much quicker, which elements of DRF we would recommend others to replicate and how DRF as a whole could be improved. We also want to know whether the existence of DRF helped make everyone else quicker. (If DRF helped other donors to be quicker, which is one of its hoped-for impacts, then this will be a very significant positive impact. However, measuring the difference in dates in too simple a way would appear to show that the benefit of DRF was reduced.)

In order for our learning to be more useful, we need to maintain the overall idea that we can look quantitatively at speed of response (ie, counting days on a calendar), but we do this by comparing how DRF is supposed to work with how it worked in practice.

The theory behind DRF

Again, we are using the Start Network's DFF as an example of drought risk financing, with the understanding that there may be some differences in the way that other risk financing works. The DFF has three components which, between them, are intended to bring about an earlier response (see Figure 2). These are: science-based risk modelling (SRM), pre-positioned finance (PPF) and contingency planning (CP).

FIGURE 2: OUTLINE MODEL OF HOW DFF IMPROVES DROUGHT RESPONSE

One of the strengths of the DFF is that it combines, and finds synergies between, these three elements. If we map out all these synergies, along with the various processes and intermediate steps aimed at improving response, the resulting conceptual model may appear complicated (Figure 3), but the logical flow is relatively simple to understand.

Although Figure 3 appears more complicated than Figure 2, finding the answers to the question 'Is the response earlier and why?' becomes simpler. This is because Figure 3 breaks down the process from cause to effect into a number of steps, each one of which can be tested. This helps us to know exactly what to examine and makes it much simpler to examine each individual causal link.

The synergy behind the three underlying components of the DFF

01

Science-based risk modelling allows us to set triggers (or indicators) when we already know the causes of crisis. These can trigger the release of funds far earlier than by relying on so-called 'humanitarian indicators', which only become evident when a crisis is already developed.

02

Science-based risk modelling also informs **contingency planning**. By understanding when, how and for whom the crisis develops, we can design response plans that take advantage of earlier windows of opportunity to prevent or mitigate the impacts of drought.

03

Pre-positioned finance gives agencies the confidence to invest in contingency planning and response plans that address these early windows of opportunity, allowing us to take action before humanitarian indicators reach levels that set off the normal alarm bells.

Apart from enabling a different kind of response process, based on contingency planning, DRF should also speed response by improving bureaucracy.

In the case of the DFF, the combination of science-based risk modelling and pre-positioned finance means that funds are released on automatic triggers (eg, soil moisture). This avoids the need for repeat assessments and 'political' arguments, meaning that funds should be released much more quickly.

Because pre-positioned finance is predictable, science-based risk modelling allows us to carry out a contingency planning process which can go beyond general response planning.

This should give agencies both the knowledge and the confidence to invest in specific response plans and even to collectively agreed project proposals in advance of any developing situation. This kind of investment in preparedness should reduce bureaucratic delays still further, speeding up response times and facilitating earlier windows of opportunity.

This quick release of funds, combined with early submission of proposals, should mean that operational agencies receive funding earlier. Contingency planning can speed up response times by several weeks at least. The usual start-up time from receiving funds to aid reaching the ground can be as long as three to five months,[7] though this is rarely, if ever, assessed in humanitarian evaluations. However, most of the tasks undertaken during this period could be done in advance, if agencies invested in preparedness.

The shorter start-up time, along with earlier receipt of funds, should result in an earlier response to need.

How to assess DRF speed of response

By teasing out all the different and interlinked ways that DRF can make response earlier (and better), it becomes much clearer what we need to test in an assessment. We are no longer faced with one big question, such as 'In what ways did DRF help speed up response?', where we might struggle to know where to begin to find an answer. Instead we have several simpler questions, each of which presents a much clearer entry-point. For example (as shown in Figure 3), we will want to look at the bureaucratic speed of releasing funds once the triggers were reached; how early the parametric insurance triggers were reached; how far contingency planning was used to meet early windows of opportunity; how far contingency and response planning were used to improve preparedness, and so on.

Any assessment should try to understand what happened (and why) for each process and causal link, and also to assess quantitatively how much quicker DRF was than other responses. The fundamental tool for assessing speed is nothing more complicated than a calendar. This can best be done by creating a retrospective Gantt chart.

A Gantt chart is usually forward looking, a planning tool: it is a simple bar chart with the calendar along one axis. It breaks down a project or other process into different tasks, to improve planning and project management. For learning about the speed of DRF, we will use a Gantt chart to show the dates when each of the processes took place that combine to produce aid on the ground. This will allow us to measure the speed of each process and also to identify which processes were limiting factors in early response, and which processes were the weak links in speeding up response.

See Annex 1 for an example of a Gantt chart for comparing speed of response.

7 *Levine et al, 2011.*

Making comparisons

What remains is to decide: if we are trying to find out whether the DRF is quicker, then what are we going to compare it to? And, what exactly needs to be put on the Gantt chart and which questions need to be asked?

In comparing the speed of the various DRF response processes, there is no single 'correct' choice of baseline against which to measure this. Three different comparisons are possible: each one could teach us something slightly different. However, none of the comparisons will give a measure of the impact of DRF from a simple read-off from the calendar. We have already seen an example of DRF causing other responses to be quicker, and how this masked DRF impact on a simplistic calendar reading. This does not invalidate any of the comparisons: it just means we need to be very clear about what they mean. (This is an important principle. There is no simple metric to answer 'How much quicker was response because of DRF?': in making any comparison, we need to use our judgement.)

Three possible baselines to compare the speed of response through DRF

01

A comparison with other agencies (not involved in DRF) during the same drought. This is the most obvious comparison or 'control'. However, different agencies have different internal mechanisms and different speeds of reaction; some might not have been trying to respond as quickly as they could, for a number of reasons. Pragmatism is also required. It might be harder to get detailed enough information from other agencies about their internal processes to make the comparison.

02

A comparison with other funding, and other projects, for DRF partners in the same drought. This will help us to see the effect of the new DRF bureaucracy/funding mechanism, but it may miss some important impacts. For example, an NGO that took part in the Start Network's DFF may find that this improved their contingency planning, risk modelling and preparedness. By participating in the scheme, the Start Network member learns how to improve their overall speed of reaction in the future.

03

A comparison with a previous drought (before DRF) for the same partners in the same country. This has the advantage that the agencies and their bureaucracies were the same. But because every drought is different (and the politics and donor reactions are also different), the comparison may not be exactly fair.

It will be useful to use more than one comparison. This depends on the resources available, which will in turn depend on what exactly we are trying to learn about most. Using one comparison in detail could be complemented by a broader overview of the other comparisons, which can serve as a check on our conclusions. Alternatively, we might spend more time on different comparisons for different parts of the process. For example, if comparison 2 is used in detail, it might be helpful to add comparison 3, focusing on the parts related to contingency planning preparedness (which would be less affected by a changing drought scenario or changing donor politics).

Whichever comparisons are made, the framework in Figure 3 can be used to guide both the information collection and its analysis. The exact questions which need to be asked, and the specific processes which need to be plotted on the Gantt chart, will vary slightly depending on the situation. DRF may work slightly differently from country to country (or from response to response), and there may be reasons why any impact assessment of DRF would want to focus more on certain mechanisms than others. Nevertheless, this is the area where this guide can most helpfully offer generic guidelines because, as we have said, the question 'Was response earlier?' is largely an internal question about the working of DRF, and not a question about what was happening in the outside world.

Annex 2 provides a list of sample questions to guide the investigation: 'Did DRF deliver earlier response?' **Much of the information is best gathered in real time:** it will be difficult and time consuming to find detailed information after the response is over. This requires an organisational commitment to support, and resource, this learning process in advance of any crisis. Piecing together a full map requires a certain amount of investigative work: sources of information will include talking to people involved, project documents, minutes from coordination meetings and other primary and secondary data.

Did DRF deliver early enough response? (the external question)

We have looked at the question of whether or not DRF helped bring about an earlier response. However, there is little point in being earlier if even an earlier response is too late: it matters little whether life-saving assistance arrives three months or three hours after the patient has died!

In many cases, there will be a window of opportunity to respond effectively. For example, if an NGO distributes seeds after the planting season has passed, the seeds cannot be planted and will most likely go to waste. However, emergency agencies rarely think in terms of deadlines for response and often continue with activities regardless of whether the period of usefulness is over.

Whether or not any individual response was early enough can only be determined by a proper evaluation of each project. While this guide is not designed for the comprehensive evaluation of humanitarian projects, the framework outlined here can help us to learn a lot. **A great deal can also be learned even in years when no DRF emergency was triggered**.

Whether or not emergency funding arrived early enough depends upon defining what early enough looks like. This will vary depending on the type of project. 'Early enough' ought to have been described and even defined (on a crisis calendar) in the contingency planning. We can assess how well and how clearly this was done by critically studying the contingency plans, along with a study of the crisis calendar, both in real time and after the fact.

DEFINING A DATE ON A CRISIS CALENDAR

A calendar date cannot be predicted long in advance for when a response would be needed in a drought. However, because drought is a slow-onset crisis, it is usually possible to identify the start and end of a window of opportunity.

For example, if it is known that many livestock die in droughts when they are too weak to be taken to market, then providing feed at way-stations may help cattle arrive to markets alive and in better condition. The drought trajectory and this kind of response plan will be described in the science-based risk modelling. This should include an estimate of the time it takes for animals to become too weak to sell, starting from when pasture runs out. A crisis calendar prediction should be drawn up, indicating roughly when the agency expects pasture to dry up, cattle to become weak, marketing to become difficult, etc. The prediction of when this date will arrive can be progressively modified as it is seen how the drought develops.

The window of opportunity for this intervention shuts when livestock are no longer in a marketable condition at all, even with support along the way. The time when this happens can also be predicted, and can be modified by regularly monitoring the drought trajectory.

If a logic and date for being early enough was not described in the response plan, we will have to fill in the blanks. This will need a good understanding of the crisis calendar and of the technical area involved (eg, livelihoods, markets in crises, livestock). The simplest questions to answer are

- **What would have made this intervention too late to be useful?**
- **When was it reasonable to predict that this would occur (ie, before the drought, without hindsight)?**

We can then look at when the intervention was carried out, and compare this with two dates: when the window of opportunity was expected to happen, and when it actually happened. (We will need to distinguish between what was sensible to predict beforehand, and what actually happened. If a drought developed in an unusual and unpredictable way, and as a result windows of opportunity were missed, it would be unfair to conclude that the planning process had failed, even if it did not achieve any positive impact.) As with the other questions, a simple yes/no response will not teach us as much as an analysis of what helped or prevented the response from being on time

Further questions exploring the value of being early are discussed in more detail in the section *Did DRF result in losses being avoided?*

AT THE END OF THIS ENQUIRY PROCESS, WE WILL HAVE:

- **A quantified answer for the difference in speed between DRF and other humanitarian response.** is unlikely to be one number, because different responses happen at different times. The number will be broken down into various steps.
- **An understanding of why the different responses were triggered at different times.** This will help us learn about any delays in the main response as well as understanding the speed of DRF better.
- **Evidence of the advantages that DRF had over the main response.** If a facility such as DFF was quicker, then how exactly? What can be copied by other agencies who are not part of the DFF partnership? What elements of the DFF could usefully be replicated? (Was it looking at much earlier signals? Was it more 'objective', because it used parametric insurance triggers? Was it earlier because it spoke with one voice and so avoided long arguments?)

We might also have some lessons for improving DRF, or lessons for the rest of the humanitarian system on how to be faster.

However, we will have not learned anything about whether or not being earlier made any difference to people affected by the drought, and whether or not the speed of DRF resulted in appropriate responses.

> Comparisons need to be fair. The humanitarian community, including the agencies involved in DRF, would not expect to use the majority of their relief funds before a crisis is well developed. There is no point in comparing the timing of a small project, designed for shorter mitigation, with a major relief response that has been scaled up because conditions became more severe.

Question 4:

Did DRF facilitate more effective response?

Many of the ALNAP/DAC evaluation criteria[8] will be relevant here, as they would to any other emergency response.

- **Was the scale of DRF response relevant to the scale of need?**
- **How good was the targeting of assistance?**
- **What was the quality of the programming? Did it respond to priority needs? Did it take advantage of opportunities? Was it market aware? Did it enhance dignity? etc.**
- **Did DRF lead to any particular kinds of programming (eg, household level transfers)? Was this a beneficial direction? Did the DRF system allow other options to be considered? Did it facilitate and encourage innovation?**
- **Did DRF improve coordination between aid agencies, government structures and other institutions?**
- **Did it help create structures for coordination and collaboration within the country? Across the countries in DRF?**
- **What are the wider impacts of any projects, beyond the immediate project area or in the longer term? (Are there impacts on markets, incentive structures, maladaptation, power balances, etc?)**

These questions are raised as issues to reflect on, but they are not the subject of this guide.

8 *Relevance/appropriateness; connectedness; coherence; coverage; efficiency; effectiveness; impact. See ALNAP 2006.*

Question 5:

Did DRF result in losses being avoided?

Earlier emergency response may be needed for two different reasons: to alleviate suffering more quickly, and to prevent suffering in the first place.

In many crises, assistance simply arrives late, after people have already suffered badly. Most readers of this guide will know of examples where they wish that assistance had begun earlier because it could have lessened suffering. This is an argument for earlier or timely response, but it does not claim that an earlier start to aid would have been qualitatively different in any way: it would have brought the same benefits, only sooner.

This first argument is quite different from the argument that if used early enough, aid can be used differently – to prevent suffering, not respond to it. This belief that early aid can transform people's own ability to cope with a crisis is part of the rationale of DRF. Most of this section addresses how to find the evidence for this argument – or not.

It is commonly argued that early emergency response helps to avoid suffering, as well as avoiding economic losses such as asset erosion which can cause progressive destitution and increasing vulnerability to the next crisis. Appropriate assistance delivered earlier in a crisis should help alleviate suffering at that time, but there is actually surprisingly little evidence that taking action earlier can have the kind of transformative effects that are often hoped for (preventing, for example, acute symptoms such as child malnutrition over the course of the crisis, or longer term asset erosion).

DRF partners have an important opportunity to contribute evidence to support or challenge these beliefs. Because there is already such a strongly held belief in the value of early response, it is particularly important that any evidence is collected in a robust way, and that the conclusions are very carefully drawn. There may be a temptawd even pressure – to find evidence, for advocacy purposes, for investment in early response. It should be borne in mind that even if early response is shown to be beneficial as a general principle, it is unlikely that it will have benefits in every situation. It will be useful to learn more about how best to use it.

WHY LEARNING CAN BE UNCOMFORTABLE

The aid world has no bottom line to measure success, only our own perceptions of success. The realities of the sector require that, as far as possible, we are able to demonstrate success. Everyone wants to do something to improve response and prevent suffering. Lots of people have put time and effort into identifying a problem and finding a way to address it. Lots of people want DRF to work, for many reasons – hope, commitment to the ideas of emergency response and better aid bureaucracy, personal effort invested in a project, agency reputation. Individuals and agencies have also invested in the ideas behind DRF – contingency planning, triggers, insurance and the like. It's very likely that not everything in DRF will ever work as the theory predicted. Hopefully, it will be possible to demonstrate positive outcomes. Agencies and their staff will always be aware of these positives – and less likely to be aware of, or to want to recognise, what is not going to plan. The purpose of the learning process is to document and learn from the positives, but also to point out where things didn't work and how improvements can be made in the future. It is highly likely that there will be lot of things that worked in unexpected ways, that some assumptions did not hold true and that, as with any innovation, some things will fail. There may be messages that are unwelcome in some quarters – lessons that reinforce learning, but which can be uncomfortable.

Types of impact

Most impact assessment methodologies set out to measure a particular kind of impact. These impacts are chosen through a prior understanding of what kinds of impacts (or losses) are to be expected, or which kinds are of importance to the person who is measuring.[9] What is measured in impact assessment depends totally on the intervention being appraised: it could be anything from the monetary value of asset losses in a flood, to the percentage of people who succeed in giving up smoking or pupils' performance in an exam.

Although impact assessment related to DRF will only consider a narrow range of actions (ie, early emergency responses to drought), the impacts may be seen in a wide variety of ways. The suffering which an intervention attempts to mitigate will vary for different beneficiary families even in the same area. It is obvious, then, that there can be no generic, or standard, methodology for measuring impact or losses avoided.

To understand how well a project succeeded, we have to know what the project was trying to do. But this is not always enough. Some projects may be intended by the government or aid agency to have one kind of impact (eg, give people money to buy food to prevent malnutrition) but the recipients may use them for different purposes (eg, paying for health care or paying for a relative to reach town so they can send remittances). It would be wrong to only consider as a benefit the impacts that the aid agency was prioritising: the priorities of the recipients, how they used the cash grant, should surely count. Equally, any unintended negative impacts have to count too, whether or not the aid agency or government had anticipated them.

In a normal project evaluation we would be interested in the impact that an intervention had on people's well-being, or perhaps their longer term ability to recover. However, if we want to learn about DRF, then finding the actual impact of the intervention does *not* give us a well-rounded answer! As well as the impact of the intervention we also need to understand whether or not giving assistance *earlier* made any difference. In some cases, this will not be too different from assessing the impact of the assistance, if the DRF project was something quite distinct.

In other cases it may be more difficult. An example could be where DRF was used to make early cash grants, before a main emergency response was scaled up. It is reasonable to expect that people will be better off if we give them more money. So if we start giving people money earlier in a crisis, and continue giving it until the end of the crisis, we would expect to find more impact. However, the impact might still have been the same if the project had started later, but just given bigger cash transfers.

Many project proposals include a log frame (logical project framework) with 'outcome indicators', intended for assessing people's wellbeing after a project. For example, this could be by measuring their food consumption or asset levels. However, these indicators do not give us any basis for a comparison which would indicate the benefit of *earlier* aid. We could only make a comparison if there happen to be identical communities who suffered the drought in the same way and who received the same amounts of aid, only later in the drought.

If everyone affected by a drought had the same priorities and goals and used the same coping mechanisms, it might make sense to think of a general set of measurements that would indicate people's well-being. However even if this were the case, it would rarely be sensible to compare the well-being after a crisis of beneficiaries who received DRF with that of beneficiaries who received other assistance. It is highly unlikely that we would find any patterns – there would simply be too many other factors affecting people's well-being. The proposed overall approach of this guide – breaking large questions into small ones, examining separately the links in causal chains – offers a more fruitful alternative. We can see the benefits of earlier response directly by understanding how people used resources and opportunities at different stages of a drought.

9 *Relevance/appropriateness; connectedness; coherence; coverage; efficiency; effectiveness; impact. See ALNAP 2006. An agency may, for example, assess the impact of a cash or food distribution by measuring diet or nutritional status, rather than expenditure on health care.*

For each project and each situation we will need to consider a different set of possible impacts and different impact pathways *(see Box 1: Programme theory evaluation)*. No guide can cover every eventuality, and it is beyond the scope of this preparatory guide to develop pre-designed evaluation plans for all likely eventualities. Instead, we will demonstrate an approach that can be adapted in any circumstances we are investigating.

UNDERSTANDING WHY EARLIER AID IS BETTER

Receiving assistance earlier might be useful for different reasons. An earlier cash grant will help alleviate suffering which occurs earlier – but it may be that, had it been delayed, it would then have alleviated later suffering. Which is more important? There is no simple answer to this question. Earlier in the crisis there may be no other forms of assistance and so there may be particularly critical periods (eg, when children were malnourished before a main emergency response came on board). Sometimes, earlier assistance will allow the recipient to take action to prevent suffering, that a later transfer could not. Examples might include stockpiling food before prices rise; paying school fees to prevent children dropping out of education; financing a family member to migrate to find work and send back money; helping prevent a family from having to sell land to buy food, or from borrowing money at exploitative rates; giving them the funds to take livestock to where there was pasture. We can assess the impact of timing if we know what people could do with the assistance: we can then see how far they were successful in achieving this, and to what effect.

Using programme theory evaluation in practice

We can measure a project's impact by looking for evidence of each link in a causal chain from intervention to outcome. In practice it is likely that, for many projects, we won't have a clear theory of change documented, let alone one for giving aid earlier.[10] It is not possible to start with a planned theory-based impact assessment (TBIA) until we can clearly see the expected pathway to bringing about impact.

Where an early response was justified by a clear theory of change, this may serve as the main basis for planning the TBIA. Where no adequate theory of change was prepared or documented before the project, we will need to add another step to the assessment: to develop one retrospectively. This will result in the following three- or four-step process:

- **A rapid scoping exercise to find out the impact pathways retrospectively**
- **Using these to plan a detailed assessment**
- **Detailed fieldwork to establish the impact**
- **Optionally, quantifying particular aspects of impact, once we have understood them.**

There is no single 'correct' way to achieve each of these steps. Much may depend on what is already known. The following suggestions are intended to be helpful, practical and widely applicable; they should enable conclusions that are rigorous enough to be convincing.

10 *This is a weakness of the normal contingency planning process: particular opportunities for earlier responses are rarely identified with their windows. It is to be hoped that the learning accumulated through the reviews of DRF by the Start Network can contribute to addressing this weakness.*

Step 1: Open-ended enquiry

This can be used to discover how resources were used by different people or households at different periods of the crisis, and how coping developed (or failed) over the course of the crisis.

It is useful to base conversations around the calendar (when did X happen? What did you do then? What happened next? How long did that last? etc). You are looking to identify the possible windows of opportunity to avoid or mitigate suffering or losses, and any critical time periods when people suffered who lacked resources or opportunities.

We can do this when we understand the logic of people's stories: if children were sick in one village, but not in another, do we understand why? Can we make sense of what they tried to do at different stages in the crisis? Does the story of 'what happened when' make sense, for example, why animals died at a certain time, or why people migrated when they did? If you can't make sense of their stories, then it is useful to ask people directly, explaining why you can't follow a certain logic or what you might have expected to happen at a certain time instead. (This kind of challenge is not done as frequently as it could be in the aid sector.)

It is very dangerous to rely on so-called key informants (people expected to have first-hand knowledge about the community), because it is usual that they understand far less about what is happening outside their own personal experience than they realise. Inevitably, for this initial work the sample will be limited. It is important to maximise the diversity rather than look for a representative sample, since there is no attempt at this stage to characterise what is normal or most frequent.

In order to minimise bias and people's (rational) tendency to promote a particular image of their suffering, it is usually best to avoid any questions that directly mention aid.

Step 2: Analysis and planning

The information gathered in step 1 is used to draw up potential theories of change for the actions taken at particular periods of the crisis. These can be explored widely to develop guides for subsequent, more detailed fieldwork using a theory-based impact assessment approach. More than one possible chain of impact, or branching chains, can be used. (For example, farmers might have been helped to grow more crops. These might be used for food over a critical period, for sale, to save time to permit some migration, etc.) Different households are likely to have used the same funding to support quite different plans of their own: these should have been identified in step 1.

Each link in the chain is investigated and analysed separately rather than to trying to answer the overall question at once. In other words, a group discussion or an individual interview will not be designed to research the whole chain, but specific parts of the chain. An assessment plan might include interviews to test some links, analysis of price data from market monitoring to test other links and an interview with a scientific expert (such as a vet or nutritionist) to verify the plausibility of other links.

If this does not provide any leads for identifying possible benefits of earlier action, then the learning exercise could stop here: if implementing the response earlier could have had no additional benefit, there is nothing more to research. Additional fieldwork may be considered to explore the issues further, but without attempting to prove or quantify any impact.

Step 3: In-depth fieldwork

At this stage, fieldwork will necessarily be qualitative. (The brief exploratory enquiry in step 1 is not enough to show all the possible impacts of a response, so we don't yet know what, if anything, can be measured quantitatively.) The aim is to explore the evidence for each link in the different possible causal chains that we identified in steps 1 and 2.

There is no methodological blueprint here, because the way to test each link depends entirely on the nature of that link. For example, we might use monitoring data or a very rapid and widespread survey to see how many people received a cash grant and when; or we might use more in-depth focus group discussions in a few places with beneficiaries of a cash grant to understand how they spent the money from different sources. We could use interviews with non-beneficiaries to look at changing expenditure patterns in the absence of a grant. We might talk to traders to verify any claimed changes in purchasing or selling patterns. Monitoring data at feeding centres might give us information on changing malnutrition rates in the area over time, and we would probably seek the input of a nutritionist to substantiate any claimed impacts on nutritional status of any reported changes in diet.

All of these inputs would be included in the crisis calendar: the calendar ensures that we analyse both cause and effect more closely (things that happen at the same time or following each other), and captures the importance of early response, identifying when action could or could not make a significant difference.

If there is good evidence that the early cash grant had particular benefits, whether by preventing suffering at a particular time or by enabling recipients to take advantage of a particular window of opportunity, then it may be useful to undertake further quantitative research. Having a larger and more representative sample would a) give us a more accurate and more precise measurement of the benefits and b) increase the rigour of the findings.

A survey can only be used if you are very sure that you have understood all the ways that different people may have used the funding and all the ways in which impact may be seen. It may be tempting to use one or two obvious indicators of suffering, expecting these to be at lower levels for a successful intervention. This is possible if you can be sure that these indicators will reliably tell you about the degree of suffering.

For example, skipping meals is a common response to lack of food, but was a household which skipped meals always coping worse than one which did not? It may be that some households chose to skip meals in order to protect productive assets for recovery. Is getting into debt a burden – or is the ability to borrow money to 'smooth' expenditure over more than one year an indicator of good coping ability? If you understand these dynamics well, then you will know how to design your survey and how to interpret the results.[11] If you are not sure, then you will not know which data to collect and you won't know for sure what they mean.

Knowing the limitations of our understanding is a fundamental strength of any research. Quantitative conclusions are often very useful, but it is more useful to have qualitative findings that we are sure of than quantitative research that is wrong or misinterpreted.

Annex 3 gives an example of one type of project to demonstrate how a difficult problem can be broken down into its component parts.

11 *Questionnaire design and the interpretation of quantitative data are beyond the scope of this guide. As discussed in the introduction, this should be carried out by a partnership between people with specialist technical knowledge and people with specialist knowledge of the issues and the context.*

Question 6:

Was DRF cost-effective?

Value for Money (VFM) calculations are a standard ex ante (before the event) requirement in the grant approval process. VFM calculations for approving current and future grants for the Start Network will be made according to the standard procedures agreed by individual donors, in general focusing on the 4Es of economy, efficiency, effectiveness and equity.

Cost-benefit analysis (CBA) is the most commonly deployed methodology for measuring the 4Es, but is almost entirely quantitative. *Ex ante* CBA (based on forecast figures) is intended to help decide whether an investment is viable. *Ex post* CBA (based on actual figures) justifies the investment based on outcomes. An example of this would be statements such as '$1 invested in a DRF instrument is equivalent to $4 of "traditional" humanitarian response.' The people who might think this question is important are donors, who feel the need for a headline figure to justify investment, but it will not be useful for learning. The calculation arrives at a price per unit measure, which, while appropriate for procurement and logistics, is insensitive when accounting for the intangibles of social development and individual community or household preferences.

The Start Network's DRF instruments provide an opportunity to innovate not only in the way that droughts are responded to, but also in the way that VFM is regarded within the network, and possibly beyond, in the future.

Value for Money should not be treated as a question in isolation, but regarded as an important potential benefit deriving from each of the different component parts or processes of DRF.

THERE ARE A NUMBER OF COMPONENTS TO VFM, SOME OF WHICH ARE RARELY, IF EVER, LOOKED AT INDEPENDENTLY.

1 **Contingency planning:** This takes time and resources and therefore has a cost. There are many benefits of contingency planning, not least that it should accelerate a response.

2 **Coordination:** Contingency planning should lead to better programming in general. With a networked approach (such as that used by the Start Network), contingency planning should lead to more efficient ways of working across and between organisations (thus improving implementation).

3 **Quality of design:** Where agencies are able to share thinking, and to plan adequately without the pressure of an already-present emergency, the quality of programme design should be enhanced (so less wastage will occur).

4 **How populations might have benefitted:** All of the above will probably have been done on the assumption that population X would benefit from a prescribed outcome Y. A VFM exercise should assess whether the outcome was as predicted. It should also examine what decisions were made that were not as expected[12] or were less easily quantifiable.

Annex 4 explores various approach to Value for Money, including traditional cost-benefit analysis along with less common approaches such as Outcomes for Money and Social Return on Investment.

12 *This includes the value that beneficiary households or communities might place on certainty, and the investment decisions they might take in the knowledge of that certainty (which might not necessarily be related to their main source of livelihood).*

THE VALUE FOR MONEY OF CONTINGENCY PLANNING

Agencies rarely try to assess the costs of contingency planning (except when they are budgeting in order to request funds). This means that the benefits of contingency planning have rarely been documented and its value has not been properly assessed. Calculating the Value for Money or benefit-cost ratio of contingency planning should not be thought of as a precise accounting exercise, since it will not be possible to quantify exactly the impact of any improvement in speed or quality of programming. *Nonetheless, considering how exactly beneficiary populations might have benefitted* is an important exercise for learning which elements of contingency planning coordination are useful, and how to improve them. Documenting the benefits of contingency planning, and understanding how they improve VFM, may be an important step in helping to draw attention to, and much greater investment in, this area.

The approach used in this guide treats Value for Money as a cross-cutting issue. We can analyse each aspect of DRF, and of the emergency responses that it finances, in terms of investment (in money and time) and of benefits.

Measuring costs

Implicit in any consideration of cost-effectiveness is an underlying system for valuing costs and benefits. For example, are the costs of recipient populations, such as their time, included? Whose judgement is used in measuring the relative value of (for example) protecting food consumption or protecting livelihood assets? Maintaining family unity or maximising income from migration? Keeping children in school versus earning money to avoid selling assets?

Measuring benefits

We have refrained from offering any way of measuring absolute levels of total losses or suffering from a drought and, as a result, also from any way of measuring the overall impact of DRF or any emergency project. People cope with droughts in different ways and with different priorities. Although we have a fairly simple definition of coping and distress strategies, showing clearly that one is acceptable and the other not, these do not always correspond to how the affected people see their lives, and the line between the two is in any case blurred. There is simply too much doubt whether our indicators of loss or suffering are actually considered as loss or suffering by drought-affected people, and too little evidence that the different forms of loss are correlated. This means we do not have enough evidence for any composite indicator of loss or suffering to make sense.

Reaching conclusions

How, at the end, do we put all this learning together to reach any conclusions about the value for money of DRF as a whole? The simple answer is that we don't. Even if it had an answer, asking about the VFM of DRF as a whole is not the most important question.

The actual benefit brought about by DRF and its costs will be very different in every drought and for every different project that it funds, and those involved will value the different positive and negative impacts in a different way. Finding out how to get the best value for money for scarce aid resources is far too important a question to be reduced to a single VFM figure.

The reasons that we need to understand VFM are to improve DRF, to replicate elements that improve the cost-effectiveness of assistance, and to have well-informed advocacy that can bring about an improvement in emergency response as whole. We cannot achieve any of these with a headline VFM figure for DRF as whole, but with a learning process that examines in detail all the elements that combine in DRF, and which incorporates a VFM perspective throughout, they can all be achieved.

Conclusion

We have looked at six key questions as part of our measurement framework. These will provide the evidence needed to measure the impact of responses facilitated through a DRF initiative. Our framework makes explicit each and every assumption, for example about what will lead to what. An appropriate way can then be found to test each assumption and each causal link individually. The individual lessons are then pieced together like a jigsaw, and the conceptual framework – the map of assumptions and causal pathways – provides the template for linking the lessons together. The individual lessons can be gathered across different countries, for different projects and by different agencies collaborating in learning over several years.

Bibliography & further reading

ALNAP (2006) Evaluating Humanitarian Action Using the OECD-DAC Criteria. London; ODI.

Clark D and R V Hill (2013) 'Cost-Benefit Analysis of the African Risk Capacity Facility', IFPRI Discussion Paper 01292 *www.IFPRI.org*

Cugelman, B and Otero, E 'Basic Efficiency Resource. A framework for measuring the relative performance of multi-unit programs, ver 2.3' Leitmotiv and Alterspark, 2010 *https://www.bond.org.uk/data/files/cugelman_otero_2010_VfM_analysis_OGB_campaign.pdf*

Fleming, F 'Evaluation methods for assessing Value for Money', BetterEvaluation, Oct 2013 *http://www.betterevaluation.org/sites/default/files/Evaluating%20methods%20for%20assessing%20VfM%20-%20Farida%20Fleming.pdf*

Funnell, SC, and Rogers, PJ 'Purposeful Program Theory', San Francisco, CA: Jossey-Bass, 2011

HM Treasury (2004) Regularity, Propriety and Value for Money *http://webarchive.nationalarchives.gov.uk/20130102173401/http://www.hm-treasury.gov.uk/psr_governance_valueformoney.htm*

Hughston, L 'Outcomes for Money (OxM): an Unorthodox Analysis of Value for Money' Plan International UK, 2016. *https://usaidlearninglab.org/library/outcomes-money-oxm-unorthodox-analysis-value-money*

Levine S with A Crosskey and M Abdinoor, 2011 'System failure? Revisiting the problems of timely response to crises in the Horn of Africa', HPN Network Paper 71. London; ODI

Levine S and A Crosskey 2008 'Can "Preparedness Auditing" Make Contingency Planning Actually Useful?' Nairobi; USAID *www.fao.org/fileadmin/user_upload/drought/docs/preparedness_auditing.pdf*

New Economics Foundation/CARE: 'Counting on Uncertainty: the Economic case for Community-based Adaptation in North-East Kenya' July 2012 *http://www.nefconsulting.com/wp-content/uploads/2014/10/Counting_on_Uncertainty_July12.pdf*

Rijneveld, W 'What does the investment return? An experiment in measuring social value in a participatory manner', Red een Kind, Lisap, Resultante. 2013 *https://docs.google.com/file/d/0B0C42UxSyIcaQW5lekRvV2JlejA/edit*

Salverda, M 'Social Return on Investment', BetterEvaluation. Retrieved October 2017 from *http://betterevaluation.org/approach/SROI*

UNICEF/WFP Return on Investment for Emergency Preparedness Study 2015 *https://www.unicef.org/publications/index_81164.html*

White, BA and Rorick, MM 'Cost-benefit Analysis for Community-Based Disaster Risk Reduction in Kailali, Nepal', Mercy Corps *https://www.mercycorps.org/sites/default/files/mc-cba_report-final-2010-2.pdf*

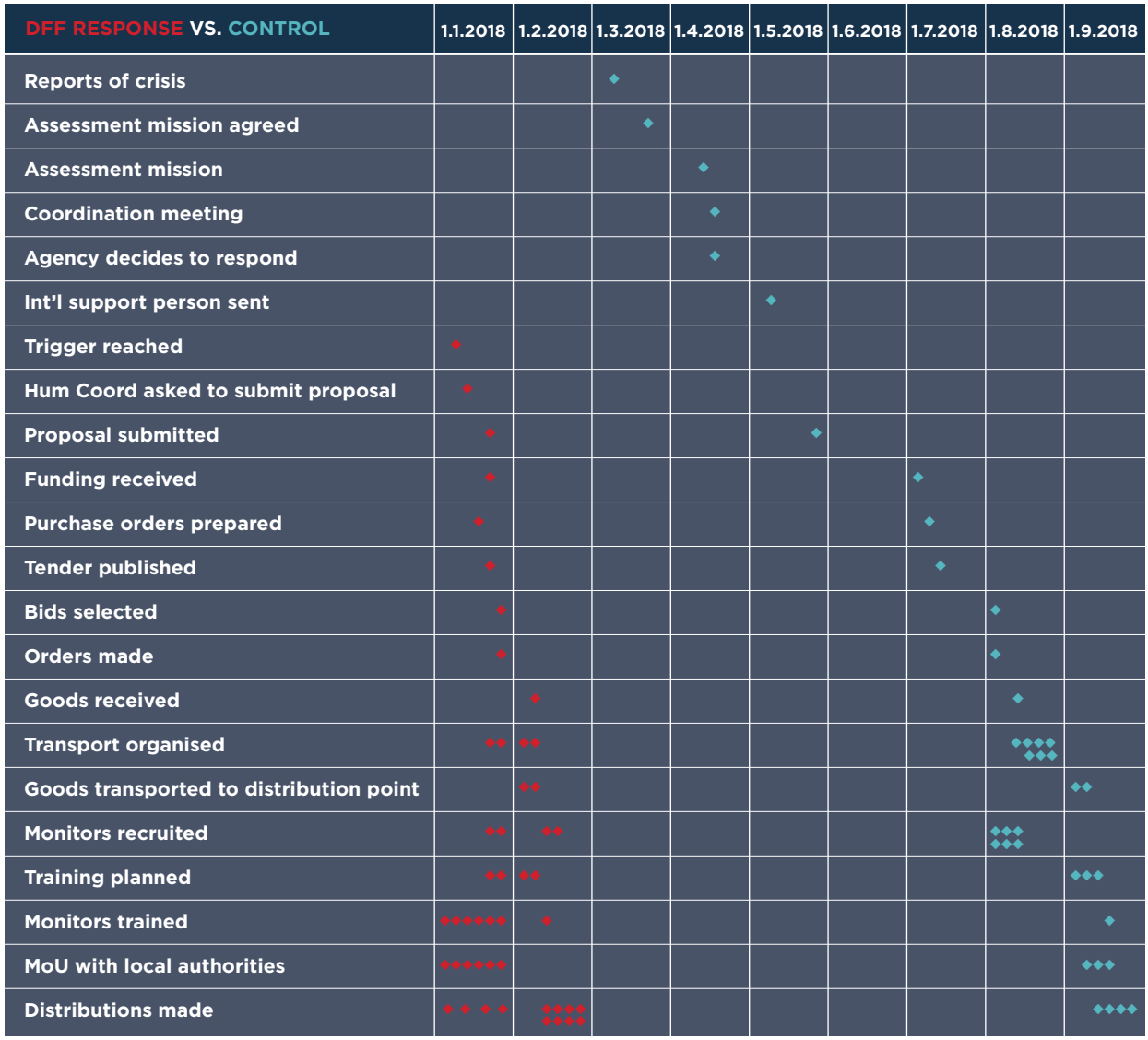
White H 'Theory Based Impact Evaluation: Principles And Practice', Working paper 3. New Delhi; International Initiative for Impact Evaluation, 2009

Annexes

Annex 1:

Example of Gantt chart for comparing speed of response

FIGURE 4: EXAMPLE GANTT CHART



In this example, we are assuming a similar kind of response in both DRF and a control project: an in-kind distribution of food. On this evidence, DRF appears to have sped up response by seven months and one week. We need to be careful in drawing conclusions, if the comparison was between different agencies (for example, one with a dedicated national humanitarian team and one that required international support). The control project may also have deliberately targeted a later, more acute stage of the crisis.

WITHOUT ANY FURTHER INFORMATION ON THE CONTEXT, THE DETAILS APPEAR TO SHOW THE FOLLOWING:

- **The actual funding bureaucracy (time from proposal submitted to funds received) saved four weeks.**
- **DRF triggers were set two months earlier than the subjective triggers that were relied on to warn of a crisis.**
- **The pre-planned DRF system avoided a delay of six weeks between first reports of crisis and a decision to act (time for assessments, etc). However, delaying to verify the situation on the ground may have been important. (Would DRF response have gone ahead if it depended on such assessment?)**
- **Most of the extra response speed was due to preparedness. A well-worked response plan avoided the need for extra international support (three weeks) and preparing proposals (three weeks); and good preparation of purchasing (eg, prequalified suppliers) and staff recruitment saved two months.**
- **Can we improve the speed of DRF response still further? We could look in more detail at why it took ten days from the triggers being reached to the humanitarian coordinator being asked to submit a proposal. ("The country director was on leave, and no one else read the emails".)**
- **The Gantt chart tells us nothing about how good or how appropriate the responses were. Perhaps the later response was needed and not the DRF response? Perhaps the DRF response still missed the window of opportunity when it could have made a difference?**

Depending on the comparison being made, different parts of the Gantt chart will be relevant. For example, if the kinds of project being compared are quite different, then (for example) the time taken for purchasing might have to be very different, but we can still compare the speed of getting funds.

This chart might suggest a clear conclusion (although we cannot be sure without more information about the comparison): DRF was enormously successful in speeding up response. However, we still need to be careful about any recommendations. Let us assume that the investment in contingency planning and preparedness was actually due to DRF. This evidence then suggests that the impact of DRF was huge, but that in future, because of the lessons learned, most of the benefits of speed could be gained by other agencies or in other countries without recreating DRF, just by making a similar investment in contingency planning and preparedness.

Annex 2:

Sample questions to guide the investigation: 'Did DRF lead to earlier response?'

The sample set of questions below are offered as a guide, rather than as a recipe or a set of rules. This list contains many questions, and more may be added according to local experience. Much of this information can be documented in real time. A great deal can be explored in a few in-depth conversations with insightful informants.

1 **Were DRF triggers set earlier than for most humanitarian decision making?**

- i What were the triggers for activating the DRF mechanism? When were these reached (date)? When was this recognised[13] (date)?
- ii How was the humanitarian response first triggered? In the absence of automatic predefined triggers, this is usually a long process with several stages (eg, early warning reports, NGO or national government alerts, UN or interagency assessments, donors asking for proposals, appeals being issued).
- iii On what dates did the different steps happen? What were the triggers for crisis modifiers (ring-fenced budget contingency lines) that might have been built into existing grants? When were these thresholds crossed? Were there different activation dates for different donors, and if so why?

2 **What was the humanitarian community predicting about the coming drought when DRF was triggered? (Sources: early warning systems, famine early warning systems, cluster meeting minutes, government reports, etc.)**

- i Were there any agencies calling for humanitarian action? What was the response? What was the view of the government at this time? What influence did this have?
- ii What was the date of an emergency appeal, if any?

3 **Was the funding bureaucracy quicker?**

- iii How long did it take from the triggers being reached to funds being received by the partner agencies? On what dates were the triggers reached, the funds requested and the funds received? Was this the same for all partner agencies? Were the funds of the predicted amounts?
- iv What was the funding process for other response? Was it possible to request emergency funds before an appeal had been launched or an emergency declared? What funds were available? How long did it take between requesting funds (submitting a proposal) and receiving them (for crisis modifiers, for new emergency funds)? What was the speed difference between different donors? How did pooled funds compare with other donors? Were there differences between sectors?

4 **How much did contingency planning speed up response?**

- i How far were response plans modified as the drought triggers neared?
- ii Were the response plans used? If they had to be changed, or were not used, why?
- iii Did any agencies have proposals already prepared in contingency plans? (Were any of these proposals modified as the triggers neared?) Were these submitted unchanged? (If not, why not?)

13 *There could be a time lag between the trigger threshold being crossed and this information being picked up, although for parametric, automatic triggers this should not happen.*

iv How much quicker were these than other proposals? How long did it take to write them? Did any proposals have to be resubmitted to any donors? Could this have been avoided? What contacts were made between agencies and donors before formally submitting proposals? Did this/could this have shortened the process?
v How far did the contingency plans result in greater preparedness? What steps in mounting a response operation had been taken in advance?
vi How many days were saved, if any, by each step of preparedness? (How long did it take for those steps in agencies which had not done them in advance?)

INVESTING IN PREPAREDNESS

An investment in preparedness would probably not be good VFM where it involves a large investment of time (for the agency, for the communities affected by drought, etc), and where there may only be a benefit every few years. In considering the VFM in any one year, we would need to consider whether there are any other spin-off benefits (the idea of 'no regrets investment', where money is not wasted if a crisis does not occur), or where any investment remains useful for a longer period. This could be activities like establishing an internal financial system that can be more rapid, or preparing job descriptions in advance. Making lists of possible beneficiaries in advance would obviously not make sense for many reasons (including VFM); prequalifying suppliers for a quick tender procedure or establishing procedures for rapid staff handovers to facilitate emergency secondments probably would. More preparedness actions would be considered as providing good VFM as a crisis nears. It is necessary to separate the actual cost-benefit calculation from a judgement about which investments made sense in advance, given what could reasonably be expected at

5 Was the contingency planning and preparedness good value for money?

i How much time was invested in contingency planning? On preparedness? If the assessment is being done in a year when no response was needed, how much of this investment was useful?
ii If a response did take place, how much of the investment had to be replicated anyway?
iii Which parts of preparedness resulted in fewer delays for lower investment?

Up to now, valuations have ignored preparedness measures. There is almost no literature on how to evaluate the contingency planning process, what to look for, or how to look for it. There is no documentation available of any evaluations of contingency planning, or of how those plans were subsequently used when a crisis was declared. There is no literature from the international humanitarian sector where preparedness has been assessed in the absence of a crisis having arisen. This is partly because aid agencies have never really taken seriously the potential for contingency planning to radically transform the way in which they think about identifying windows of opportunity, and prepare to take advantage of them. The Start Network, through DFF, has the opportunity to catalyse a major shift in practice in this area. We are able to study changes brought about in operational readiness through contingency planning and preparedness, whether or not they are ever used directly. Because it is unlikely that DFF will be triggered very often over the next few years, it is important that every opportunity to develop learning in this area is exploited.

Annex 3:

How to assess impact of an earlier cash transfer

This annex gives an example of one type of project to demonstrate how a difficult problem can be broken down into its component parts. The specific questions used here are not intended to be copied: they are only to illustrate an approach to investigation, which, we hope, is of wider practical application. The example used is of an early cash grant to drought affected households. This is only an example and there is no implication that such grants, or indeed any intervention targeted at household level, would or should be a part of any DRF response in practice.

PRECISION OR ACCURACY?

Whenever we make a measurement (whether this is the number of people who have benefitted from an intervention, the breakdown of this number to men and women, how much each beneficiary has gained or lost), it is important to be accurate in our measurements. Without accuracy our conclusions are not rigorous. However, it is important not to confuse precision with accuracy. Precision means that we try and give an exact measure. In fact, the information we will collect is almost certainly not of good enough quality to make the precise assessments which often appear in reports. (It is common to see reports which give percentages even to two decimal places, as if the authors were really sure that their figures are so precise.)The more precise we try to be, the less accurate we are likely to be. For example, if I say that it is two minutes to five o'clock and in fact it is four minutes before five, then I am inaccurate. However, if I say that it is about five o'clock when it is four minutes before five, then I am completely accurate – just not very precise. Our accuracy comes from not claiming more precision than we have.

When we are trying to learn about the workings or the impact of DRF, each time we have to measure something we need to ask ourselves how precise it is *important* to be – and how precise it is *possible* to be. If we are trying to quantify delays in project implementation due to administrative bureaucracy, it may be useful to be precise even to the nearest day. However, similar precision in trying to assess how long a food distribution was able to sustain a family would usually be unnecessary and almost certainly lead to a wrong answer – false precision.

In our example, the rains failed between October and February. Following this, a rapidly worsening crisis led to a government initiative in November subsidising staple food, plus a large scale emergency response (including food distributions and cash grants) starting in November and continuing until the following year.

A rapid scoping study and subsequent analysis of an early cash grant will have given us some leads on why the early grant might have been useful. Let us suppose that we think we are finding three different positive impacts.

1. **Some of those who received the cash reported having much lower debt burdens at the end of the crisis, and so were able to recover more quickly.**
2. **Some recipients said they were able to feed their children better, and it seems this avoided widespread malnutrition (a problem which was very prevalent in the period just before the government started to subsidise staples in November).**
3. **Livestock owners were apparently able to hire transport to take their animals, which would have been too weak to walk, to market. This allowed them to receive some income (and thus buy food) from animals that would have otherwise died.**

What we are investigating

First, we have to be clear what we are investigating. The lower level of debt at the end of the crisis might be because beneficiaries received more, rather than earlier, aid. We would hope that the scoping exercise gave us some reason why receiving the cash earlier made a difference. If not, we would try and establish this at the beginning of the in-depth field work.

For example, perhaps we hear that recipients of the early cash were able to pay off previous debts, and were therefore able to buy food on credit at a particularly critical time, in September/October. Those who had not paid off debts had reached their credit limit and so went hungry until the subsidised food was available. If we cannot find any reason why the timing of the aid made a difference to their debt levels at the end of the crisis, then we may abandon this thread and conclude that it was not relevant to DRF and being early. We would then concentrate on the other two possible impacts.

We will want to understand the different sources of food and income available to different households in the period leading up to November and beyond, in order to investigate further the purchase of food at the critical time (just before November).

We will also want to understand the different expenditure patterns of different households, to ensure that there is a logic matching income and expenditure. This means that income (which includes credit) and expenditure should match over fairly short time periods (which could be as little as one or two days). The expenditure also has to make sense in terms of the crisis.[14]

FIGURE 5: CRISIS CALENDAR FOR EARLY CASH GRANTS (PARTIAL ONLY)

	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
Rains	**Failed rains**											**Rains**
Harvest			Harvest c 20% of normal				Hh stocks low		Most stocks over			
								DFF cash			**Gov subsidy**	
Malnutrition							Rising		Feeding centres open			Reduc-ing
Grain prices	5/kg	5/kg	4/kg	4/kg	5/kg	6/kg	8/kg	11/kg	11/kg	11/kg	5/kg	4/kg
Hh credit level (non-benf)	0-100	0-100	0-20	0-20	0-100	0-150	75-250	0-150	0	150-200	250	200-250
Hh credit level (benf)	0-100	0-100	0-20	0-20	0-100	0-150	75-250	100-250	150-250	250	250	200-250
Food purchases					**Increasing ·····>**			Credit limit reached	Benefs+ rich only	Benefs+ rich only	**High**	
Other Income	Local labour			Local opps limited	Migration to town	Migration increases ·····>			No work opps. return fron town			

14 *For example, if malnutrition among non-beneficiaries was particularly high, we would not expect to find many purchases of non-essentials at that period.*

The crisis calendar shows visually how far the stories that were heard have a coherent logic behind them. In Figure 5, there is a clear link in time and logic between the end of household food stocks, the rise in food prices, people reaching a credit ceiling and increasing acute malnutrition. There is also a clear link in time and logic between the cash grant, the credit ceiling and increasing food purchases by beneficiaries. A full crisis calendar could also add food consumption patterns, typical monthly income levels, etc.

Finding and testing information

Information about malnutrition may come from personal testimonies from individual interviews, from estimates given by interviewees about what they know about malnutrition in the area in different weeks and months, and information from key informant interviews, as well as any secondary data, for example from admissions to nutritional centres or from nutritional surveys. Information on purchases should be checked by talking to traders: they can confirm quantities bought, the range of customers, prices, credit patterns, etc. Further verification can come from a Household Economy Approach. The scale of household food purchases, combined with reports of any other food sources, should match the quantities which we would expect people to eat during a crisis: income should match expenditure, etc.

Asking respondents how they used money from any particular source, including DFF, is rarely useful. If people spend money from one source to cover one need, it frees up other money for something else. What is important (for all of us) is the total money we have and what we can do with it. So it is not important what any recipient of a transfer happened to use 'your' money for: the additional purchases they were able to make as a result of having the cash grant is more significant. We learn this better by taking a more indirect route, trying to understand how people used and needed money, and about what they added as they could find more money. (This may also reduce the problem of recipients telling you what they think you would like to hear about how they spent 'your' money).

Let us suppose that recipients claimed to have bought more food as a result of the cash grant relieving their credit ceiling, and that this was substantiated. They further claimed that the extra food prevented the malnutrition that they saw in other places. This requires careful examination. Is it plausible that the quantities of food items they were able to buy could significantly affect their nutritional status? It will not be possible to get precise measurements of the quantities of food eaten, but the estimates available from their purchases should suffice (eg 'an extra 20 kg of grain for a household of five, eaten over six weeks'). A nutritionist can be consulted on the plausibility of the link to nutrition.

We may be testing whether providing a cash grant early enough (before the government took action to subsidise food prices) has an added advantage in preventing malnutrition. First, we will need to have testimonies and evidence that are well researched and validated. We will also need a convincing logic behind the story that early cash recipients were more likely to be able to make larger food purchases in the period immediately before October, and that non-recipients were frequently unable to find enough to eat in the period and had higher malnutrition rates than the early recipients. Only then can we make a strong case for the theory.

Although we will not yet have a clear estimate of how much malnutrition was prevented, the case that DRF cash grant helped to prevent malnutrition is actually much stronger than one which is based solely on nutritional surveys of beneficiary and non-beneficiary populations.

Possible questions for investigating the impact of a cash grant on consumption by alleviating debt

- What sources of income did people have in each month? How much money?
- If they received money, what was their priority in spending it, in different months? Why?
- In different months, what coping strategies could people who had a little money engage in? People who had less money?
- What stocks of food did they have?
- How much did they eat? Of what kinds of food? Where did food come from? How much did this cost, each week/month? Where did the money come from?
  (This must all match, following a Household Economy Approach: food requirements = food consumed, purchases = income + debt.)
- What did they have that could be sold to buy food? Did anyone sell anything? In what months? Who/why/why not?
- Where did they buy food? How did prices change over the period ? (Verify with traders.) How much did they buy? How much on credit? What was their debt each month?
- What was the highest debt they were allowed? What happened to people who reached this level of credit and had no money? What did they do then to find food/cash? How many times a day did they eat (men, women, children)?
- What was the history of credit before the drought – what credit would be normal? What was the highest credit they have been given before? (These questions are for both the borrowers and the traders giving credit.)
- What were malnutrition rates in the area over the period from June to December? If there is no data from feeding centres or surveys: what symptoms of malnutrition did people in the village see? In which months?

The analytical question is: Are there patterns that link credit, purchases and consumption with cash grants in the August–September period?

A similar approach would be followed to investigate the impact of early cash in helping livestock keepers to sell their animals before they died. The same crisis calendar would be used to map how livestock fared as the crisis developed. This could include changes in body condition, marketability, livestock prices, the existence of a market for animals at different times and any livestock aid, such as a destocking programme.

The calendar analysis of livestock prices and marketability should show a logic to the specific timing of the grant (eg in September) in relation to the economics of marketing animals. An economic analysis will be needed to confirm both the plausibility of the reports on marketing and also on the possible impacts that this could have on household well-being. This would involve finding out the costs of animal transport hire and any other costs of marketing, how many animals people owned, how many they were able to sell, their net profit, etc. We may be able to verify this from the transport sector: for example, interviews with truck owners can confirm any increase in vehicle hire around September/October.

We also need to make sense of the marketing strategy of those who did not receive the early cash. Why were they unable to sell animals? Did they describe the same market conditions, but complain that they were unable to take advantage of them? If selling animals made economic sense, why did they not borrow money to hire transport (and then repay the loan with the money from the animal sales)? If we understand this logic, it will give us the confidence to assert that the cash grant distributed at a particular time transformed the market opportunities of the grant recipients.

Calculating the exact benefit from selling animals may be complicated, because it also depends on what happened to animals that were not sold. If many of them died, then the benefit of the grant could roughly be equated to the net income raised from selling the animals (which would otherwise have had no value at all). If most animals survived the drought, then the people who sold them, while being able to buy food, will have lost assets. However, there is no need for us to make our own assertion of the relative value of protecting food consumption versus protecting livelihood sources for the future. This would be necessary if we wanted to calculate a benefit-cost ratio for the project. If we simply describe what we have discovered and quantified, we are presenting objective evidence of the impact of the project, without imposing our own value judgements and without reducing a complex narrative to a single number.

Annex 4:

Perceptions of value

Traditional cost-benefit analysis for Value for Money (VFM)

In standard VFM calculations, the perception of value stems from the provider (the donor), encouraging the optimisation of resource use, but with the caveat that "... 'optimal' means 'the most desirable possible, given expressed or implied restrictions or constraints'. Value for money is not about achieving the lowest initial price."[15]

Almost inevitably, the calculation arrives at price per unit measure, which, while appropriate for procurement and logistics, is insensitive when accounting for the intangibles of social development and individual community or household preferences.

The 'value chain' of the DFF

Similar to the causal chain viewed through the learning framework, there is a chain of perceptions of value. These change according to the position of the individual or entity on that chain, and will include a number of economic intangibles or 'difficult to value' perceptions.

USING THE ASSUMPTIONS IDENTIFIED ABOVE, PERCEPTIONS OF VALUE OF THE DFF CAN BE SEEN AS FOLLOWS.

For the DFF itself: the VFM calculation is purely actuarial. The facility should provide a return on the initial premium invested. This assures a fund that covers any claim and leaves a residue for future demands supplemented by further premiums.

For the donor community: mitigates the need for early decision making? A risk-sharing mechanism through co-financing? A catalyst for internal decision-making? A bargaining chip with national government to accelerate decision-taking?

For national government: an assurance of early/earlier action through co-financing.

For local government: a guarantee of a response to an expected or identified need, a planning tool and uncertainty regulator, and, potentially, a bargaining chip for additional resources from central government.

For the Start partner (country programme and/or national NGO): a tool to mitigate uncertainty that allows planning with a reasonably predictable level of financial resourcing; a negotiating platform for work with local government and communities.

For the community or household: a recognition of the priorities for coping with and adapting to an emerging problem, identified through the contingency planning process, and a potential means to *diversify* in order to mitigate the worst effects of that problem ahead of time.

Assuming that the contingency planning process starts at the beneficiary level, the perceptions of value chain should start from there and work upwards.

15 *HM Treasury, 2004*

Approaches in common use

VFM calculations are used prospectively to justify an initial investment or retrospectively to show that the investment worked. Rarely, if ever, do prospective calculations take into account the value that the beneficiary might apply to any investment intended to assist them.

Cost-benefit analysis (CBA)

This is the most commonly deployed methodology. It provides a measure of the 4Es, but almost entirely in a quantitative manner. *Ex ante* CBA is intended to help decide whether an investment is viable. *Ex post* CBA justifies the investment on the outcomes.

Mercy Corps noted in their ex post CBA analysis of the community-based disaster risk reduction programme in Kailali, Nepal[16] that the intangible benefits of things such as community or individual empowerment, or the indirect impacts on economic capital brought about by the programme, could only be captured through qualitative enquiry. They list what Mercy Corps considers should be included in any benefit-cost calculation.

The New Economics Foundation (NEF) observed in its Counting on Uncertainty study for CARE International in NE Kenya:[17] "Donors are often concerned with the opportunity cost of investment decisions. When considering the opportunity cost of investing in community based adaptation... the [value] is virtually zero because investment in community based adaptation is synonymous with good development practice." Again, NEF points out that the intangibles are lost in standard approaches.

Return on Investment (ROI)

The Return on Investment approach is a financial measure in which "an ROI rate of 1.0 indicates that future costs will be reduced by the same initial investment amount" (Boston Consulting Group). Thus, all returns with a value greater than 1.0 denote a higher cost saving than the original investment.

Boston Consulting Group (BCG) in its DFID-funded 2014 study for UNICEF and World Food Programme of the ROI for emergency preparedness[18] found that all investments studied saved significant time or costs in an emergency:

75% SHOWED COSTS SAVINGS GREATER THAN THE INITIAL INVESTMENT (ROI>1.0)

64% SAVED BOTH TIME AND COST

93% SAVED TIME TOWARD RESPONSE

However, the calculations were measured in purely material terms: logistics, procurement, staffing (employment, training etc) and partnerships (external contracting to NGOs and others). It discounted any of the downstream impacts and any of the externalities of partner activities, influencing etc.

In other words, the ROI analysis is insensitive to the finer social issues that are the concern of both humanitarian and development activities.

16 *White and Rorick, 2010*
17 *New Economics Foundation, 2012*
18 *UNICEF/WFP, 2015*

Less common approaches

Outcomes for Money (OxM)

Developed by Plan UK, the OxM methodology is an ex ante approach that defines value as "the changes experienced in the communities" benefitting from their Building Skills for Life programme. These changes are measured in terms of knowledge, attitudes and behaviour. OxM uses visual as well as more mainstream analytical tools to colour perceptions of value, and shifts the emphasis away from the purely economic impact of a project to look at its outcomes; where the impact of any particular investment has been greatest; and what strategies and resource allocation priorities arise for future activities to either sustain the impact of the project as a whole or to even up impacts between activities within a project.

Social return on investment (SROI)

The SROI approach supports thinking along the lines of a results chain, but does not assume that components of that chain have a linear relationship. It accepts that development takes place in complex situations. The SROI ratio is not the deciding factor in an investment decision, but can be used as a tool to aid explanation and communication of the progress of an initiative.

While SROI is generally described as an ex post approach, it can be adapted for ex ante analysis of value.

PARAMETERS OF SROI[19]

- **Define the boundaries.**
- **Identify the influencers** (positive and negative)**.**
- **Develop the Theory of Change** - who benefits (or not); why; what; how? This will be perceptions-based.
- **What goes in and what comes out** (the input and output variables)**.**
- **Valuation:** indicators that help to turn perceived benefit-cost into a monetary value. Tangibles (eg, increased or diversified income) and intangibles (eg, increased or decreased status): use of different measures such as ranking or opportunity cost. Include:
  - **Deadweight** (what would have happened anyway)
  - **Attribution** (who/what else helped)
- **SROI:** comparison is made of the investments (inputs) on the one hand and the financial, social and, for example, environmental returns (outcomes and impact of an intervention) on the other. Triangulation through, for example, semi-structured stakeholder interviews or narratives**.**
- **Verification.**

WHERE SROI COULD BE USED[20]

- **The method can be used to focus on real outcomes for real people, since it forces participants to consider what really changes for people and to apply a value to this change. This would bring monitoring from the output level to the outcome level.**
- **However, it can be overcomplicated for beneficiaries, especially if asked to 'monetise their lives'. As with all VFM approaches, there are also ethical questions around health, bereavement, etc (net present value of a life lost, for example).**
- **SROI analyses could also be used as the overarching framework for outcome level reporting. It would not be possible to set fixed outcomes and fixed indicators, but each location, or each project or each partner could be asked to use participatory SROI analyses to report on changes effected and apply a value to those changes.**

19 *Salverda, retrieved 6 June 2017*
20 *Rijneveld, 2013*

PROS AND CONS OF THE METHODOLOGY[21]

- **SROI appears to have value for NGOs in identifying and quantifying the benefits (and detriments) of an intervention that may otherwise be overlooked or understated. Its greatest strengths are the objectivity that it brings to a comparison when looking at diverse outcomes within a project.**
- **Its economic language may provide a useful and reassuring medium of communication to donors and critics.**
- **SROI data collection and analysis is very time consuming, so is best done as a discrete evaluation approach rather than a component within a mixed method approach.**
- **Its generation of a simplistic cost-benefit ratio is likely to be misused by decision-makers inside and outside NGOs, perhaps to the detriment of richer descriptive interpretation.**

According to NEF Consulting, Social Cost Benefit Analysis (and, by extension, SROI) "is a particularly effective tool at valuing what matters to stakeholders, capturing avoided losses and placing externalities within a decision-making framework".[22]

Basic efficiency resource (BER)

BER was originally designed by Cugelman and Otero[23] to evaluate Oxfam GB's global climate change campaign. It aims to provide a measure of value for multi-unit programmes, where each unit might be of a different scale (and subject to 'importance ranking' as a consequence).

It is not a stand-alone methodology but is useful at the end stage of an evaluation, to triangulate and to investigate further why certain units performed differently (perhaps unexpectedly) from others. It is adapted from a BCG portfolio management analysis technique, customer satisfaction quadrant analysis and some aspects of social network analysis. It builds upon the SROI approach.

Although designed for end use, there are elements of the technique, in particular the weight that is put on *perceptions* of impact and value, that might make it adaptable for both project design and evaluation. It also allows for counterintuitive interpretations.

Cost-effectiveness analysis (CEA)[24]

An alternative to CBA, this is most useful when costs are not easily applicable and outputs not easily monetisable (eg, 'better health' in an individual or community).

It measures costs in, for example, dollars but provides outputs in units. These are therefore incommensurable and can only be expressed as a ratio.

- **CE Ratio (cost/effectiveness) measures the cost/unit saved.**
- **EC ratio (effectiveness/cost) measures $ spent/life saved.**

CEA is most useful at the design stage (ie, before a programme has commenced). It can be used to develop counter-factual scenarios comparing the programme to similar activities or to options that have not been tried.

21 *Fleming, 2013*
22 *ibid*
23 *Cugelman and Otero, 2010*
24 *http://www.betterevaluation.org/en/evaluation-options/CostEffectivenessAnalysis, accessed 3 September 2017*

Cost utility analysis (CUA)[25]

This is useful for comparing and evaluating alternatives for programmes that have no intended monetary value in their outputs, for example, the health sector and Quality Adjusted Life Years (QALYS) which reflect the preferences of the patient to the treatments on offer. The tool, therefore, is a useful way to reflect the desires of the individual or household, and might be applicable to the DFF.

- **CUA makes careful attempts to consider individual preferences.**
- **It allows the inclusion of a large number of potential outcomes in any evaluation of options.**
- **It can contribute to consensus building and participatory decision-making as stakeholders are called upon to assess their preferences for different outcomes.**

It is traditionally dependent upon the quantification of 'hardware' to reach its conclusions, but it can equally be used to value 'intangibles' and the costs that the beneficiary has to bear to benefit from any intervention.

Conclusion

The choice of approaches is not exhaustive. It allows reflection as to what can or cannot be captured when considering 'value' as something more than simply a cost per unit measure of effectiveness and equity. The choice does not exclude the use of current 'standard' approaches. CBA and/or ROI are, in the current environment, necessary and well-understood methodologies for estimating VFM gains (and potential losses).

If, as we suggest, the standard approaches are accepted as being insensitive to the finer social or qualitative issues that are our prime concern, then parallel exercises might be deployed that satisfy donor needs and aid agency concerns. DRF provides the opportunity to experiment and innovate in this regard.

25 *http://www.betterevaluation.org/en/evaluation-options/cost_utility_analysis accessed 3 September 2017*


START
NETWORK


ACTION AGAINST HUNGER
actionaid
Adeso
christian aid
CAFOD
Just one world
care
CRS
CATHOLIC RELIEF SERVICES
WAR child
ARDD-LEGAL AID
NRC
NORWEGIAN REFUGEE COUNCIL
Save the Children
International Medical Corps
community
WORLD service
RELIEF INTERNATIONAL
welt hunger hilfe
Caritas
SRI LANKA
HelpAge International
PLAN INTERNATIONAL
GOAL
tearfund
Qatar Charity
Oxfam
INTERNATIONAL RESCUE COMMITTEE
MERCY CORPS
HANDICAP INTERNATIONAL
SOLIDARITÉS INTERNATIONAL
DORCAS
RELIEF & DEVELOPMENT
world jewish relief
MAG
saves lives builds futures
Trócaire
ISLAMIC RELIEF
RED VIDA
ACTED
CONCERN worldwide
ALIMA
Muslim Aid
Serving Humanity
CADENA
COMUNIDAD JUDIA DE MÉXICO
DOCTORS OF THE WORLD
Cordaid
World Vision
For children. For change. For good.