

# **Building resilience and managing risk in fragile and conflict-affected states: A thematic evaluation of DFID's multi-year approaches to humanitarian action in the Democratic Republic of Congo, Ethiopia, Sudan and Yemen.**

---

VALID INTERNATIONAL

## **INCEPTION REPORT**

Bill Gray, Courtenay Cabot Venton, Lewis Sida and Simon Levine

## Contents

<u>1. Introduction and background</u>	<u>page 4</u>
1.1 Introduction	page 4
1.2 Background	page 4
1.3 Context	page 4
1.4 DFID portfolio	page 5
1.5 Evaluability assessments	page 7
1.6 Pilot	page 8
1.7 Structure of the report	page 11
<u>2. Purpose and scope</u>	<u>page 12</u>
2.1 Purpose	page 12
2.2 Evaluation matrix	page 12
<u>3. Approach</u>	<u>page 14</u>
3.1 Evaluation question 1a: resilience	page 16
3.1.1 Logic model	page 16
3.1.2 Methods	page 21
3.2 Evaluation question 1b: humanitarian outcomes	page 27
3.2.1 Logic model	page 27
3.2.2 Methods	page 28
3.3 Evaluation question 2: contingency	page 29
3.3.1 Logic model	page 29
3.3.2 Methods	page 31
3.4 Evaluation question 3: value for money	page 36
3.4.1 Logic model	page 36
3.4.2 Methods	page 38
3.5 Data management and analysis	page 41
<u>4. Evaluation team and management</u>	<u>page 42</u>
4.1 Management of the evaluation	page 42
4.2 Evaluation team and roles	page 42
4.3 Evaluation uptake strategy	page 46
4.4 Quality assurance	page 47
4.5 Risks	page 48
4.6 Constraints and limitations	page 49
4.7 Ethical guidelines	page 49
<u>5. Work-plan and outputs</u>	<u>page 50</u>
5.1 Outputs and milestones	page 50
5.2 Work-plan	page 50
5.3 Synthesis report	page 50
<u>Annexes</u>	
1. Evaluation matrix	
2. Bibliography	
3. Work-plan	
4. Valid ethical guidelines	

## List of Acronyms

<b>ACF</b>	Action Contre la Faim
<b>ARCC</b>	Alternative Responses to Communities in Crisis
<b>BRACED</b>	Building resilience and adaptation to climate extremes and disasters programme
<b>BRCiS</b>	Building Resilient Communities in Somalia
<b>CRS</b>	Catholic Relief Services
<b>DAC</b>	Development Assistance Committee
<b>DFID</b>	Department for Foreign and International Development
<b>ECHO</b>	European Community Humanitarian Office
<b>EQ</b>	Evaluation Question(s)
<b>EWS</b>	Early Warning System
<b>FAO</b>	Food and Agriculture Organization
<b>HEIP</b>	Humanitarian Evaluation and Innovation Programme
<b>HPG</b>	Humanitarian Policy Group
<b>HQ</b>	Head Quarters
<b>HRF</b>	Humanitarian Response Fund
<b>IDP</b>	Internally Displaced Person
<b>IDS</b>	Institute for Development Studies
<b>IOM</b>	International Organization for Migration
<b>M&amp;E</b>	Monitoring and Evaluation
<b>MSF</b>	Medecins Sans Frontiers
<b>MY</b>	Multi Year
<b>MYHF</b>	Multi Year Humanitarian Funding
<b>NCA</b>	Norwegian Church Aid
<b>NGO</b>	Non Governmental Organization
<b>NRC</b>	Norwegian Refugee Council
<b>OCHA</b>	Office for the Coordination of Humanitarian Affairs
<b>ODI</b>	Oversees Development Institute
<b>PBS</b>	Promoting Basic Services
<b>PDP</b>	Peace and Development Programme
<b>PSNP</b>	Productive Safety Net Programme
<b>QA</b>	Quality Assurance
<b>SCI</b>	Save the Children International
<b>ToC</b>	Theory of Change
<b>ToR</b>	Terms of References
<b>UN</b>	United Nations
<b>UNHCR</b>	United Nations High Commissioner for Refugees
<b>UNICEF</b>	United Nations International Children's Emergency Fund
<b>UMCOR</b>	United Methodist Committee on Relief
<b>USAID</b>	United States Agency for International Development
<b>VFM</b>	Value for Money
<b>WFP</b>	World Food Programme
<b>WHS</b>	World Humanitarian Summit
<b>WVI</b>	World Vision International

# 1. Introduction and background

## 1.1 Introduction

This inception report sets out the approach, work-plan and practical arrangements for the thematic evaluation of DFID's multi-year humanitarian funding. The thematic evaluation will take place in the Democratic Republic of Congo, Ethiopia and Sudan, and possibly one other country. It builds on the terms of reference and six months of scoping work, including evaluability assessments of the four countries.

The evaluation is complex, and as a result the inception period has been invaluable in working out both conceptual and practical arrangements for its implementation. These are reflected in the inception report.

In the course of this scoping work, with security deteriorating in Yemen, the difficult decision not to proceed with this part of the study was taken. Alternatives are still being worked through and an updated version of this report will be needed once this decision has been made.

## 1.2 Background

Multi-year humanitarian funding (MYHF) and the accompanying contingency arrangements are a bold and innovative approach to the challenge of protracted crisis. DFID is increasingly recognising the long term nature of these crises and the inadequacy of short term approaches. MYHF is potentially a way of working better in such contexts – retaining the flexibility of the humanitarian instrument whilst also taking the longer term view. The assumption is that this allows partners to plan better, leading to more effective programming, and potentially better value for money.

Whilst intuitively the new approach should improve effectiveness (and anecdotal evidence from at least two funded partners suggests that predictability is leading to better planning), there is currently little hard evidence that this is actually the case. Mostly this is because such approaches are so new there has not been time to see what differences, if any, occur. However, it is also the case that DFID responses are mediated by partners. It is the recipients of DFID funding – primarily UN agencies and NGOs – who do the work, and thus any change in programme effectiveness is dependent on the partners changing in response to the new DFID approach.

A wider question that the evaluation will seek to answer is whether by taking a longer term approach DFID and its partners can help communities become resilient in the face of shocks.

## 1.3 Context

By some estimates 80% of the global finance for humanitarian action goes to protracted crises<sup>1</sup>. Since 1992 - the beginning of the modern UN appeals system - this financing has been annual in nature. And yet most of the countries in the appeals

---

<sup>1</sup> Global Humanitarian Assistance (2013). Development Initiatives.

system end up needing assistance over multiple years. In fact, there are very few examples of countries that are included in the UN complex emergency appeals system (currently called ‘strategic response plans’) that only stay for one year.

The practical implications of this are that UN agencies and their partner non-government agencies end up running programmes over multiple years that rely on annual funding. This is inherently unpredictable, and usually means that longer term investments can’t be made. Often it means staff are on short term contracts, a disincentive sometimes for attracting the right skills. And it can mean valuable management time absorbed in fund raising that might be better spent running programmes.

For donors too, the annual funding process can be inefficient. At the very least it means an additional bureaucratic burden – having to spend time every year going through internal processes to secure and release finance. Given the increasing rigour required for securing funding and ensuring the right oversight framework is in place, this can be extremely time-consuming; often taking weeks if not months. This means time lost for overseeing programmes, and for engaging strategically with partners.

The inefficiencies may well be broader than this. One of the hypothesis under-pinning the move to multi-year funding and planning is that this can result in different types of programmes. UNHCR in Ethiopia is a good initial example of this. They report that MYHF in the Somali based Dollo Ado camps has allowed them to invest in more durable shelter, rather than shorter term solutions. Over a period of three years they calculate this to have saved them money, and saved refugees distress. Clearly, the evaluation will attempt to quantify this and similar examples.

## 1.4 DFID portfolio

The combined humanitarian portfolio across the four countries is in excess of £400m (see table 1), with time frames ranging from two years (in Yemen) to five years (in DRC).

**Table 1: combined humanitarian portfolio in the four DFID countries**

DRC	£135m	2012-2017
Ethiopia	£142m	2012-2016
Sudan	£88m	2013-2016
Yemen	£38m	2013-2015
TOTAL	£403m	

Overall, this breaks down as shown in table 2 below.

**Table 2: overall humanitarian portfolios by country and partner**

DR Congo	£158m	Sudan	£88m
Pooled Fund	£87.5m	CHF	£31m
UNICEF	£21m	ICRC	£7.5m
ICRC	£18.3m	UN agencies	£15.5m
NGOs (emergency)	£5.8m	INGOs	£12m

WFP	£5.4m	Evaluation	£1m
NRC	£6.1m	Contingency	£21m
ACF	£4.5m	Yemen	£52.4m
OCHA M&E + UNHAS	£4m	Oxfam	£10m
Flexible response	£5.5m	Save the Children	£10.2m
Ethiopia	£142m	IOM led consortium	£8.5m
WFP	£95m	CARE led consortium.	£8.5m
UNHCR	£22m	OCHA	£1m
OCHA	£25m	OCHA ERF	£4.9m
		ICRC	£9.3m

Three of the four countries have a mixed portfolio of annual and multi-year humanitarian funding; Ethiopia has an exclusively multi-year programme, although with staggered timeframes across the funded period. Assuming multi-year to mean longer than two years' duration from the start of the evaluation, the total MYHF portfolio to be evaluated at present across the four countries is roughly \$220m (£122m)

**Table 3: DFID partners in the four countries, with programme areas, value and timeframes**

	Partner	Programme area	Value	Timeframe
DRC	• NRC.	IDPs	£6.1m.	2013-2017
	• ACF	Nutrition	£4.4m	2012-2015
	• UNICEF.	Cash	£5.98m	2014-2016
Ethiopia	• WFP.	Food	£95m.	2012-2015
	• UNHCR.	Refugees	£22m.	2012-2015
	• OCHA	Emergency fund (HRF)	£25m.	2012-2016
Sudan	• CRS, Oxfam, NCA, WVI, UMCOR.	Agriculture, nutrition & DRR.	£12m.	2014-2016
	• UNICEF/ WFP/ FAO	Food, nutrition & agriculture	£15.5m	2014-2017
Yemen	• SCF.	Multi-sector	£10.2m	2013-2015
	• Oxfam.		£10m.	2013-2015
	• CARE.		£8.5m	2013-2015
	• IOM		£8.5m	2013-2015

While Ethiopia is both the largest programme and exclusively multi-year, DRC by contrast is just over £16m or about 12% of the portfolio. Sudan's share is 40% (the rest being contingency). The WFP grant in Ethiopia of £95m constitutes half the multi-year funding across the portfolio, especially once Yemen is removed.

The exclusion of Yemen means that the MYHF portfolio comprises only eight grants. However, within these, there is significant diversity of sectors covered. As can be seen from Table 2, the study takes in all of the big humanitarian UN agencies (WFP, UNICEF, UNHCR, OCHA and FAO) and through the CRS led consortium in Sudan, the funding for ACF and NRC in DRC, and the OCHA HRF in Ethiopia there is also

healthy INGO coverage. There is a similar diversity of programming for cash and cash transfers, agriculture, nutrition, refugees and IDPs. The humanitarian response fund (HRF) – a pooled emergency fund managed by OCHA - also funds most sectors, albeit for periods from 3-9 months (presenting challenges with tracking significant change). In all four countries there is an explicit linkage between MYHF and ‘resilience’ with at least one MYHF partner looking specifically at the issue<sup>2</sup>. The wider DFID portfolio (see Table 3) in each of these countries is also either explicitly or implicitly focused on building resilience. This needs to be factored into the evaluation as the dividing line between MYHF and development programmes focused on building resilience (such as the PSNP and PDP programmes in Ethiopia) is fine and, to some degree, deliberately blurred in order to join funding and programming across administrative lines.

**Table 4: wider DFID portfolio programmes of direct relevance to the evaluation.**

DRC	Tuungane. Post conflict recovery programme in the east. £95m over four years focused on community reconstruction through linking with service provision.
Ethiopia	PDP, PBS and PSNP (approx. £560m in total)
Sudan	Tbd
Yemen	n/a

Contingency funding is also diverse across the portfolio with each country having taken a slightly different approach. In Ethiopia, the contingency is committed to the partner agencies (OCHA and UNHCR) although DFID still has to approve; in Sudan (the largest at £21m) it is at the discretion of the country office. In DRC there is no contingency arrangement but funds are allocated on an annual basis to respond to acute emergency. Yemen was under design at the time of the inception report (see table 4).

**Table 5: Contingency Funding.**

	Amount	Approach
DRC	None	But un-programmed funds in the BC for emergencies.
Ethiopia	£7m	£1m with HCR, £6m in the HRF
Sudan	£21m (£7m/ year)	Programmed by DFID Sudan
Yemen	£5m	Under design.

The final synthesis report for this evaluation will also include findings from a similar exercise being carried out in Somalia. This is being done under a different contract and includes a substantial third party monitoring component; however there is also provision for the evaluation of multi-year humanitarian financing along very similar lines to the framework for this exercise. The Somalia portfolio is still evolving, but in essence looks like:

<sup>2</sup> In Ethiopia UNHCR is looking to build refugee resilience; in DRC, both NRC and UNICEF are focused on building IDP resilience; and, in Sudan, the focus is on post- conflict displaced communities.

**Table 6: Somalia humanitarian portfolio, including multi-year humanitarian components.**

	2013/14	2014/15	2015/16	2016/17	Total (£ million)	Annual or Multi year (MY)
ICRC	3	3	2	2	10	Annual
CHF	8	7	6	5	26	Annual
Livelihoods (Cash & Cash for Work)	3	2	2	2	9	MY
NGO resilience	3	3	3	3	12	MY
UN resilience programme	6	6	6	6	24	MY
Nutrition (ACF / UNICEF / SCF)	4	3	3	2	12	MY
Aid Enablers	2	2	2	2	8	MY
Research, M&E	1	1	1	1	4	MY
<b>Total</b>	<b>30</b>	<b>27</b>	<b>25</b>	<b>23</b>	<b>105</b>	
Internal Risk Facility (IRF)	10	10	10	10	40	Annual
<b>Total With IRF</b>	<b>40</b>	<b>37</b>	<b>35</b>	<b>33</b>	<b>145</b>	

The Valid team have been in regular contact with the DFID Somalia humanitarian advisor and with Transtec who are implementing the monitoring and evaluation. Valid evaluations' director also sits on the advisory board for the Transtec exercise, and the ODI/Valid methodology expert is also a part of the Somalia extended team to try and ensure harmonisation between the two evaluations.

### 1.5 Evaluability assessments.

Three of the four evaluability assessments found that there was sufficient information, access and enthusiasm for the evaluation to proceed. The Yemen exercise did not materialise due to continuing security issues.

In each of DRC, Ethiopia and Sudan there are – as the portfolio analysis above shows – sufficiently diverse and relevant programmes for the evaluation to generate evidence. However the evaluability assessments concluded that the small sample size of projects would raise problems in attributing definitive impact directly to DFID-funded activities. This is particularly the case for the first evaluation question on whether individuals, households and communities were more resilient to shocks as a result of DFID funded interventions.

However, if impact can't be definitively measured, the evaluability assessments concluded that there would be sufficient evidence generated to reach some robust conclusions about impact, while for the question of value for money the conclusions might be quite comprehensive.

In all three countries, one or more of the DFID partners has a relevant – often substantial – research component included in their work. The most significant of these are:

- UNICEF DRC.
- Norwegian Refugee Council DRC.
- CRS-led ‘tadoud’ consortium Sudan.
- UNHCR Ethiopia.

All the above offers real synergy, and could potentially expand the evidence base of the evaluation, assuming methods can be harmonised to some degree (or at least are compatible). A fuller explanation of these can be found in section 3.2.2.

The wider DFID portfolios in all three countries can provide additional sources of data and evidence. Most obvious amongst these is the Productive Safety Nets Programme (PSNP) in Ethiopia, a multi-year programme entering its third phase. Whilst it is not ‘humanitarian’ in name, it targets populations at risk of humanitarian need. There are several comprehensive studies of the PSNP, and an ongoing longitudinal research effort led by the Institute for Development Studies (IDS). These could well provide valuable contextual information.

## 1.6 Pilot method study in Ethiopia

A small pilot was undertaken in Ethiopia in September 2014 to test Valid’s proposed methodology. It was not intended to generate evidence; rather it concentrated on qualitative data gathering and aimed to ascertain that the iterative questioning approach proposed elicited viable responses that could be analysed.

The pilot brought together a group of experienced research practitioners with some junior researchers to try and understand the right skills profile for this work. The team was multi-disciplinary for the same reason including backgrounds in health, nutrition, economics, agriculture and livelihoods. A key finding was that it will be important to have a range of skills in teams for the substantive research, as the blend of perspectives offers interesting insights. Not surprisingly, it was also of great benefit having more experienced practitioners in the team. The less experienced researchers were able to adopt this type of very open ended enquiry, but it was new and required time to adapt.

The team undertook both group and individual interviews, initially in equal measure. However it quickly became clear that individual interviews yielded much richer information, and latterly more emphasis was put on these. Group interviews were still useful however, largely in setting the context. In future, these may be further tailored towards this purpose.

Initial results suggest that using very open ended questions yielded rich and varied results. There were a number of key lessons, and questions, that were highlighted by the exercise however. These are now being factored into the next iteration:

- The training of the researchers needs to be better – longer and more structured. In the pilot exercise we took two days for training, and this was not long enough.

Having experienced researchers on the team partly countered this, but in future training should be for a week, especially where we envisage teams will see the work through for the duration of the evaluation.

- Training and profile of the interviewers is more important than in a normal survey. This kind of approach requires experienced researchers, and an intensive few days training are required to change a mind-set that tends to want to fill boxes.
- The results vary geographically, as livelihood strategies vary and wealth and access changes. This makes standardisation challenging, and means quantitative tools need to be designed for context.
- Data capture is also challenging as it requires significant time investment in recoding the interviews afterwards. The same holds true for analysis; for the pilot this was basically done manually, by having someone read through everyone's notes. For later rounds of interviewing this capture and analysis system will have to be more 'mechanised'; work on this is ongoing.
- The questions, and the method, will necessarily be iterative. After each successive round of interviewing we will have to set aside time to learn lessons and adapt the method. As a result initial rounds in all four countries will continue to have a 'pilot' feel to them.

The initial questions are set out below:

Box 1 (below) gives an indication of the questions used in the pilot.

#### Box 1: Guiding questions for the pilot qualitative field work

##### Problems people face

This is presented in a very open way, to capture both common and individual shocks.

1. What are the problems that people have experienced/ foresee? why can't they avoid them? who faces them and who doesn't?
2. What is the impact of these problems on different people?

##### Solutions

3. On what do they/ will they rely to get through hard times? (note: need to discuss BOTH coping and the avoiding crisis.)
4. When you're trying to get through hard times, what are you actually trying to do? (what are the core minimum objectives/plans people have for themselves/their families?)
5. What makes it more or less likely that you'll be able to use these 'solutions' to get by?

##### Assistance

6. What help do you get? can you rely on? (both from within the 'community ' and outside)

7. Terms and conditions for getting this help? [It will be really hard to get at any but the most formal conditions. I don't think we need to expect the pilot to throw much up, but we should certainly try. it'll need very 'sociological' interviewing, as terms and conditions are usually implicit, unrecognised. (Is it a term and condition of being invited to dinner that you bring a bottle/flowers? reciprocate? How often will you be invited back if you never do..?)

#### Recovery

8. What does 'getting through' look like? For different kinds of household crisis, how long does it take to get back to situation where you can say you have 'got through'? and to where you were before? What does that depend on?

Whilst the pilot was mostly concentrated on getting the method right, it did however throw up some interesting insights.

Chiro, or Aseba Teferi zone, is 600kms south east of Addis Ababa, towards Jijiga, in the middle highlands (2,000m). It was chosen for its accessibility, for its known vulnerability to nutritional emergencies, and because the team knew it well. It is a predominantly cash-cropping area (chat and coffee in particular), with staples grown for household consumption, bordered by agro-pastoral communities in the lowlands. Livestock is kept in small numbers, primarily for fattening and sale and there is a range of petty trade, commerce and day labour. The PSNP, food aid and other aid programmes operate in the area; labour for unemployed men on the PSNP tends to be soil conservation and similar basic tasks.

The pilot found that despite a potential increase in cash income from chat, an intensification of drought cycles (or, put another way, the unpredictability of rain, as much as the quantity falling) and land fragmentation has made people feel less secure than they have historically been. Chat production generates around Birr 4000 (\$200) per annum for an "average" household. However, land holding sizes have declined significantly over the last 30 years – one community proposing by as much as two thirds, meaning that one hectare is now considered a sizeable holding. Most people, as result, supplement this income with day labour, and in harder times petty trading, charcoal production, firewood sale and, increasingly, migration (for work). Whereas historically resettlement was not a subject for discussion, it is now seen as a real option (with conditions). Education is now seen as essential for personal improvement but, with poor employment prospects, a bad family investment.

### 1.7 Structure of the inception report

This inception report is structured in four broad sections.

- The first looks at the background and context to the evaluation.
- The second sets out the purpose and scope of the evaluation.
- The third section is the most substantial. It develops the approach to the three research and evaluation questions, and goes into some detail about the methods to be used.

- The fourth section sets out a timetable and work-plan, with key milestones and outputs.

## 2. Purpose and scope

### 2.1 Purpose, objectives and scope

The purpose of the thematic evaluation is to generate learning and evidence on whether and how a multi-year humanitarian funding approach has enabled DFID programmes: 1) to ensure timely and effective humanitarian response; 2) to build disaster resilience; and 3) achieve better value for money.

The evaluation has three core questions:

- Are vulnerable individuals and households more resilient to shocks and stresses as a result of the work of DFID’ funded interventions? How do investments in resilience contribute to/compromise delivery of humanitarian outcomes?
- Has the availability of pre-approved DFID funding linked to specific triggers enabled DFID to respond more quickly and effectively when conditions deteriorate? [this component of the evaluation will only be applied where emergency funds have been triggered]
- To what extent does DFID multi-year and pre-approved contingency funding provide better value for money than annual funding for DFID and partners?

### 2.2 Evaluation questions and matrix

The evaluation terms of reference set out a detailed set of sub-questions. This inception report proposes amending these slightly as a result of work done during the evaluability phase. The revised sub-questions are set out in an evaluation matrix, which is developed in full in annex 1. The original terms of reference, and the original sub-questions can be found at annex 2.

**Table 7: short version of the evaluation matrix, setting out the main questions and sub-questions.**

EQ1: Are vulnerable individuals and households more resilient to shocks and stresses as a result of the work of DFID funded interventions? How do investments in resilience contribute to/compromise delivery of humanitarian outcomes?	
Have DFID-funded interventions made people more resilient?	<ul style="list-style-type: none"> <li>• How well have DFID funded interventions addressed priority causes of vulnerability?</li> <li>• How far have the underlying logic models of the interventions given reason to believe that they can have an impact on resilience?</li> <li>• How successful have projects been at achieving their outputs?</li> </ul>
Has MYHF contributed to this?	<ul style="list-style-type: none"> <li>• How different are the uses of MYHF from 'normal' HF?</li> <li>• What are the constraints on using non-humanitarian funds for resilience building?</li> </ul>

How do investments in resilience contribute to or compromise delivery of humanitarian assistance? <sup>3</sup>	<ul style="list-style-type: none"> <li>• Has DFID MYHF helped build resilience?</li> <li>• Has any contribution to resilience reduced likely humanitarian needs?</li> <li>• Have humanitarian funds been diverted from meeting acute needs?</li> <li>• Has this reduced overall humanitarian response capacity?</li> </ul>
EQ2: Has the availability of contingency funding enabled DFID and its partners to respond more quickly and effectively when conditions deteriorate?	
Link between pre-agreement on triggers and response	<ul style="list-style-type: none"> <li>• Does planning pre-crisis result in a better understanding of which triggers are appropriate?</li> <li>• Does pre-agreement on triggers result in decisions being taken in accordance with triggers?</li> <li>• Agreed thresholds trigger response at an appropriate time, and at an early enough stage in the development of a crisis to permit more effective response.</li> </ul>
EW system is functioning well enough to permit response according to triggers	<ul style="list-style-type: none"> <li>• Quality of EW/surveillance</li> <li>• Speed of dissemination, clarity of analysis and conclusions.</li> <li>• Does EW give implications of non-action, for specific actions and in a specific time-frame?</li> </ul>
Contingency funds lead to investment in contingency planning	<ul style="list-style-type: none"> <li>• Is it possible to know (reliably) if triggers have been reached?</li> <li>• How trusted is EW for decision making?</li> <li>• Levels of preparedness at agency level, changes as a result of contingency planning?</li> </ul>
Decisions were taken in a timely way based on early triggers.**	<ul style="list-style-type: none"> <li>• Actual levels of preparedness (e.g. start-up times) are recognised and explicitly incorporated into decision making.</li> <li>• How much faster was contingency funding than other funding?</li> <li>• How far did actual response match that planned?</li> </ul>
Earlier analysis and earlier response lead to more effective responses	<ul style="list-style-type: none"> <li>• How far was analysis of crisis developed in advance?</li> <li>• If response was earlier, were more preventive response options considered?</li> </ul>
Scale of response facilitated by contingency funds	<ul style="list-style-type: none"> <li>• Is the scale of response funded by contingency funds sufficient to be relevant?</li> </ul>
EQ3: To what extent does DFID multi-year and pre-approved contingency funding provide better value for money than annual funding for DFID and partners?	
How far have MYHF funds actually operated as MY funding with MY perspectives?	<ul style="list-style-type: none"> <li>• Role played by DFID MYHF funds in humanitarian system</li> <li>• Ability of partner systems to incorporate MY perspective?</li> <li>• Ability of DFID systems to work with partner systems on MYHF.</li> <li>• Has there been a change in programming as a result of MYHF?</li> <li>• Has MYHF led to a change in programme management?</li> </ul>
Are programmes more effective as a result of MYHF?	<ul style="list-style-type: none"> <li>• Earlier response</li> <li>• Better quality in project design</li> <li>• More effective implementation</li> </ul>
Are costs lower as a result of MYHF?	<ul style="list-style-type: none"> <li>• Lower administrative costs</li> <li>• Lower operational costs</li> </ul>

<sup>3</sup> Note that this question has been changed from the evaluation question in the TOR which asked about the delivery of humanitarian outcomes. See below.

### 3. Approach

The evaluation will use an enquiry-oriented approach, rooted in the various theories of change for multi-year humanitarian funding and the country programmes. It will use mixed methods to gather and analyse data, generating its own primary data as well as substantially using secondary data from DFID partners and other sources.

In essence, the evaluation will seek to answer the three evaluation questions by following a ‘logic chain’ from the approval of multi-year humanitarian funding (MYHF), to the outcomes of various programmes.

This is captured in both a ‘composite’ theory of change, outlined below (figure 1), and for each of the three evaluation/ research questions.

Action: MYHF made available to partners			
Outputs			
Reduced operational costs for DFID and partners setting up programmes/contracts, and through efficiency gains	Interventions that aim to build assets, address structural challenges, reduce extreme vulnerability and build resilience in the mid-term become more common	More comprehensive and sustained dialogue with national and local authorities is recorded	Better access to beneficiaries and better feedback mechanisms are put in place
Partners report tailored admin processes (planning cycle etc.) for MY funds	Partners step up investment in developing staff capacity, esp. those likely to remain long-term i.e. national staff	Pre-agreed trigger indicators are set for contingency	Partners demonstrate more flexibility to changing conditions e.g. scale-up to meet higher need
Outcomes			
More resources routinely spent on monitoring, evaluation, research, and VFM tracking	Longer-term, tailored, & strategic programmes emerge	Partners advocate on contentious issues more confidently	Transition to development is smoother (due to relationships, advocacy, & programmes)
More timely interventions based on triggers	Improved financial management within DFID	More skilled and knowledgeable staff managing and implementing projects’ performance	

IMPACT: Higher quality, more relevant programming manages risk, addresses root causes, and ultimately shortens suffering			
Assumptions			
MY approval is passed through to end implementers	Funds used for long-term activities, not e.g. urgent gap-filling (contingency is for that)	Implementers invest in long-term relationships (even if staff contracts are short)	Partners can set up adequate finance/admin support for MY funds
Early warning data is sufficient to support triggers for contingency	Cost savings go to improve programme quality, not to other areas	Oversight of triggers does not slow action	
More M&E, research, and VFM raise quality	Frequency of reviews, level of optimism bias & remedial action based on findings are the same (or higher) as for annual funds	Implementers learn over time and programmes are adapted accordingly	

**Figure 1: composite theory of change for multi-year humanitarian financing.** Source: Claire Devlin (DFID). The MYHF TOC is an amalgamation of country based MYHF TOCs.

The importance of theories of change to the methodological approach proposed in this inception report by VALID is shaped by DFID’s request for a *thematic*, learning evaluation rather than an impact evaluation. The purpose of the evaluation is not simply to establish to what extent a certain portfolio of programmes has achieved impact: it is rather to understand how and in which circumstances DFID is able to change the way in which humanitarian aid functions, and the consequences of that for its impact, by making a change to its structural bureaucracy (i.e. the system of MYHF).

In other words, the evaluation has to explore and test the entire length of the causal pathways by which a move to MYHF makes a difference, establishing in which circumstances this is facilitated and in which circumstances it is constrained. The implication of this is that the logical structure of the evaluation itself is created not by the projects and their implementation but by the theory of change of MYHF itself. Since the range of issues raised by the ToC is extremely wide, no single methodological approach is possible for this evaluation. Different tools, ranging from primary qualitative interviewing with people affected by crises to financial analysis of partner accounts, will be needed. Coherence is brought to this methodological *smorgasbord* by the internal logical skeleton provided by the theory or theories of change, as detailed below.

An overview of the various elements that are combined in this evaluation is given in a comprehensive evaluation matrix, in annex 1. This outlines the different evaluation questions, the lines of enquiry that are necessary to answer them and how those lines of enquiry will be pursued. The following sections go into more detail for each evaluation question, detailing as comprehensively as possible the relevant impact pathway(s). The DAC criteria have been used as lenses to develop this analysis. The resulting *logic models*, which include any assumptions that are necessary for the impact pathways to work, will then be tested, link by link. The aim of the evaluation

is thus not simply to give an answer to the EQs but to increase understanding about how, why or in which circumstances DFID's theory of change has been realistic and where any constraints have been which have created weak links in the impact chain.

Sections 3.1-3.4 outline the logic models relevant to each of the three research questions, as they have been analysed by the VALID evaluation team in the preliminary work undertaken during the inception phase. (It is anticipated that these will be refined during the evaluation from our own analysis and from continued interaction with DFID, partners and others.) These logic models are followed by a description of the methodologies that will be appropriate for investigating the different links of the chains under each EQ.

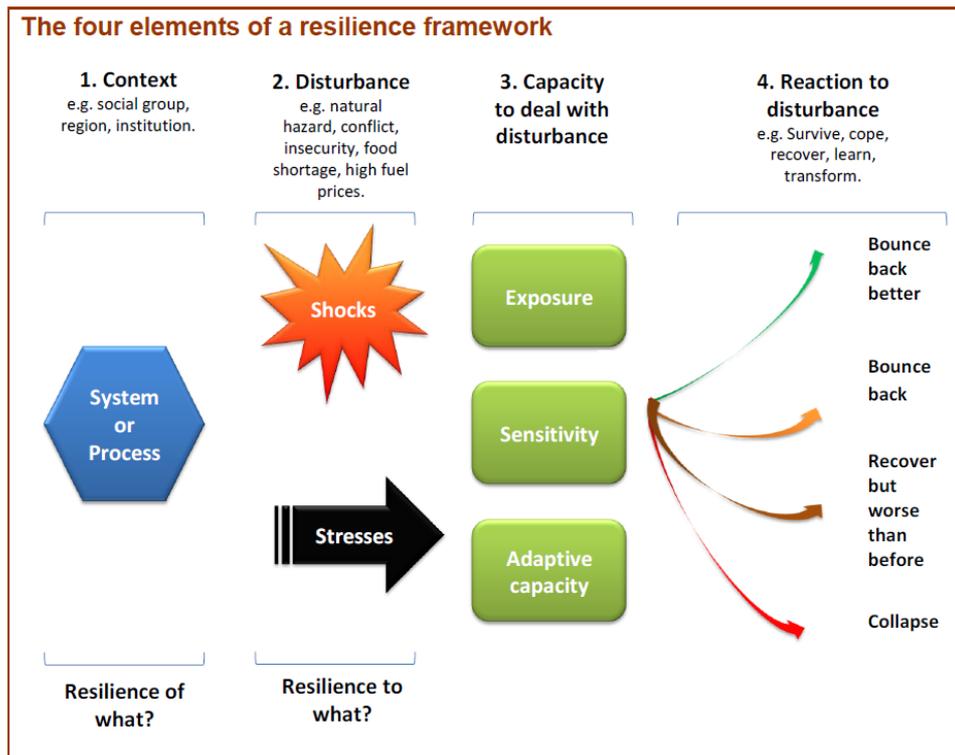
It is important to note that the evaluation concentrates exclusively on humanitarian issues, and in that context humanitarian disaster resilience, as a pose to the wider structural problems that communities undoubtedly face. These sorts of problems will invariably register in responses from interviewees, however DFID humanitarian programmes cannot be evaluated on whether they resolve these wider issues.

### 3.1 Evaluations question 1a: resilience

**Are vulnerable individuals, households and communities more resilient to shocks and stresses as a result of the work of DFID' funded interventions?**

#### 3.1.1 Logic model

The DFID disaster resilience framework, set out in figure 2 below, provides a framework for looking at the first part of question one, namely whether DFID programmes are contributing to people's resilience.



**Fig 2. DFID resilience framework.** Source: DFID 2011.

The resilience framework indicates four broad areas of enquiry to consider: 1) context, 2) disturbance, 3) capacity to deal with a disturbance and 4) reaction to disturbance. Where the objective is to assess the impact of aid on resilience, an essential fifth element must be added, that relating to the assistance provided by DFID partners. The resulting five areas of enquiry might be called:

- a) Context
- b) The problems people face
- c) The solutions, or how people cope
- d) Assistance
- e) Outcomes, including (hopefully) recovery

There are three relatively clear steps which the evaluation team must take in order to answer EQ1.

**Step 1:** The team must gain an independent understanding of the vulnerabilities (and how to address them) of different population groups in the different contexts where MYHF interventions are run.

Using the DFID resilience framework, this entails investigating areas of enquiry a)-c) above.

**Step 2:** The team must then assess the MYHF-funded interventions to see to what extent it is plausible to think of the projects as improving resilience.

This involves addressing areas of enquiry d) and e). This will be undertaken by matching the logic of the projects (their implicit theories of change, assumptions and resilience models) against what was learnt in step 1 about how people face their life challenges and opportunities.

**Step 3:** The team must then evaluate what has been the contribution of DFID funding to achieving these potential impacts on the resilience of people affected by crisis.

Specifically, the evaluation will focus on assessing the contribution of the bureaucratic change in humanitarian funding (to multi-year) to these impacts. The evaluation will also need to put the use of humanitarian resources in context, by including an appreciation of the potential role of both humanitarian and non-humanitarian resources.

### **Step 1: Understanding people's resilience**

Methodological approaches for assessing resilience are only starting to be developed: no consensus has yet been reached even on which parameters need to be considered and analysed, let alone the tools that are appropriate or valid for assessing them.

Resilience is essentially a latent capacity, i.e. the (current) ability to deal with a future difficulty. (Once any difficulty or crisis is over, the person or system being considered will have a new level of resilience – which relates to their ability to deal with the next, future crisis and not the crisis just past.) The extent of this capacity called resilience, or the degree of success in dealing with difficulty, can only ever be seen in retrospect, after its application. Unlike measurements of well-being, resilience can only be directly assessed by considering the comparison between a situation before and a situation after (or during) any difficulty.

Although there is a broad, but largely theoretically derived, consensus about the *kinds* of issues which “must be” (i.e. are assumed to be) important for resilience (e.g. access to services, assets, social safety nets, adaptive capacity), there is still a dearth of empirical evidence that explains exactly which elements people most rely in different situations. There is no evidence base for making reliable inferences about resilience from assessments at a single point in time of known characteristics or parameters, even if the volume of data held about people’s lives were far greater than is possible to achieve in the contexts under discussion. The evaluation thus has to first learn how resilience actually ‘works’ for different people in different parts of the case study countries (i.e. what parameters to study and how to analyse them), and then to assess how well it is working for them (i.e. assess resilience according to those parameters and the analytical frameworks developed). In particular, the evaluation must ensure a properly disaggregated analysis to explain what *different* people rely on for their resilience or the factors that are making them more vulnerable.

The evaluation will take a two stage approach to understanding local resilience; first defining the parameters of enquiry qualitatively, and then using the findings from qualitative enquiry to determine any further quantitative work.

- A first round of qualitative interviews will be used to determine the parameters of resilience: it is important at the outset to understand (as opposed to measure) the

dimensions of resilience for different people in the different contexts where DFID funded interventions are implemented. This includes an understanding of the problems and the context including the ethno-social dimensions. This means starting with very open ended questioning, based around the 5 areas of enquiry set out above. This somewhat more anthropological approach will also help ensure that questioning does not force respondents to construct answers using the concepts of the evaluation (e.g. 'resilience'), but rather using the concepts with which they construct their lives. Concepts such as resilience will be used as organising constructs in subsequent analysis of how people relate their lives' struggles and hopes.

- VALID will particularly investigate how far different people in different circumstances rely on similar or different factors, strategies or processes for their resilience; how far they even have the same objectives when faced with difficulties; and how far the forces or factors maintaining people in vulnerability can be addressed at individual or household level, rather than addressing causes of vulnerability which lie in processes and institutions which work at a very different level.
- Subsequent rounds of qualitative interviewing will build a more accurate picture: and in particular will help us understand how people cope over time with changing circumstances, giving more detail as to how coping works.
- Quantitative enquiry will be shaped by the qualitative results. Once specific vulnerabilities and constraints, or opportunities and features of resilience are well understood for specific population groups, it may be necessary to quantify how widespread such issues are, and how important the factors are in shaping resilience as a whole. Quantitative enquiry may also be used to explore factors associated with various parameters relevant to resilience, e.g. as possible underlying causes or avenues for addressing vulnerability, etc.
- VALID will also seek to use partner M&E and their own research for collaborative quantitative analysis. In some cases, this will be with data collected from existing M&E systems. In other cases, VALID will seek to work with the partners to influence the data being collected.
- VALID will continue with qualitative enquiry throughout the three years of the evaluation in order to understand how resilience changes for people and in order to understand the perspectives of the people affected by crisis on the 'resilience impact' of interventions.

Managing and analysing the qualitative information on resilience is expected to be the most challenging part of the whole evaluation. The quality of the analysis will depend upon the quality of the field researchers used, the training they receive, VALID's ability to ensure the quality of the work and VALID's ability to manage the information, in order to bring together a vast array of very different nuggets of information from different people in different places in order to see patterns and find insight.

Those engaged for the research will have to show experience and competence in social research. They will be trained for at least one week before going to the field in a custom-designed training programme, that will encourage them to think on their feet in interviewing rather than simply to follow checklists as questionnaires. They will be recruited as locally as possible to the study areas to maximise the cultural familiarity of the team with the context under study.

Researchers will be responsible for shaping how questions are posed. Members of the core VALID team (i.e. the international team and the senior national consultants) will work alongside them in the field as much as possible, in particular for the first few days of any research. Long debriefing sessions will be held each day that will focus on analysing the conversations/interviews held, rather than merely recording the information gathered. The local researchers will be encouraged to record answers in the interviewing language used and will be given the responsibility for interpreting these answers culturally and not merely linguistically. Rigorous questioning of the interviewees each day will help to ensure that interviews evolve each day to build upon findings as they are accumulated and to ensure that gaps in understanding are quickly identified and addressed. The use of feedback sessions based upon a team critique (rather than on reporting back) should also ensure that the quality of interviewing improves each day.

### **Step 2: Assessing the potential impact of DFID funded interventions**

One of the most important contributions which VALID aims to make during its evaluation is precisely to make explicit the (largely implicit) theories of resilience that underpin interventions and then to test them. Where projects have assumed that certain characteristics underpin or endow people with resilience, their success can only be assessed by including a challenge to these assumptions. If change in resilience were measured using these same characteristics, then evidence of outputs (giving people more of those characteristics) would be taken as proof of changed outcomes (more resilience). The use of such circular logic creates a situation where a project would be regarded (or would regard itself) as being successful in changing people's lives merely by delivering on its plans.

In keeping with the learning objectives of this thematic evaluation (which are much broader than those an evaluation designed to verify the extent to which certain projects reached their objectives), much of the analysis will be on the potential impact of the interventions based on their logic models. This will ensure that much lessons of much wider relevance are learned, since, taken in isolation, the results of impact assessments of each individual intervention would depend mainly on the partner's capacity to design and implement a project, and would yield little of interest about how the longer term perspective of MYHF had influenced impacts.

The evaluation will answer the following three questions about the DFID funded aid interventions.

- a. How do the findings of the evaluation about what is important for the resilience of different people (step 1) compare with what agencies are currently using in their planning, and in their M&E?
- b. How far are DFID partners addressing the identified issues in their programmes?
- c. How far are DFID partners contributing to change in the identified issues through their programmes? What trade-offs, if any, are people having to make to benefit from any progress?

### **Step 3: assessing the extent to which MYHF has supported resilience building**

This step focuses on how the working of the aid machinery has changed as a result of the introduction of a new kind of funding. The scope of the enquiry is thus actors in the humanitarian and non-humanitarian aid sectors (DFID, its partners and others) rather than on beneficiary populations.

The evaluation will look at the extent to which partners have changed their strategies, interventions, organisation and coordination, implementation modalities, presence, operational management and administration as a result of the availability of multi-year funding. The various factors which have enabled or constrained them to take advantage of the potential opportunities of multi-year financing will be identified and analysed.

Following this information gathering step, a subsequent stage of analysis will identify the changes that they have been able to make as a result of MYHF, and by combining this with the analysis in step 2 (of the potential contribution of the interventions to resilience), the team will assess how far the interventions have a greater or lesser potential to support resilience as a result of MYHF. It will also be necessary to put this potential contribution in context – situating it both within the context of the scale of need and in the context of broader (non-humanitarian) aid flows. It cannot be assumed, for example, that resilience is built incrementally. If resilience functions in some senses in quanta, and if there appear to be thresholds which must be crossed to reach a new quantum level of resilience, then the evaluation must try and assess the extent to which it is plausible to believe that the scale of interventions funded by MYHF can help people to attain a genuinely more resilient state.

Further to this, it will be very difficult to ascertain the degree to which DFID funding contributed to increases in resilience. In any given context or area, a multitude of actors will be operating, as well as external forces that contribute to or undermine resilience, and therefore attributing changes to DFID will be confounded. However, this evaluation is structured around a logic model, where each step in the logic is evaluated. Therefore, the intention of Step 2 will be to investigate the potential contribution of interventions to resilience more broadly, and this step will then link those interventions that contribute most to resilience to those interventions funded under MYHF under DFID funding.

### 3.1.2 Method

Section 3.2.1 detailed how the evaluation team will go about undertaking the three steps to answer EQ1a. This included a description of the role of primary field work at community level, the use of information collected by partners and the information to be collected through interviews with the staff of DFID, partners and other agencies. Section 3 began with a description of the overall programme theory approach to answering the evaluation questions, based on uncovering a variety of logic models – underpinning projects, the case for MYHF, the justification for contingency funds, etc. Evidence will be looked for to test each link of these logic models: the nature of the evidence and the methods and tools used to find it will vary from link to link.

Annex 1 gives the overall evaluation matrix for the study, detailing how the three evaluation questions are broken down into 71 sub-questions or ‘key issues’, and the different research methodologies that will be used for each key issue. Table 6 (below)

gives a summary overview of how the evaluation team will address the question of the contribution of MYHF to resilience, by giving a breakdown of the logical premises which must be true for MYHF to bring positive change to resilience building support, the evidence that will be looked for to find out the degree to which each premise is true and the research approaches used for each one.

**Table 8: contribution of MYHF to resilience and humanitarian outcomes**

<b>Logic model/impact pathway</b>	<b>Observable evidence</b>	<b>Research approach</b>
1. DFID partners have identified and correctly prioritised the causes of vulnerability to crisis of different population groups	Analysis with populations affected by crisis of vulnerability and sources of resilience (disaggregated)	<ul style="list-style-type: none"> <li>• Primary field work</li> <li>• Analysis of project documentation</li> <li>• Agency interviews</li> <li>• secondary literature</li> </ul>
2. a) DFID partners were unable to find funds to address these problems except from HF, and b) short-term HF did not allow DFID partners to address these underlying causes of vulnerability	Funds available in country for partners and others Uses of donor money in country Proposals and strategies developed by partners and donors	<ul style="list-style-type: none"> <li>• Analysis of previous interventions by partners and non-MYHF humanitarian interventions in country</li> <li>• Analysis of use of non-humanitarian funds in country</li> </ul>
3. Longer term funding, on scale made available by DFID MYHF, does enable partners to implement interventions with a realistic chance of successfully addressing some of these causes	Logic models of partner interventions are plausible for building resilience	<ul style="list-style-type: none"> <li>• Analysis of project documentation</li> <li>• Findings from initial primary field work</li> <li>• Information from partner M&amp;E</li> <li>• Ongoing primary fieldwork in project areas</li> </ul>
4. Partners use the funds in order to implement such interventions, and in ways conducive to successful impact (e.g. well targeted, good monitoring, flexible management, etc.)	Successful partner implementation of proposals (inc. adaptations to plans in light of changing circumstances and increasing understanding)	<ul style="list-style-type: none"> <li>• Partner M&amp;E reports</li> </ul>
5. The interventions are successful in meeting desired impacts	Project outputs lead to desired outcomes	<ul style="list-style-type: none"> <li>• Partner M&amp;E</li> <li>• Primary field work</li> <li>• Analysis workshops with partners at country level</li> </ul>
6. Implementing such interventions does not compromise the ability to respond to acute needs	Adequacy of actual humanitarian response for acute needs Use of MYHF in event of spike in acute needs Resources available to humanitarian sector in event of spike in acute needs Surge capacity of partners in absence of spike of acute needs	<i>See EQ 1b</i>

### Longitudinal panel research

The evaluation will include a longitudinal qualitative study to understand resilience by looking at how people cope with shocks and stresses and adapt to changing circumstances. Where possible, this will be based on a panel methodology, i.e. going back to the same households over time. (Displacement or insecurity may make it impossible to use a panel in some places.) The intention here is to get beyond the

‘single point in time’ issue, and track changes over time. Furthermore, the possibility of creating a rich dataset that can be accessed by others is seen as a desirable outcome.

Given the time taken to undertake the first round of interviewing, it is likely that the total length of time over which the panel will be followed will not be more than two and a half years. Although this is not a very long term cohort study following people over a lifetime, this period is believed to be enough to gain insights about the dynamics of change, especially since it is highly likely that some of those being followed will experience shocks of some degree during the lifetime of the study. The duration of the study is in line with some other longitudinal studies such as that used by the DFID funded Secure Livelihoods Research Consortium. If the approach proves valuable, it is to be hoped that others, most probably DFID’s partners working in the study areas, will continue to follow the panels as part of their regular monitoring in future programming. (The incorporation to project monitoring of such research approaches, especially going beyond interviews with actual ‘beneficiaries’, would be deemed to be a positive impact in itself of the VALID evaluation.)

It is estimated that VALID’s own qualitative enquiry will cover two geographic areas in each country. In each area, sixty people will be interviewed roughly every six months (subject to access, and with small modifications if demanded by people's seasonal availability). Initial interviewing would be guided by the questions in Box 1, and developments and changes in their lives would subsequently be explored in follow up interviews, together with any impacts they see from the interventions in their area. Since the use of the panel is purely for qualitative interviewing, and since the sampling will be purposive (in order to capture the range of diversity rather than to be representative, see below), there is no minimum ‘legitimate’ sample size. The scale of the use of the panel is dictated by time and resource constraints, but following sixty case-studies is in fact a very extensive piece of social research and a relatively ambitious target. Useful understanding would be expected to be gained by fewer than this. Informants would be selected individually, i.e. the same household member would be interviewed each time to gain understanding of that individual’s perspective. The objective of speaking to as diverse a range of people as possible, which guides the use of purposive sampling, will be extended to the intra-household sampling, which will be used to ensure that people who play a wide range of family roles (e.g. household head, wife, dependent children, etc.) will be heard.

It is hoped that this methodology of a ‘qualitative panel study’ will also be adopted by partner agencies, where they work in the field in contact with the populations they are intending to help on a regular basis. Where this occurs, the VALID team will work collaboratively with the partners to facilitate an overall analysis of the combined set of interviews.

### Sampling

For the qualitative enquiry, sampling will be purposive in order to capture as wide a range of situations as is considered useful. The exact prevalence of the different situations or population groups is not related to their importance for the evaluation (i.e. for understanding the issues), and it will not be inferred from the research. It is necessary that the research covers a range of livelihood zones and hazard profiles; and

that adequate disaggregation within villages/communities takes place. The exact way in which disaggregation is made will be developed as understanding is gained. It is obvious that the evaluation must speak to both men and women, and to people of a range of ages and of different socio-economic status. However, as the research progresses it will become clearer how far these groups need to be broken down (e.g. even excluding age, education and wealth as factors, how far do married women, single mothers, widows, divorced women and separated women share the same vulnerabilities and resilience?)

In quantitative research, it is critical that selection of interviewees is made only according to identified characteristics, and that unidentified characteristics do not play an unrecognised role in shaping the responses received. Although in a case-study approach it is still important that the researcher understands what each case is and so does not extrapolate the analysis inappropriately, the actual sampling methodology plays a much less critical role in this. As the research and analysis proceeds, the team will progressively understand the individual characteristics of each informant and of their situation that are important in shaping people's vulnerability and resilience. If a selection of informants has been unduly narrowed by unintended filters (e.g. only those near roads, in secure areas, confident to speak to foreigners, etc.) it is important that the role of these filters in determining outcomes is then identified and analysed. If this is successfully achieved, then the research will not be invalidated by 'bias', even if any lessons are restricted in their application.

Any quantitative work that arises from this qualitative enquiry will be undertaken where there is a specific need, e.g. to have a better understanding of the prevalence of certain issues, to look for correlation to test very specific proposed causal theories, or to scope possible further lines of enquiry. The choice of sampling and statistical procedures of this quantitative work will depend upon the objectives (e.g. sample sizes will depend upon factors such as the range of contexts in which the enquiry is deemed to be important, the estimates of variability which determine power calculations, etc.). It is not useful to try to predict in advance what parameters may need to be measured to gain this understanding. These will be discussed with DFID as the study progresses.

#### Discrete studies

It is likely that specific questions will arise during the course of the evaluation, where a discrete and narrowly focused study would be highly informative for the much broader enquiry of the evaluation as a whole. The evaluation will reserve resources to conduct one/two such studies in each year across the three countries (i.e. a total of 3/6 in 3 years). These studies would be limited to a total 20 days each, intended to produce short, focused reports that are of specific policy significance.

The choice of research topics under this budget line will be made in consultation with DFID country offices and partners. The topics will emerge from the research, where the answer to very specific questions would make a particular contribution to overall understanding provided by the evaluation. The studies would be designed to contribute better contextual understanding in which to set the overall evaluation research. Without additional resources, they would probably be too limited in their scope to provide the kind of definitive answers needed for DFID or its partners to

make a specific policy decision on a directly related topic. The kinds of issue which might be appropriate for such study (without in any way predicting which questions would actually be seen to be important) could include questions around the rights and situation of women in divorce or separation, resource sharing in a polygamous context, how priorities are made in resource allocation around child feeding, the role and function of a particular local institution, etc.

#### Review of literature and secondary data

Understanding the context is a key part of asking the questions in the most productive way. People are not necessarily forthcoming with information, especially when they perceive that additional assistance may be at stake. Understanding from agencies exactly what is being done is an important part of framing initial interviews, and it will be a standard part of the research to liaise with the implementing offices of agencies before starting enquiry.

The use of secondary sources of information will be vital for establishing an understanding of any trends in resilience in three ways.

- a) There is a growing body of work on resilience, and though this has largely been conceptual until now, it is anticipated that more and more empirical studies will become available during the course of the evaluation. The VALID team will need to keep abreast of any insights which emerge from resilience studies in different countries which may offer useful perspectives and leads for the countries in this study.
- b) It will be necessary to have a very deep understanding of each of the specific contexts and previous studies and research by others (supplemented where necessary by some interviewing of key informants) will be the main way of gaining this understanding.
- c) As already discussed, in some places there will be quantitative data already available, or created during the evaluation, to help understand trends, food security and vulnerability.

#### Collaborating with others

There are a number of studies ongoing that are either similar or have common components in all four countries.

The evaluability exercises have indicated that a number of DFID partners in all four countries covered by the valuation are either engaged in specific learning/research exercises, or are committed to learning more about resilience and the impact of their work.

- UNHCR Ethiopia: UNHCR have chosen Ethiopia, and Dollo Ado refugee camp in particular to be a global pilot study for multi-year humanitarian financing. The current UNHCR programme manager has been tasked with designing research and gathering evidence around the impact of MYHF, and several conversations have now taken place on possible collaboration. It is likely Valid will be able to shape the research questions with UNHCR and split the primary data gathering.

- UNICEF DRC: The Alternative Responses to Communities in Crisis (ARCC) programme, now in its second phase, is essentially a cash and vouchers scheme for IDPs. Within this, there is a budget for research into whether such approaches enhance people's resilience. Valid is part of the steering group for the research, providing potential integration.
- Norwegian Refugee Council DRC: the NRC programme in DRC looks to support individuals and households who have experienced multiple displacement, and takes an innovative, research-led approach. The programme sets out to look at how people cope with multiple displacement, and then reinforce their coping strategies. Although their computer-based modelling method is very different to the Valid research approach, there is nevertheless a lot of interest in collaboration. The NRC grant has one of the largest research components in the entire portfolio (£6m over three years), and as such may offer valuable evidence. Active discussion is ongoing between VALID and NRC on how to collaborate more fully.

Further, these DFID partners have demonstrated a considerable degree of openness and interest in a collaborative relationship with the thematic evaluation, extending to a willingness to share ideas and to cooperate in designing M&E and other knowledge generation systems.

There are other DFID-funded initiatives with which the VALID team has also established relationships.

- In Somalia, DFID has commissioned a separate but connected MYHF study. Findings from this study will be incorporated into the synthesis report at the end of this study. Whilst the team is not directly included in this evaluation, there are strong linkages. HPG is the learning partner with the NRC-led BRCiS consortium, and has a formal advisory role with Transtec (the company running the evaluation). This inception study and earlier methodology papers have been shared with Transtec and collaboration is ongoing, facilitated by the DFID research advisor.
- The Secure Livelihoods Research Consortium (SLRC) is conducting a longitudinal panel survey on recovery, looking at state services, perceptions of the state and livelihoods, and includes DR Congo as one of its eight study countries. The overall research project and survey is led by ODI, and the work in DR Congo by Wageningen University. One of the VALID team members has been involved in designing the survey instrument and working with Wageningen University research staff on SLRC research.
- CRS-led 'tadoud' consortium: In Sudan, the CRS-led consortium in Darfur is in the process of appointing Tufts as their research partner, and has already used TANGO for their baseline. Valid have had good initial conversations on potential collaboration, both with CRS and with Tufts University. Research methods appear to be closely aligned and VALID has already established a relationship of trust and there is mutual agreement to share progress, findings and analysis with the Tufts study team.
- DFID-Sudan Darfur stock-take work.
- A member of the VALID team is also linked in to the learning/research, monitoring and evaluation work of the DFID-funded BRACED programme,

which has a much wider geographical coverage, through cross-institution work at ODI.

There are a number of other initiatives with other donors with which the VALID team is and will be sharing ideas.

- HPG has been in dialogue with USAID's Resilience programme in Somalia (SomRep).
- In Ethiopia, the PSNP has an IDS longitudinal study attached. There is also a wealth of similar research commissioned by USAID, ECHO and DFID.

VALID international will work with DFID partners to maximise the extent to which their existing knowledge gathering exercises, in particular data collection (i.e. quantitative research), can contribute to a collaborative learning process. This learning should at a minimum lead to a shared, greater understanding on: the determinants of resilience for different people in different situations; how people make decisions about trade-offs when faced with constraints to their freedoms; how far external interventions can have lasting impact on resilience, and for whom.

VALID will use its own insights into resilience (through the qualitative enquiry) in its work with partners to ensure that the most relevant information is included in data sets and other information; will help the different partners across the different countries to achieve the best possible degree of compatibility in the information collected, given the different needs and situations within which information collection takes place; will help catalyse greater collaboration in the sharing of raw data and in the analysis of all the information from the various different sources. In helping to create greater collaborative learning and information sharing, VALID will not limit itself to working with DFID partners, if other agencies are interested in being a part for this shared effort, either contributing to information collection or information analysis.

The information which partner agencies will already be collecting, or which they may come to collect, is likely to cover many of the parameters associated with food security and its determinants. Partner agencies will also be collecting information specifically related to their intended outputs and outcomes as part of their M&E. Much of this work is going to be quantitative. In order to maximise synergies, for its own information gathering, VALID will therefore prioritise the in-depth qualitative investigation which will enable a much richer interpretation of the information collected by partners. Independent quantitative data collection will be used where either:

- a) There is a lack of data from partners, and such data is needed to answer the evaluation questions
- b) Where qualitative investigation throws up hypotheses or models which need to be quantified or tested through quantitative enquiry, which necessitates the complementary very and very targeted collection of specific data.

More insights on what shapes how resilient people are could also be gained by a much larger scale longitudinal data collection exercise of several years. This could generate evidence for more specific research leads into factors which can be seen to correlate

with people and households ability to maintain well-being (income, asset holdings, etc.) over time, including through hard times. This would, of course, be beyond the scope of this evaluation but VALID will explore possibilities to help establish such a collaboration and to use the learning from its own research to inform the data collection.

## 3.2 Evaluation question 1b: humanitarian response

<b>How do investments in resilience contribute to, or compromise, delivery of humanitarian assistance?</b>
--

### 3.2.1 Logic model

The impact of *humanitarian* attention to resilience on emergency response cannot be measured directly, but, in keeping with the overall approach of this evaluation, a number of discrete questions can be identified whose answers will go a long way towards answering the evaluation question.

On the one hand, there is a proposition that investment in resilience may contribute to the delivery of humanitarian assistance in two ways:

- a) a longer term approach and presence will improve the ability of agencies to deliver more timely and more effective humanitarian assistance in times of acute need;
- b) investment in resilience will lead to a reduction in people's need for humanitarian assistance even when subject to shocks and stresses.

Against this there is a counter-argument:

- c) investing humanitarian resources (time, attention and funds) in non-acute needs by using them to address structural causes of vulnerability will reduce the availability of those same resources for acute emergency response.

The first question (a) will be addressed in looking at EQ3. The second question (b) will be addressed in looking at EQ1a. Although the three threads will all be brought together analytically to answer the overall question (EQ1b), this section here only treats the additional informational requirement posed by the counter argument (c) above.

For (c) to be the case, there are a number of assumptions implicit:

- i That people do not cope in a given emergency (i.e. that resilience investments have not resulted in lessened humanitarian need).
- ii That resources have actually been diverted (from acute response to resilience).
- iii That the 'humanitarian' response has an impact on acute needs (because if it doesn't then diversion of funds is moot).

### 3.2.2 Method

The methods will be chosen in order to assess each of these three assumptions. This will allow for an analysis of the contribution/ compromise of resilience investments. If it is found that resources have been diverted, and the response is both less than previous responses with the result that people suffer more, then it seems logical to surmise that such investments can, in certain cases, do harm. Conversely, if there is no ‘diversion’ of funds, the change cannot have this effect (and if people cope better than before then there is a positive contribution). Table 7 summarises this approach and sets out evidence to be gathered.

**Table 9: resilience building impact on meeting acute needs**

	<b>Methodological approach</b>
1. Addressing longer term needs reduces acute needs on an adequate scale and time frame.	<ul style="list-style-type: none"> <li>• Evidence from EQ1a</li> <li>• In case of acute spike in humanitarian needs, primary field work</li> </ul>
2. Using MYHF, acute needs can be met in ways that at the same time support resilience in the longer term.	<ul style="list-style-type: none"> <li>• Evidence from EQ1a</li> <li>• Analysis of actual emergency response: agency staffing and capacity, funding flows, proposals.</li> <li>• Interviews at agency level</li> </ul>
3. There is sufficient flexibility in the management of resources to permit them to be prioritised for acute needs as and when necessary without (entirely?) prejudicing the chances of success of the resilience building work.	<ul style="list-style-type: none"> <li>• Stakeholder interviews with agencies.</li> <li>• Evidence from EQ1a.</li> <li>• Analysis of resource flows both from DFID and within agencies.</li> </ul>
4. Resources have been clearly diverted from acute response to resilience building.	<ul style="list-style-type: none"> <li>• Analysis financial flows.</li> </ul>
5. Reduced DFID resources for acute humanitarian response leads to decrease in help for those in acute need/ suffering.	<ul style="list-style-type: none"> <li>• Agency M&amp;E.</li> <li>• Independent studies/ evaluations.</li> <li>• Evidence from EQ1a.</li> </ul>

Humanitarian outcomes are theoretically connected to resilience, if one uses the DFID framework as a starting point. Resilience is precisely that quality, or latent capacity that allows people to cope with crisis, or to be confounded by it. Understanding what actually happens in crisis – as the research set out above sets out to do – should also give rise to understanding about how people find themselves unable to cope and what then happens.

### 3.3 Evaluation question 2: Contingency preparations and earlier response.

The original evaluation question in the TOR was “*has the availability of pre-approved DFID funding linked to specific triggers enabled DFID to respond more quickly and effectively when conditions deteriorate?*”. This refers specifically a) to

pre-agreed triggers linked to the release of pre-approved contingency funds and b) to faster response only by DFID, and not by its partners.

VALID proposes to change this evaluation question for two reasons. Firstly, since it is not clear how far DFID contingency funds will be tied to pre-agreed triggers over the next three years in each of the four study countries, it is more useful to derive broader learning from the use of contingency funding in the different ways in which they may be established. In fact, there are several different models of contingency funding, with funds being held in different places, and with different rules. Funds which are flexible enough to be available to be used in different ways as needs change are also a form of built-in contingency, and this study will try to understand the various strengths and weaknesses of these different arrangements. Secondly, contingency funding is intended not only to speed up the bureaucratic response of the donor, but also of the response system as a whole. Part of the logic of making contingency funding available should be that partner (or potential partner) agencies can use their knowledge of the existence of contingency funds to prepare faster and more appropriate response.

VALID therefore proposes to treat the question more broadly than phrased in the TOR, using the following evaluation question.

**Has the availability of contingency funding enabled DFID and its partners to respond more quickly and effectively when conditions deteriorate?**

### 3.3.1 Logic model

Late responses to deteriorating situations, especially to slow-onset crises, have been frequently identified as a problem that, it is argued, has prevented the use of early interventions, which could prevent a difficult situation from becoming a humanitarian crisis. As a result, the range of response options is much more limited, response is much more expensive than it needs to be if earlier action were taken, unnecessary suffering is caused, and, even when emergency interventions prevent mass mortality or tragedy, many people are left much more vulnerable to future crises.

There have been attempts in several countries to speed up the timeframe for emergency response by the creation of contingency funds which can be called upon rapidly when certain conditions exist. By having mechanisms that sets aside funds in advance, establishing a clear set of criteria for their release and putting in place mechanisms for the speedy release of such funds (whether this is pre-approval in agreed conditions or a fast-track decision-making procedure), it is believed that valuable time can be saved in financing response. This, it is believed, should lead to earlier response.

The attention which contingency funds could bring to contingency planning as whole could be expected to have a much larger impact on the speed of response: the planning process (deciding what kinds of response would make sense when) ought to lead to much better contingency planning overall, and as a result, to better levels of preparedness, which have the potential to reduce the time taken from decision making to response on the ground by three or more months (Levine et al 2011). Response should also become more appropriate, both as a direct result of response being earlier,

and because of the attention and analysis given pre-crisis to how response could be triggered and the kinds of response that could be funded.

Where contingency funds have been created, possibly including the pre-approval of triggers for the release of the funds, their contribution to faster and more appropriate response can be evaluated in two ways. Where a crisis occurs during the time-frame of the evaluation, it will be possible to assess the impact of these contingency funds, with much useful information being gathered in real time or close to real time. Where such funds are not used, it will still be possible to learn much from their creation: the very logic of contingency funds is that time invested before a crisis brings benefits after one strikes. This means that even when benefits cannot be directly assessed because there was no application of the funding instruments, much learning can take place by examining what was or was not done before (or in the absence of) any crisis.

Whether the evaluation is looking at the actual use of contingency funds in any country or assessing the potential contribution of measures taken pre-crisis, the same 'logic model approach' will be used. This is an approach by which the implicit causal or logical chain by which the funds can lead to better response is identified in detail, and each individual link of the logic chain (see below) is then investigated. This approach does not attempt to find a statistical correlation between improved outcomes and the use of contingency funds or pre-agreed triggers or to measure the size of impact that can be attributed to the contingency funds. Instead, because the most appropriate tool can be chosen for learning about each individual link in the logical chain, it permits the use of a wide range of tools and a much richer learning about how to improve the timeliness of emergency response. Learning is not limited simply to the use of the funds, but a rich understanding can be gained touching on many of the factors that determine the speed and nature of response (e.g. how early warning affects decision making, how bureaucratic procedures influence delays, etc. See table 9 for a fuller description.)

The impact chain and associated conditions which lead from contingency funds to earlier and more effective response are as follows:

- a. A system for using the contingency funding is established in advance
- b. The system can respond to triggers early enough in the crisis calendar to enable early response.
- c. A functioning system that can provide early warning (EW) or surveillance is in place
- d. There is improved contingency planning, including planning linked to the criteria by which the funds can be released.
- e. There is an increase in investment in preparedness, including adequate linkage to the contingency plans and contingency funds.
- f. Decisions are taken in a timely way and sufficiently on 'objective' (rather than political) criteria, including an appropriate basis in EW.
- g. Decisions about response are qualitatively different

### 3.3.2 Method

The evaluation methodology for EQ2 is much simpler than for EQ1 because it relates purely to the response of the actors in the emergency response system.

Learning about this evaluation question is about how the practice of agencies and of the emergency system as a whole can improve as a result of a bureaucratic change in how resources are made available. The question does not relate to the quality of the actual implementation of any specific intervention and so the impact of emergency interventions is not the focus of the evaluation research. This part of the evaluation will look beyond DFID's MYHF partners, and will include in its remit the whole range of actors relevant to the part of the response system for which any DFID contingency funds are relevant. This includes governmental, UN and non-governmental agencies: those responsible for early warning, analysis, decision makers, donors, coordination and implementing agencies.

These actors first have to be identified. The logic model or impact chain (see Table 8) is, as for other EQs, then a fairly straightforward basis for the evaluation. In some cases, links in Table 8 are *necessary* conditions for the logic model to work; in other cases, they are secondary pathways that could bring about *additional* impact. For each of these logical links, Table 8 also describes the evidence which can be assessed by an evaluation in order to see how far expected and desired changes are actually occurring. The research approach to answering the questions will largely be based upon interviewing of the staff of the range of agencies identified as crucial (see above) and reviewing documentation. The nature of this interviewing and analysis is described in detail below.

The interviews and documentary analysis will be informed by a preliminary review of literature to develop an appreciation and critique of the state of the art thinking on early response. The study of country level documentation will then include: contingency plans (national/state plans, those of individual humanitarian agencies, clusters/sectors, etc.); early warning reports (internal, if available, and external communication) and situations analysis; actual emergency proposals submitted for funding; minutes of coordination meetings and of any other forums for analysis.

An 'early warning' function does not necessarily have to be formally called as such, and systems called 'early warning' do not necessarily provide warnings that are early. The question relates to the function of making sure that decision makers at the appropriate level are aware of what is on the horizon with sufficient notice to be able to make their decisions. Similarly contingency funds may not always be labelled as such: what will be studied is the speed with which funds can be used for unexpected needs. Where partner agencies have sufficient flexibility in their funding arrangements to move money to increase spending on unforeseen needs, such flexibility may function as form of contingency funding. The evaluation will examine the functioning in practice of the different arrangements which permit operational agencies to deal with unforeseen needs.

Although the use of pre-determined triggers is only one mechanism for managing such funds, it is one that needs particular attention. It would be naive to assume that only 'objective' criteria relating to needs are relevant in decision making on resource allocation. Indeed, the more decisions are about early response, and thus rely to some extent on predictions, the less they can be divorced from broader political considerations. The situation is complicated further in places where the distinction between chronic and acute need is hard to draw, where the scale and urgency of

chronic needs may be just as great as those created by a discrete acute crisis and where resources for addressing chronic needs are insufficient. In other words, there will always be pressure to use any contingency funds even in the absence of a specific new ‘crisis’, and it is likely to be easy to ‘justify’ the use of contingency funds for meeting chronic needs. On the other hand, the more funds are tied to very precisely drawn ‘objective’ criteria, the less useful they will be for any crisis that does not exactly follow a predicted trajectory and the more difficulty there will be in using them to support response which is genuinely early enough. Understanding how the politics of decision making play out will therefore be an essential angle for understanding how best to improve funding for early response: incorporating such a political lens will of course make it less likely that a simple answer on the ‘best’ technical arrangement can be provided.

Preparedness will be audited<sup>4</sup> by creating a Gantt chart with each agency for each proposed intervention in their contingency plan. The Gantt chart is constructed by first identifying all the separate actions and steps which need to be taken between decision making and response, and then estimating with the agency how long each step would realistically take, given the agency’s current state of readiness and preparation. Improvement in preparedness will be measured by the change in the start-up time that can be demonstrated to have occurred as a result of actions taken by the agency. In order to establish how far contingency funding influenced or contributed to any improvement, this quantitative analysis will be supplemented by qualitative interviewing to explore the reasons for any changes. This investigation of the reasons for change will apply to all the evidence detailed in Table 8.

In order to avoid drawing an unduly general lesson from over-attribution regarding any positive changes, the analysis will also explore the extent to which any of these changes could have been brought about without specifically committing contingency funds or pre-agreeing triggers or other criteria for the release of pre-approved funds. The answers to this latter question will come partly from an analysis of delays in previous responses in the country, and through a participatory exploration of the various specific constraints and measures that could be taken to address them. This exploration may be particularly important in achieving much wider change in the sector even where the specific arrangements being evaluated cannot be replicated (e.g. for other governments and donors).

---

<sup>4</sup> For further details on preparedness auditing, see COMESA-CAADP Technical Briefing Paper 1 (May 2009), *The Use of “Preparedness Auditing” to Strengthen Contingency Planning*.

**Table 8: Evaluation Framework for ‘contingency funding will lead to earlier and more effective response’**

<b>Logic model/impact pathway</b>	<b>Observable Indicator</b>
<p>1. <b>A system for using the contingency funding is established in advance</b>, when there is time to consult and analyse.</p>	<p>a) Existence of agreement on new or altered criteria for releasing funds for emergency response as a result of the change in funding policy / practice.  b) Quality of the process by which the criteria, including any trigger-thresholds, were agreed.</p>
<p>2. <b>The system can respond to triggers early enough</b> in the crisis calendar to enable early response.</p>	<p>a) Change in timing of ability to release funds, relative to the development of typical crisis (compared to previous practice).  b) Where triggers are agreed, they can permit response in the windows of opportunity for different interventions, taking into consideration the predicted delay between agreement on releasing funds and action materialising on the ground.  c) Funds will not be released if crisis is unlikely. The decision on how likely a crisis ought to be to justify use of contingency funds is subjective. Evidence will relate to donors (and other decision-makers on the release of funds) having a analysis of the development of a typical crisis calendar such that the release criteria reflect where they wish to maintain the balance between acting early and reserving emergency funds for genuine emergencies.</p>
<p>3. <b>A functioning EW or surveillance system is in place:</b></p> <ul style="list-style-type: none"> <li>• It is possible to know with reasonable reliability how likely a crisis is becoming; when the windows of opportunity for different interventions will be; and how prevailing conditions and predictions relate to the criteria for releasing funds.</li> <li>• The EW or surveillance system must translate the information it collects into predictions, conclusions and/or recommendations that are understandable and trusted by decision makers.</li> <li>• EW must achieve this in timely way.</li> </ul>	<p>a) Functioning of early warning or surveillance system in respect of speed of information flow, reliability, geographical coverage and coverage of the range of parameters needed for decision making, including triggers.  b) EW/surveillance information is being analysed, and is disseminating predictions about likely scenarios with an appropriate degree of certainty to the necessary audience.  c) Dissemination takes place very quickly after the information is collected.  d) Decision-makers regularly read, understand and trust the predictions, conclusions and recommendations that are given.</p>
<p>4. <b>There is improved contingency planning, including planning, linked to the criteria by which the funds can be released.</b></p>	<p>a) Evidence that contingency planning has improved as a result of changed funding policy/practice – considering both plans (i.e. documents) and planning processes;  b) Plans are for more appropriate and earlier interventions than previous practice  c) Interventions have been timed at the relevant period in the livelihood and crisis calendar and</p>

	are clearly linked to triggers that could realistically result in them being implemented at the appropriate times.
<p><b>5. There is an increase in investment in preparedness</b>, including adequate linkage to the contingency plans and contingency funds.*</p> <p>In the event of a crisis, <b>start-up is as efficient as expected</b>, permitting response which is genuinely earlier, and meeting windows of opportunity**</p>	<p>a) The estimated time taken from decision making on emergency response to relief arriving on the ground has shortened.</p> <p>b) Shorter start up times relate to interventions relevant to contingency funds and their criteria.</p> <p><i>In the event of funds being used:</i></p> <p>c) Measurement of the time actually taken between decision making for early response and relief arriving for people affected by crisis, compared to previous practice and state of preparedness.</p> <p>d) Comparison of timeliness of actual intervention on the ground with that promised in proposals and the timeliness demanded by the logic of the intervention itself.</p>
<p><b>6. Decisions are taken in a timely way based on EW and agreed criteria.**</b></p> <ul style="list-style-type: none"> <li>• effective early response is taken on the basis of the agreed criteria (e.g. triggers)</li> <li>• reliance on pre-designated criteria does not lead to a situation where a looming crisis is ignored because those specific conditions were not met.</li> </ul>	<p>a) Time actually taken from warnings being given until decisions taken for action</p> <p>b) Explanations for any delays</p> <p>c) Comparison of time lag from EW to decision making to time-lag in previous crises or in crises in countries where such pre-approved contingency funds have not been agreed.</p> <p>d) Evidence from written reports and testimony of an appropriate level of analysis of the unfolding of a crisis and that discussion around early action was not focused solely on the pre-determined criteria.</p>
<p><b>7. Decisions about response are qualitatively different</b> as a result of being early/ better prepared for.**</p>	<p>a) Comparison of kind of interventions funded and implemented compared to previous practice.</p> <p>b) Evidence that a wider range of interventions, including those capable of protecting livelihoods, have been implemented at appropriate stages of the livelihoods calendar and of the crisis.</p>

The evaluation will gather data and information on early warning, contingency planning, disbursements and the other range of material identified in table 8 throughout the implementation period. It will also use a staggered approach to look in depth country by country. Currently there are two countries that have funds specifically labeled contingency – Sudan and Ethiopia. As an interesting comparison, in Sudan the funds are held by DFID and in Ethiopia by the agencies. This should provide some interesting lessons about pros and cons of these approaches.

Information gathered through the inception period seems to suggest the contingency funds being used regularly. This may suggest regular acute, unforeseen emergencies; or that such funds are partly used to fill gaps. They important questions relate both to the ‘political economy’ of decision making (see above) and to the ‘bureaucracy’ of response. The former has already been discussed: the latter will also be the object of the study, recognising that bureaucracy and politics cannot be completely separated. The study will trace as precisely as possible the timing of each decision and activity by the chain of actors from donor and Government down to actual response on the ground. The reasons for decision and the reasons for the time taken by each step will be explored with the actors in as near to real time as possible. (It is recognised that in a major emergency, it may not be appropriate or possible to be asking questions where these can adequately be answered or analysed shortly afterwards.)

The VALID evaluation team will produce a short country-level report on lessons learned and any recommendations for future action to be incorporated into formative evaluation reports, but also for more immediate use by country offices. In the event of the Evaluation team analysing actual response during a crisis, any lessons which could be of immediate relevance will be communicated at the time directly to the DFID country office, prior to the preparation of any report.

#### Review of literature and secondary data

Because EQ2 relates to the specific changes brought about in each country, the role of secondary literature will be more limited in answering EQ2, except where early response and contingency preparations have recently been or will be a subject of study in one of the study countries. The literature will be used to understand what changes have been, or are theoretically, possible in order to look specifically at these changes in the study countries.

### 3.4 Evaluation question 3: Value for money

<b>To what extent does DFID multi-year and pre-approved contingency funding provide better value for money than annual funding for DFID and partners?</b>
---

#### 3.4.1 Logic Model

DFID recently commissioned a paper on multi-year humanitarian funding<sup>5</sup> to establish the evidence base that already exists on the value for money (VFM) of such an approach. The paper found that multi-year humanitarian funding and/or contingency funding has the *potential* for numerous benefits:

<sup>5</sup> Cabot Venton (2013). “Value for Money of Multi-Year Approaches to Humanitarian Funding.” DFID

- **Predictability** of funding allows more strategic partnerships and better planning, and can facilitate the choice of the most appropriate interventions as well as cost savings from making long term investments, leverage of additional funds, pre-positioning of stocks and pooling orders.
- **Flexibility for early response** – Agencies can react more appropriately and/or quickly to changing conditions, resulting in reduced caseloads, levels of needs, and loss of life.
- **Lower operational costs** – multi-year funding can result in decreased costs of aid, for example through reduced procurement, reduced staff costs, savings on proposal writing and reduced currency risk.

However, the paper also found that value for money varied depending on the type of crisis and the type of response. Numerous data gaps exist.

This proposed methodology builds on the multi-year scoping paper, by using the framework it developed for investigating the different changes brought about and their impacts, and then going on to address some of the data gaps for each of the country studies. This study will have the scope and time frame to work with a broader range of partner organisations in country, allowing it both to verify the findings from the scoping study and to refine them as necessary.

VfM is made up of three key components, referred to as the 3 e's, and these are described below:

- **Economy** relates to the price at which inputs are purchased – staff time, goods and services, etc. Economy in procurement will be particularly important for humanitarian response and multi-year financing.
- **Efficiency** relates to how well inputs are converted to given outputs.
- **Effectiveness** relates to how well outputs are converted to outcomes and impacts (e.g. reduction in poverty gap and inequality, improved nutrition, reduction in school drop-out, increased use of health services, asset accumulation by the poor, increased smallholder productivity, social cohesion). **Cost-effectiveness** analysis measures the cost of achieving intended programme outcomes and impacts, and can compare the costs of alternative ways of producing the same or similar benefits.

Ultimately, a VfM analysis requires a comparison of costs against outcomes. In the case of the evaluation of MYHF, the following two tables describe the logic model related to the both outcomes and costs.

**Table 10: Logic Model: MYHF leads to better outcomes**

Category	Evidence Sought
Changes in programming	<ul style="list-style-type: none"> <li>• Better analysis - partners have more time to study the context more carefully; this analysis is used in programming.</li> <li>• Development of longer term relationships with the same population groups, leading to more participatory approaches</li> <li>• Projects can learn, and evolve or adapt over a longer time horizon, permitting more effective strategies</li> </ul>

Earlier response	<ul style="list-style-type: none"> <li>• Increased preparedness leads to earlier response</li> <li>• Earlier response leads to the use of interventions that are more effective</li> <li>• Early response requires less support as asset depletion has not declined significantly.</li> </ul>
More effective implementation	<ul style="list-style-type: none"> <li>• Better trained staff, less staff turnover, particularly for national staff.</li> <li>• Better monitoring</li> <li>• Better relations with local populations.</li> <li>• Greater flexibility</li> </ul>

**Table 11: Logic Model: MYHF leads to lower costs**

Category	Evidence Sought
<i>Lower administrative costs</i>	
Staff costs, e.g. due to less recruitment	Staff turn-over and recruitment; total number of staff required
Proposal writing and reporting. Time required may be higher for multi-year proposals but lower because of fewer proposals.	Time invested in proposal writing and reporting. (Note: the ‘bureaucratic’ time of proposal writing and reporting will need to be separated from the ‘productive’ time of analysis and making decisions about programming. This will inevitably be inexact and fairly subjective.)
Improved currency conversions	Savings made from greater control over timings of transfers
Leverage of additional funds from guarantee of longer term funds in place	Quantity of additional funds raised
<i>Lower operational costs</i>	
Adoption of more cost-efficient strategies	Cost savings by using improved strategies to achieve (at least) the same outcomes.
Economy/efficiency savings (cost per input/output) due to better procurement plan and implementation.	Cost savings due to ability to pre-position stocks, changes in procurement costs, changes in transport and logistics costs.

### 3.4.2 Methodology

VFM is a function of both outcomes and costs, and can be achieved through:

- improving outcomes for the same costs; and/or
- lowering costs to achieve an equivalent outcome.

The ability of MYHF to lower costs is easier to document in quantitative terms. This can allow for a cost efficiency analysis – in other words, for a given output, what is the cost of MYHF as compared with annual funding.

The ability of MYHF to lead to better outcomes is much harder to quantify. The potential outcomes are many, covering a full range of livelihood outcomes that are influenced by many different factors. Outcomes and programme 'value' include wider, and possibly unforeseen and unintended impacts beyond those contained in specific objectives, many of which may not be seen immediately. However, there are nonetheless many examples in the literature that suggest improved outcomes, that we may be able to draw upon even if context/sector specific.

As a result, the evidence base will be composed of two components:

- Qualitative assessment of the logic model using interviews with stakeholders based on the questions listed above, as well as data from the longitudinal study around qualitative changes in people's ability to cope as a result of better programming, etc.
- Quantitative VfM analysis of specific examples where MYHF led to a specific intervention that has documented improved outcomes. These will be identified on a country by country basis, depending on the availability of robust evidence relevant to this evaluation. At a minimum, analysis of reductions in costs will be documented, but these will also be offset against changes in outcomes where this linkage can be made.

1) In-depth partner interviews. The first step will be to gather in-depth data for each MYHF partner on the use of DFID multi-year humanitarian funds and the systems being used to record this. Implementing partners will be interviewed to identify how they are using the multi-year funding, as well as their reporting systems.

2) For each country, depending on the type of hazard and the specific details around delivery of multi-year funding, a data collection framework will be developed, for both qualitative as well as quantitative data collection. This framework will very much depend on the findings from Step 1.

Nonetheless, broadly speaking, in each country, evidence will be collected around:

- a. Administrative cost savings;
- b. Operational cost savings; and
- c. Outcome savings (e.g. improved outcomes such as reduced morbidity and mortality).

Data on administrative and operational cost savings due to multi-year humanitarian response will be collected from DFID and its implementing partners. The approach will very much focus on agency level analysis and data gathering. Cost efficiency data will be collected by comparing changes due to multi-year by comparing what agencies have done in the past, as well as what they are doing now with annual funding from other donors.

The evaluability assessments suggest that obtaining the data needed will not always be straightforward, due to the complexity of internal systems for some of the bigger agencies. However, it is also the case that agencies are generally keen to prove the (assumed) benefits of multi-year funding, and therefore are likely to be open and collaborative in data sharing. It will only be once the team has started to request

detailed information that it will be possible to see whether systems are configured to provide this (and if not what level of work it will entail for agencies to gather additional data).

It is also likely that the quality and consistency of the data will vary significantly by partner organisation. In order to mitigate this risk, the Valid team will work closely with key implementing partners to assimilate the data in a consistent format, particularly working with some of the larger agencies that may have more standardised reporting.

Further to this, as the evaluation progresses, the team will gather and report on the types of VFM data that are collected, and the formats used, as part of VFM indicators linked to DFID annual reviews.

Based on the consultation for the scoping paper on cost effectiveness of multi-year humanitarian response, the following categories of cost savings were identified and these will be investigated with DFID and implementing partners. A request for data will be developed that outlines the key categories and used to facilitate a discussion around potential cost savings.

**Table 12: Cost Data Requirements**

Administrative cost savings:	Operational cost savings:
<ul style="list-style-type: none"> <li>• Reduced staff costs;</li> <li>• Reduced proposal writing;</li> <li>• Improved currency conversions; and</li> <li>• Leverage of additional funds.</li> </ul>	<ul style="list-style-type: none"> <li>• Pre-positioning of stocks;</li> <li>• Changes in procurement costs;</li> <li>• Changes in transport and logistics costs;</li> <li>• Cost efficiencies from greater long term investment and planning; and</li> <li>• Efficiencies from greater flexibility to choose most appropriate interventions.</li> </ul>

Some of these categories will be more readily available, for example staff costs, proposal writing, currency conversions, pre-positioning and lower costs for procurement, transport and logistics, should be fairly easy to document. The team will work with implementing partners to identify categories of cost depending on their specific programming, and develop mechanisms to estimate what these items would have cost without and with multi-year humanitarian funding. This exercise can use historic data for comparison purposes, as well as estimates based on existing programming needs.

Other categories will be harder to document, for example leverage of other funding, greater long-term investment and planning, and ability to choose the most appropriate interventions. It is likely that these categories will require a more qualitative evaluation, using interviews with implementing partners, on what they would have done under annual funding, and how this has changed with multi-year funding. Having said this, where partners can identify specific activities that could only happen with multi-year funding, a quantitative evaluation of the cost of response under both annual and multi-year funding will be undertaken. This will have to be undertaken on a case-by-case basis, depending on the strength of evidence, but will be used to showcase value for money of multi year humanitarian response wherever possible.

The final category of impact of multi-year humanitarian funding is outcome - i.e. decreased case loads and avoided lives lost as a result of earlier and more effective programming. Quantification of changes in outcomes is very difficult: 1) it requires comparison of outcomes under annual and multi-year funding all else being equally, whereas in reality all disaster events are different and 2) it is very difficult to attribute a change in outcomes to one specific project or partner, whereas in reality changes in outcomes are typically a result of a variety of changes, including a move to multi-year funding.

As stated previously, the VfM framework will be used to document economy, efficiency and cost effectiveness where this data exists and outcomes have been documented. It is likely that a number of case studies will be developed that provide a quantitative analysis, and this will sit within a wider qualitative assessment that explores the benefits of MYHF and the theory of change.

The panel survey used in Question 1 will also provide evidence to feed into Q3 and specifically identifying and understanding the types of interventions that are facilitated by MYHF that lead to better outcomes, to the extent that this data can be collected in the survey. Cost effectiveness analysis could be a particularly useful tool to examine the comparative cost of annual and multiyear interventions that lead to similar outcomes.

### 3.5 Data management and analysis

The evaluation will generate a substantial amount of data including potentially quite valuable primary datasets from the 'resilience' research. This will require a data management strategy. Some elements of this are clear; others continue to evolve.

- Ensuring the evaluation stays focused on the questions: the longitudinal panel work will undoubtedly generate a huge number of issues that impact people's resilience. Many of these will be beyond the reach of humanitarian assistance, however smart or well targeted. Filtering out the 'pertinent' issues will be the responsibility of the team leader and the 'research leader' within the team. It will require careful analysis after each round of interviews to ensure that such issues are identified, and fed back into subsequent rounds of enquiry.
- Preserving the dataset for future researchers: if done well the panel dataset will be important for other researchers. This is the principal rationale for involving the University of Sussex. The team intend to ensure the dataset is deemed 'academically rigorous', that it is compiled in an accessible way and that (in agreement with DFID) it is made accessible widely through appropriate mechanisms.

Capturing and analysis of qualitative data: one of the most challenging issues for the team with regard to the panel research will be how to record, analyse and present the findings from interviews. In the pilot exercise this was done through a paper based system. Two interviewers worked together, with one taking the lead and the other recording (written). Notes were then further written up at the end of the day after a joint debrief session. These were then shared with the team leader who compiled and did some analysis.

While this system worked, the team is keen to try additional ways to capture the fine detail of interviews. One obvious addition will be to record people (using simple digital voice recorders), and then have interviews transcribed verbatim. There are several software packages (such as NVivo) that can encode plain text. This is much costlier and more labour intensive however, and there are translation challenges. Several variants of this type of approach will be trialled during the first round of panel interviews.

- Data confidentiality: some of the value for money data in particular will be commercially sensitive, or at least ‘internal’ to organisations. The evaluation will be in place confidentiality protocols to ensure that data is safe.
- Data storage: a system of archiving and storage is currently being worked on to make sure it is easy to retrieve and will not degrade or be lost over time. This is likely to be a mix of cloud based storage (for ease of sharing) with several layers of backup on hard drives and disk.

## 4. Evaluation team and management

### 4.1 Management of the evaluation

The evaluation will be managed by DFID, through the East Africa Research Hub, of the Humanitarian Innovation and Evidence Programme (HIEP). A management group, comprising relevant and involved internal stakeholders has already been active in the inception period. This will continue throughout the implementation phase. The management group comprises:

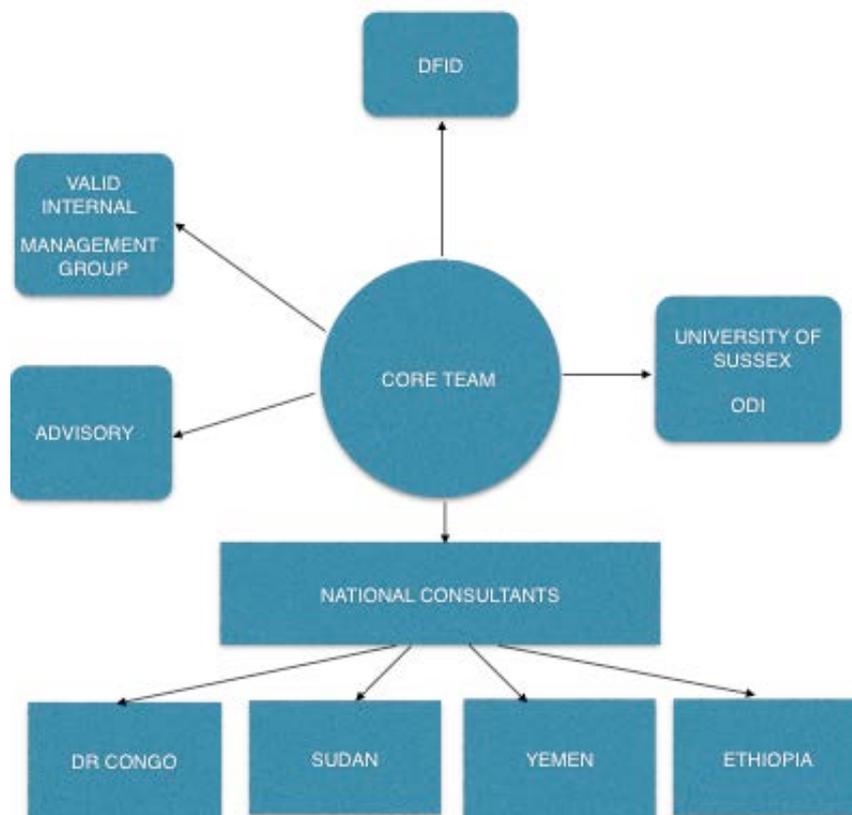
- Head of the Humanitarian Innovation and Evidence Programme (RED);
- CHASE senior resilience advisor
- ARD resilience advisor
- Programme Manager in the East Africa Research Hub (RED).

There is also a wider advisory group consisting:

- Humanitarian advisors from participating country offices;
- The head of humanitarian policy;
- CHASE senior evaluation advisors;
- Country office resilience advisors where applicable; and
- Representation from the Africa Regional Division.

### 4.2 Evaluation team and roles

The evaluation team combines a range of specialist skills and in-depth local knowledge. The Core Team consists of four international consultants, a senior national consultant from each of the subject countries, a VALID project management team and a number of independent advisors for the core team to refer to (Figure 3)



**Figure 2: The evaluation team structure.**

Team roles:

- Lewis Sida: team leader. Substantial experience of leading complex evaluations, combined with research and policy development, management and programme implementation in the humanitarian sector. Responsible for:
  - Managing overall delivery of the evaluation and products;
  - Managing teams and personnel, including being responsible for overall duty of care;
  - Liaising with DFID to provide regular updates on progress and discuss any issues encountered;
  - Regular country visits and field work;
  - Responsible for oversight of all components and contributions, including ensuring QA carried out;
  - Working closely with the project manager to keep evaluation on schedule and on budget.
- Courtenay Cabot-Venton: economist. Leading author on economics of resilience in humanitarian responses. Good research and evaluation experience. Responsible for oversight of VFM work, delivery of some primary field work and contributor to written reports.
- Bill Gray: humanitarian leader. Track record leading humanitarian responses and teams in the non-government and multi-lateral sectors for almost 30 years and

leading or contributing to humanitarian policy, evaluation and research efforts, Substantial advisory experience at national line ministry level. Responsible for field work across the four countries, ensuring that it is consistent and timely. Substantial contributor to written reports.

- Simon Levine: research leader. Leading expert on resilience in complex emergencies, and excellent track record in food security and east Africa related research. Currently part of the humanitarian policy group at ODI. Responsible for methodology and oversight of the resilience work, delivery of some primary field work and contributor to written reports.
- Basia Benda: project manager. Responsible for all reporting, administration and process management, including regular written reports, contracting, travel and payments. Supported by VALID team.
- Dr Eleni Asmare. Senior national consultant for Ethiopia. Phd nutritionist with substantial experience delivering complex research and evaluation projects (some for VALID). Hands on nutrition experience in humanitarian emergencies prior to consultancy. Responsible for delivery of all primary data work in Ethiopia, as well as practical and logistical arrangements regarding the evaluation more widely. Liaises with DFID office and partners.
- Dr Khalid Abdelsalam. Senior national consultant for Sudan. Medical doctor with over 20 years experience in humanitarian response. Has worked internationally for the UN and MSF, and extensively within Sudan. Responsible for delivery of all primary data work in Sudan, as well as practical and logistical arrangements regarding the evaluation more widely. Liaises with DFID office and partners.
- John Ntalemwa. Senior national consultant for DR Congo. New to the VALID team, John comes highly recommended after 20 years of working on humanitarian and organisational development issues for international and national organisations in Congo and elsewhere in Africa. Responsible for delivery of all primary data work in Congo, as well as practical and logistical arrangements regarding the evaluation more widely. Liaises with DFID office and partners.

At the bid stage it had been envisaged that the four international consultants would take a lead on one country each. They would have a 'second' or backup. Through the inception phase it was decided that this was unnecessarily restrictive and did not necessarily play to the strengths of the team. The senior national consultants also proved to be eminently capable of managing the 'day to day' somewhat reducing the need for a managerial (as a pose to technical) lead for each country. Roughly then, the new division of labour is that Lewis Sida and Bill Gray have more of a 'management' focus, with Simon Levine and Courtenay Cabot Venton having more of a 'technical' focus. Lewis Sida will cover DR Congo and any new country; Bill Gray will cover Ethiopia and Sudan, although these leads

#### Primary data gathering

The original VALID bid document had envisaged that primary data gathering would be done by national research institutes. This has been slightly amended with the team now proising to take a more 'mixed' approach. So far this means:

- In Sudan, the team will work with the national Academy of Health Sciences. Negotiations are still ongoing as to the exact nature of the relationship, with the aspiration being that named individuals will become part of the Valid team over

time, rather than Valid contracting AHS for a ‘product’. These individuals will be trained by Valid and supervised by Valid senior national consultant and potentially other Valid appointed experts.

- In DRC the model will be extremely similar, using the Université Libre du Pays du Grand Lacs (ULPGL). Researchers will be appointed to the Valid team and they will be supervised by Valid. Feedback, analysis and training will be run by Valid with input from ULPGL senior staff and the researchers. This model will be replicated in the next site chosen in DRC, potentially south Kivu.
- In Ethiopia this model will be applied more fluidly, with the Valid team comprising mostly research consultants, but still using researchers from local Universities in the study sites.

### Advisory group

The evaluation will have a number of leading thinkers to advise throughout the lifetime of the project, and to help with analysis of emerging findings.

*Paul Venton* has over a decade of experience working on resilience issues, starting with his PhD in sustainable community based disaster risk management. He has worked across the humanitarian sector (from World Bank to UNISDR to the Red Cross) and across the country contexts included in this study. He will provide a link to the ongoing discourse on resilience globally and new and emerging thinking, as well as modelling resilience.

*John Seaman* was one of the founders of the Household Economy Approach to food security analysis and before that was an epidemiologist, nutritionist, medical doctor and more. He is acknowledged as one of the humanitarian sectors’ original thinkers and an expert on research and modelling what are now being thought of as fragile environments. He currently heads an NGO called evidence for development.

*Ernest Guevara and Pakulu Bahwere* are Valid staff members with strong research methodology backgrounds. They will provide ongoing input into the design of survey instruments and research methods, as well as the design of analytical tools. Hugo Slim will advise on the ethics of evaluation, and ethics more generally. He will contribute to the thinking of the evaluation on the balance between humanitarian action in its purest sense and the aspects of DFID partners work that are financed through humanitarian budgets but have longer term horizons.

*Alistair Hallam* is one of the founders and Directors of Valid International and is well known as an evaluation expert, particularly in the field of humanitarianism. He has recently published a paper for ALNAP on evaluation uptake and will provide advice on this key area throughout the evaluation, as well as generally being involved in quality assurance.

### Sussex University

In the original bid, it was envisaged that Sussex University, Global Studies Department would oversee and ensure the rigour of the evidence base. This is still very much the intention, but has not been developed further during the inception phase. In part this was availability of various key people, in part it was a result of not

having a clear picture of the type of data that would need analysing until close to the date of submission of the inception report.

It is still planned that Sussex will work on data analysis, as well as helping to develop tools to interpret data. Research data collected in the four country contexts will be collected and analysed in databases at Sussex (as well as within the team) as further guarantee of the robustness of findings.

### 4.3 Evaluation uptake strategy

The evaluation uptake strategy will target a broad spectrum of stakeholders, seeking to simultaneously influence policy and practice. This will require strategies for influencing ongoing work, as well as feeding into broader policy debates.

Essentially desired change may occur on two levels; during the lifetime of the evaluation with implementing partners (practice), and over a longer time period affecting policy.

- Practice: primarily uptake strategies will be about feedback loops allowing for course correction. Helping partners understand where their work is having impact, or at least areas of programming that appear to have most potential, and conversely where areas of work do not appear promising. This will require close dialogue with partners, and where possible joint research efforts.
- Policy: there are several areas where the evaluation has potential to influence policy. Most obviously the growing practice of multi-year humanitarian financing; robust evidence showing gains from this approach has the potential to add to ‘critical mass’, making such approaches standard. There is also potential to add to the growing literature and evidence on resilience, and in particular in fragile contexts, helping agencies, donors and affected governments understand what works. There is also potential to further develop research methods and establish best practice in action-research on humanitarian programming in complex situations.

There are a number of categories of stakeholders:

**Table 13: stakeholders for evaluation uptake strategy.**

Stakeholder	Target	Method
DFID	As the principal client for the evaluation DFID is also the main stakeholder identified as part of the uptake strategy. The evaluation will seek to influence DFID policy and practice at both the headquarters and country levels.	Main opportunities for influence at the country level will include regular, ongoing briefings; participation by key staff in field work and written input (via formative and

	<p>At the country level targets will include business cases under development through the lifetime of the evaluation, and therefore future, follow on programming.</p> <p>At the HQ level there are a number of interested divisions or departments, including CHASE, RED and the regional teams. There will be a particular effort to link to other DFID learning initiatives such as ICF/ BRACED.</p>	<p>summative reports) into business cases.</p> <p>Main opportunities for influence at HQ include periodic briefings and workshops, written (formal) output. Targeted products (either written or briefing) at specific departments/ events. Findings and methods will be shared with other learning initiatives (as has already been the case) regularly.</p>
Implementing agencies	<p>At a strategic level targets include the HQs of large implementing agencies as well as senior in-country leadership. For HQs of large implementing agencies, changing systems to better use multiyear funding may be an important advocacy target.</p> <p>At an operational level, the main targets are implementing partners of DFID.</p>	<p>Strategic level influencing will be a mixture of ad hoc, ongoing opportunities such as briefings for UNCT in Ethiopia and DRC, as well as using formal products at HQ level.</p> <p>At an operational level influencing will be primarily based on joint working – involving partners as much as possible in the research and joint analysis of results.</p>
Other donors	<p>Important target for broadening the multiyear funding base; key donors include the US and EU who constitute a major share of humanitarian funding.</p>	<p>To be determined with DFID as results become clearer.</p> <p>Possibly using international conferences such as the World Humanitarian Summit.</p>
Governments	<p>May be a target for making policy changes on the basis of emerging evidence.</p>	<p>Will depend on nature of findings. Strategy for now is to inform and include where possible.</p>
Academia	<p>Potential to add significant evidence base to ongoing development/ humanitarian research around resilience.</p>	<p>Formal products such as formative and summative evaluations. Potential for peer review articles. Making the longitudinal dataset available for other researchers.</p>

#### 4.4 Quality assurance

Valid is well known for the high quality of its evaluations. It is committed to producing work of integrity and quality, primarily for the sake of the client and programme beneficiaries, as well as to safeguard its own reputation. Quality assurance is an integral part of any assignment undertaken by Valid employees and associates. The quality assurance system has a number components:

- *In-house review* of all evaluation products at key stages in their development: this will be led by Alistair Hallam (VALID Director), drawing on contributions from in-house method and subject experts.
- *External peer review* by a small team of advisors, comprising subject-specific expertise in food security and resilience, methodology development and research ethics will be referred to at key stages of the evaluation.
- Regular briefing and contact with DFID management to ensure that the evaluation is delivering on time and on track.
- Participation in subject/ expert meetings to keep up to date with current practice and thinking.
- Quality assurance by the University of Sussex, which will verify the process for data gathering and analysis. This will be done through periodic sampling of data ensure the sustainability of conclusions reached, and oversight of the process and the method.

#### 4.5 Risks

**Table 14: Risks**

Risk	Reason	Likely outcome	Mitigation
1. Data inconclusive	Timeframe is too short, the number of variables too complex, or attribution to DFID is not possible	A collection of conflicting results, with some indicators improving noticeably, others deteriorating and some staying static	The qualitative method chosen should largely mitigate this risk. The initial evidence will be qualitative rather than quantitative. This, together with quantitative data gathered on the basis of the qualitative enquiry should lessen the risk of confusing and conflict results.
2. Programme implementation impossible over the full lifetime of the research	Insecurity or other disruptive factors	Research and data collection compromised.	Local and national researchers can partially mitigate the effect of insecurity- related disruption (without a transfer of risk or duty of care). Localisation allows a dynamic

			assessment of risk easier travel in difficult areas.
3. Reliance on national capacity and skills	Security and other risks, and an in-depth knowledge of local conditions	Research and data quality possibly compromised	Supervision by senior national consultants and on-going close- and distant support from the core international consultant team
4. Team membership changes over the evaluation period	Illness, changed personal circumstances etc	Reduced quality of work, or work not completed	Core international lead and support consultants to ensure continuity
5. Access constraints (especially for international staff)	Visas, permissions, insecurity	Some elements of the work unable to be carried out	Senior national staff are the main mitigation strategy.

#### 4.6 Constraints and limitations

Methodological constraints have been described in some detail in the approach section above. Aside from these, the main constraints envisaged at this stage of the evaluation are around access, either from logistical, bureaucratic or security hurdles. A secondary concern is the willingness of partner agencies to share data, especially detailed information on cost. Where this is the case the evaluation team will work with DFID to overcome such obstacles.

#### 4.7 Ethical guidelines

VALID's ethical guidelines are inspired by, and draw upon, the UN Evaluation Group's ethical guidelines for evaluations (which are themselves based on commonly-held and internationally-recognised professional ideals). The evaluation will comply with the ethical guidelines, attached to this report as annex 5.

VALID will take a strong ethical line on any research undertaken, ensuring that all participants in primary research take part on the basis of informed consent. This will be particularly the case with panel participants. Similarly, all primary research will be subjected to a risk assessment before proceeding, to ensure that neither researchers nor respondents are put in harms way.

As noted above, all will be verified by an independent peer reviewer.

## 5. Work-plan and outputs

### 5.1 Outputs and milestones

During the implementation phase between 1<sup>st</sup> of October 2014 and 31<sup>st</sup> December 2017 (implementation period extended by 3 months with no cost extension) a number of outputs will be delivered by the evaluation team. For the entire evaluation period there will be a series of reports including:

- 3 or 4 formative country reports (2015). Currently these are planned for Ethiopia (June 2015); DRC (October 2015) and Sudan (December 2015).
- 3 or 4 summative country reports, planned for Ethiopia (September 2016); DRC (January 2017); Sudan (May 2017).
- 2 annual reports (2015,2016).
- Final synthesis report (2017).
- Possible additional products (peer reviewed journal articles) 2017/18.

Additionally each year there will be four quarterly narrative and financial reports submitted to DFID (for each year a narrative and financial report will be submitted by 20<sup>th</sup> January, 20<sup>th</sup> April, 20<sup>th</sup> July and 20<sup>th</sup> October). The reports will summarise activities undertaken in the reporting period and financial expenditure to date.

### 5.2 Work-plan

A detailed work-plan is attached in Annex 4 which includes a suggested schedule for the field work, meetings/workshops and outputs delivery. Timing of the implementation phase is scheduled between 1<sup>st</sup> of October 2014 and 31<sup>st</sup> of December 2017 (including a no cost extension of 3 months).

### 5.3 Synthesis

The final synthesis report will naturally draw on the country evaluation reports, in particular the four summative reports. It will also draw on the Somalia MYHF evaluation work (this is a separate MYHF study, see section 3.2.2). The synthesis report will also draw on other work generated over the course of the evaluation, including any discreet reports and work done outside the frame of this evaluation by team members of DFID country offices and partners that is directly relevant.

The synthesis process will consist of a number of internal team workshops to examine the evidence, as well as at least one involving DFID and key partners. An annual VALID meeting will also aim to take stock of emerging findings across the case studies, allowing for these to be recorded and for methods to be adapted.

