ABOUT THIS WORKING PAPER

The Child Labour: Action-Research-Innovation in South and South-Eastern Asia (CLARISSA) programme focuses on participatory generation of evidence and innovation. It is operating in conditions of complexity that pose significant challenges for programme evaluation. In this Working Paper we share our response to this challenge as action researchers reflecting on our experience with co-design. As we move from clarifying our espoused programme theory to building intentional practice, we are learning about how to build co-ownership and capacity within the context of emergent design. As we begin our journey of reflexivity with nested theories of change, combining contribution analysis with participatory adaptive management, we are learning about the importance of sequencing nested theory of change development to enable effective revisiting and refining of assumptions. Finally, we share our experiences of striving for the right balance between process and content in the use of an evaluative rubric to both evaluate and strengthen our partnership.

Authors
Marina Apgar, Research Fellow, IDS and Director of Evaluation and Learning, CLARISSA; Mieke Snijder, Postdoctoral Research Fellow, IDS and Evaluation and Learning researcher, CLARISSA; Shanta Kakri, National Coordinator – Monitoring, Evaluation, Accountability and Learning, Tdh-Nepal and Nepal MEL lead, CLARISSA; Shona Macleod, Research and Evaluation Officer, Consortium for Street Children, UK and Consortium MEL team member, CLARISSA; Sukanta Paul, National Coordinator – Monitoring, Evaluation, Accountability and Learning, Tdh-Bangladesh and Bangladesh MEL lead, CLARISSA; Anna Sambo, Deputy Manager Asia zone, Tdh Lausanne and Consortium MEL team member, CLARISSA; and Giel Ton, Research Fellow, IDS and impact evaluation expert, CLARISSA.

Suggested citation

Copyright
© Institute of Development Studies 2020

This report has been funded with UK aid from the UK government. The opinions expressed are those of the authors and do not necessarily reflect the views or policies of IDS or the UK government.

This is an Open Access paper distributed under the terms of the Creative Commons Attribution 4.0 International licence (CC BY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited and any modifications or adaptations are indicated.
http://creativecommons.org/licenses/by/4.0/legalcode

The Child Labour: Action-Research-Innovation in South and South-Eastern Asia (CLARISSA) is a consortium of organisations committed to building a participatory evidence base and generating innovative solutions to the worst forms of child labour in Bangladesh, Myanmar, and Nepal.
ACKNOWLEDGEMENTS

This Working Paper was made possible by funding from the UK Department for International Development (DFID). The authors are especially grateful to all the participants of the AAR workshops and the insights they shared to support programme learning, as well as to those who worked hard to organise these effective workshops. We thank Louise Clark, Pauline Oosterhoff, and Danny Burns for their comments that have improved the document.
ACRONYMS

AAR After Action Review

AM Adaptive management

CLARISSA The Child Labour: Action-Research-Innovation in South and South-Eastern Asia [programme]

IDS Institute of Development Studies

MEL monitoring, evaluation, and learning

NGO non-governmental organisations

PAM participatory adaptive management

PAR participatory action research

ToC theory of change

WFCL worst forms of child labour
Section 1:

INTRODUCTION
1 INTRODUCTION

Children end up in child labour as a result of many, often unknown or hidden, interactions between multiple actors and multiple factors within households, communities, and labour systems, leading to unpredictable outcomes for children and other sector stakeholders (Johnson 2017) and sometimes resulting in the worst forms of child labour (WFCL). It is a complex problem, and interventions aimed at tackling it are also, inevitably, complex and challenging. The way they influence change is non-linear, causality is uncertain, and unintended consequences may result. Programmes such as the Child Labour: Action-Research-Innovation in South and South-Eastern Asia (CLARISSA) that are engaging with such intractable challenges and aim to reach the most left behind (children in WFCL) are operating in conditions of complexity (Burns and Worsley 2015; Ramalingam 2013). This complexity poses significant challenges to the way programmes are designed, planned, implemented, and evaluated, and requires a move away from linear and predetermined models. In this Working Paper, we share our experience and early learning about how to design and implement monitoring, evaluation and learning that intentionally embraces the challenge of complexity.

1 The term ‘research for development’ is common in the field of agricultural development, but also is increasingly issued across a range of sectors to refer to programmes that use research as the main mechanism for achieving development outcomes (e.g. Horton and Mackay 2003; DFID 2016).

1.1 CLARISSA: a participatory evidence and innovation generation programme

CLARISSA is a participatory evidence and innovation generation programme, built in response to the rationale that there is little understanding of the complex underlying drivers of WFCL and evidence of what interventions work to reduce them (Oosterhoff et al. 2018; Idris, Oosterhoff and Pocock 2019). Further, current programming often fails to include the experiences of children themselves in understanding drivers and the development of appropriate solutions (Miljeteig 2000; Imoh and Okyere 2020; Sändig, Von Bernstorff and Hasenclever 2018). The CLARISSA programme design is radically inclusive and works through three nested levels of activities (see Figure 1).

Delivered through a consortium of research and development partners, the overarching programme approach is to use participatory and child-centred action research. In this sense, CLARISSA is a research-for-development programme. Action research with children and stakeholders (the programme response) is the main mechanism for generating evidence and innovative solutions. The groups take action to intervene...
in the specific problem areas related to drivers of WFCL (the actions). Through participatory evaluation of the actions taken by the groups, those seen to be effective at a small scale may feed into programmatic decision-making and inform investments in larger-scale interventions (the interventions) that will be piloted and evaluated to generate further evidence on effectiveness. The interventions sit within four thematic workstreams that respond to key drivers of WFCL:

- Testing targeted interventions in social protection;
- Understanding and supporting shifts in underlying social norms that contribute to children ending up in WFCL;
- Revealing and reducing harm in supply chains; and
- Building children’s voice and agency through child-generated interventions and advocacy.

The focus of the programme, and consequently evaluation of it, is not solely on ‘what’ achieves the reduction of WFCL – as would be the case in a conventional intervention programme and its impact evaluation – but rather, is also concerned with ‘how’ the innovative responses to reducing WFCL are generated and can be taken to scale.

1.2 The need for theory-driven and complexity-aware evaluation approaches

When programmes start from a weak evidence base of what seems to work, evaluation becomes an important vehicle for building new theory about how interventions are implemented (implementation research) and why they work differently in different contexts (evaluation research). This approach to evaluation is commonly referred to as theory-based (or theory-driven) evaluation (e.g. Weiss 1997) and is increasingly used in the international development sector in response to the challenges of complex evaluations of large multifaceted programmes (Vogel 2012). It is an example of how the sector is broadening appropriate impact evaluation designs (e.g. Stern et al. 2012). The focus is on verifying and refining understanding about the links between short-term and longer-term outcomes, and does not treat the processes which lead to change as a black box as, for example, done in (quasi) experimental impact evaluation that only measures impact on ultimate outcomes irrespective of the way that these are being generated. This theory-driven approach is compatible with the growing use of contribution analysis\(^2\) which includes the creative and appropriate mixing of methods to address specific links in the theory of change (ToC) for which insufficient evidence is available. Other approaches, such as realist evaluation\(^3\) (Pawson and Tilley 1997) are also increasingly popular in the sector, as they emphasise theorising and learning about social change processes by studying the causal mechanisms that are triggered under certain conditions for specific groups of people. They see this theoretical refinement as an iterative process to build an evidence base for future interventions. Across all of these approaches, the programme’s ToC becomes a fundamental building block for designing appropriate impact evaluation, where every evaluation is uniquely designed to fit the key information needs of the specific programme theory.

The overarching response of the CLARISSA programme to the challenge of children ending up in WFCL is participatory action research (PAR), which is the main mechanism for generating innovative interventions (e.g. Burns and Worsley 2015; Apgar and Douthwaite 2013). This implies that it essentially does not specify what will be done in advance but leaves this to the stakeholders involved in the research process to determine. This emergent participatory design and implementation makes it challenging to determine a baseline situation, defined by measurable indicators that can be tracked in time, as it is not possible to predetermine which indicators are likely to be changed by the PAR actions. The open process of change has significant implications for the way that the programme can be evaluated. Emergent impact pathways require that we learn as we go. CLARISSA’s ToC may well change dramatically in the course of the programme. Therefore, alongside the theory-driven ToC approach in CLARISSA, there is a need to learn as change happens and face this intrinsic uncertainty and complexity (Douthwaite et al. 2017; Patton 2010).

---

\(^2\) Contribution analysis is an approach to identify the attribution a programme makes to observed impacts (see Figure 2 for further details) or read Mayne (2008) or Ton et al. (2019) for detailed steps in this approach.

\(^3\) A realist evaluation explores how and why an intervention is generating changes, in what circumstances, and for who. A realist evaluator explores the context within which the intervention is implemented, the mechanisms that are triggered by the interventions, and the outcomes that the intervention generates in this context and through these mechanisms. For a practical reflection on the use of realist evaluation in development programme evaluations see Punton, Vogel and Lloyd (2020).
To be truly complexity aware, proponents of the approach argue, evaluation needs to be driven by a learning agenda which includes real-time monitoring,\(^4\) goal-independent evaluations,\(^5\) and the regular revisiting of initial assumptions in the ToC. To maximise adaptive management, monitoring and evaluation should be embedded within the operational management structure of the programme to ensure that they can contribute to learning and adapt to meet long-term goals (Arkesteijn, van Mierlo and Leeuwis 2015; Douthwaite et al. 2017).

As with many emerging fields, there is a lot of theorising about how to be complexity aware with programme evaluation and how this might support adaptive management; however, there is also a dearth of practical examples that evaluators or programme developers can use to guide their work (Douthwaite et al. 2017). Even fewer are the documented examples of complexity-aware evaluations of programmes that use participatory approaches, and specifically lacking is detail on how they are built upon the experiences of marginalised people as well as of programme implementation teams, and how the tensions around multiple forms of evidence are navigated. As there are few programmes the size of CLARISSA that fully embrace a participatory ethos from the outset, we are in a unique position to contribute to methodological innovation in this area of complexity-aware theory-based evaluation.

### 1.3 A collective paper sharing methodological innovations

The co-authors of this paper are all members of the CLARISSA Monitoring, Evaluation and Learning (MEL) team located in-country and at programme level across all consortium partners, and are directly involved in designing and operationalising the programme’s espoused approach. The programme has now completed the inception and in-country set-up phases and is moving into full implementation of action research when all MEL tools will become fully operational. In this paper, we share the beginning of our journey through first and second person enquiry (e.g. Reason and Torbert 2001; Coghlan and Brannick 2005) into how to design, operationalise, and learn from an intentionally participatory and complexity-aware programme evaluation. We do so through first describing the participatory design (our espoused theory) through three interconnected areas of methodological innovation: (i) a participatory adaptive management approach; (ii) applying reflexive use of ToC for rigorous impact evaluation; and (iii) a participatory rubric-based partnership evaluation. Next, we share early learning from our experience of building the design through the first year of programme implementation and reflect on challenges and opportunities we have encountered along the way. We end with a discussion on preliminary implications for both programming and evaluation specialists embarking on similar evaluation designs.

---

\(^4\) Real-time monitoring in this context refers to observing what is happening in the programme as it is taking place, rather than monitoring at certain specified time points.

\(^5\) A goal-independent evaluation sets out to understand and evaluate what actually changes independently of what the programme was aiming to achieve; in other words, it does not assess outcomes against predetermined goals and so is well suited to capturing unintended change and surprise.
Section 2:

INNOVATIONS IN PROGRAMME MEL
2 INNOVATIONS IN PROGRAMME MEL

2.1 Participatory adaptive management

Adaptive management (AM) in the international development sector is embraced as an approach that helps programmes become more learning oriented and so more effective in addressing complex development challenges. A multitude of approaches are being developed and tested in practice and are leading to new communities of practice (Prieto Martin, Apgar and Hernandez 2020; Bain, Booth and Wild 2016; Wild, Booth and Valters 2017; USAID 2018; Honig and Gulrajani 2018; Vowles 2013). The current AM landscape in the sector, having myriad approaches and few practical examples, risks AM becoming simply another development fad of little consequence (e.g. Shutt 2016).

The participatory adaptive management (PAM) design in CLARISSA builds on and aims to deepen current practice and debate in the field. Much AM practice is problem driven, aligned with the general focus of development interventions as vehicles for solving well-defined and often technically oriented problems. An influential example is the Problem Driven Iterative Adaptation approach (Andrews, Pritchett and Woolcock 2012), which supports use of active learning through real-world experimentation with the aim of getting better at solving a predefined ‘problem’. In conditions of complexity, and when drivers of systemic challenges are not well known at the outset, the problem definition, we argue, must include experience of the change makers themselves – the stakeholders of the system in which a programme intervenes – if the proposed solution is to be fit for purpose. Based on evidence that ‘empowered people’ are a core characteristic of successful AM in practice (Prieto Martin et al. 2017), there is a move towards centring decision-making around local change makers’ experiences. Development actors advocating for local ownership of AM (Yanguas 2018) suggest that programmatic decisions should be based on local perspectives of how the problem is framed and learning about what is working and how through the real-life experience of change on the ground.

For CLARISSA, local ownership of AM means enabling a definition of the drivers of WFCL (the ‘problem’ space) by children, their parents and guardians, and other actors in the supply chains (such as small business owners) themselves, as well as listening to their experiences of change as it unfolds to inform programmatic decision-making throughout implementation. We argue that this starting point and the participatory monitoring and evaluation, made possible through the participatory and child-centred approach of CLARISSA, will provide an opportunity to develop a people-driven approach to adaptive management.

2.1.1 Building participatory adaptive management (PAM)

The CLARISSA MEL system responds to both upward and downward accountability needs (Apgar et al. 2019). It is widely acknowledged that there are tensions between these two objectives of MEL (e.g. Guijt 2010). A critical design feature is therefore the creation of spaces that navigate potential tensions between ensuring the programme is ‘on track’ based on its broad definition of results as committed to the donor, while also being certain it is responding to what is emerging on the ground to make sure our work remains based on the experiences of children, parents, and other stakeholders. Bearing this in mind, we have structured the PAM processes following the useful distinction made by Green and Guijt (2019) building on Punton and Burge (2018) around three interconnected levels of learning and decision-making for AM: (i) adaptive delivery, (ii) adaptive programming, and (iii) adaptive governance (see Table 1 for a definition of each). At each of these levels, monitoring data is fed in, with the intention of stimulating reflection and learning that in turn should support intentional response and adaptation. Our design of modes of participation, to move towards being people driven, is summarised in the second column of Table 1 and described below.

It is important that there is a supportive culture to facilitate building PAM within a programme (Ramalingam, Wild and Buffardi 2019). As the various organisations in the consortium are at different starting points regarding PAM, building this supportive culture is made explicit across the CLARISSA partnership. This includes that the partnership values adaptability, an entrepreneurial culture and risk taking, and openness and trust to learn from failures. How well the partnership is actually functioning in creating a supportive culture for PAM is described in Section 2.3.

Field-level monitoring of programme activities includes ongoing documentation of and by the PAR groups operating in specific sites within the three countries of operation (Bangladesh, Nepal, and Myanmar). The programme aims to run approximately 18 PAR groups.

---

6 Detail on the PAR methodology is in Burns et al. (2020, forthcoming), built on Burns (2018, 2019). The PAR groups will have 6–12 participants and be identified through a participatory systems analysis of children’s life stories.
## Table 1: Participatory adaptive management design

<table>
<thead>
<tr>
<th>Aspect of AM</th>
<th>Description and implications for practice</th>
<th>CLARISSA participatory design</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Adaptive delivery</strong></td>
<td>Activities that adapt programmes on the ‘front line’ and rely on field staff applying local evidence, emotional intelligence, and curiosity to stay nimble and flexible in the face of ever-changing conditions. This requires an iterative process of engagement and learning for rapid adaptation.</td>
<td>Learning from PAR groups (children and stakeholders) is captured through programme-supported documentation of the PAR process, including monitoring qualitative and quantitative indicators to assess how the groups are working (performance and facilitation) and what they are achieving (innovations and outcomes). Local implementation teams (facilitators and documenters) periodically review the learning to adapt implementation in consultation with the country-level team. Beyond this production of evidence and learning, a participant feedback mechanism will also include the opportunity for broader independent feedback to be captured and fed into programme sense making.</td>
</tr>
<tr>
<td><strong>Adaptive programming</strong></td>
<td>A learning process that is structured, often through a programme’s MEL system and is implemented by programme managers. The timescale for cycles of iterative learning are longer than for adaptive delivery, fitting to regular evaluation moments to reflect and decide on whether to continue/stop activities. It is a more in-depth analysis to reflect on programme implementation and changes in context and to make decisions for responding.</td>
<td>Facilitated AARs are implemented on a six-monthly and annual basis within each country and across all countries. Monitoring data and learning from programme activities are the main inputs for the AARs – this includes learning from PAR groups and synthesis of findings from the programme’s participant feedback mechanism. A core element of the AARs is the use of a partnership’s self-evaluation process (see Section 2.3). Learning reports are produced as outputs of the AARs to make the learning actionable. The sequencing of AARs is critical to ensure learning can be ‘fed upwards’ in the programme from country to consortium level. The programme plans are adapted in response to the actionable learning and in turn fed upwards to the donor through the annual reporting process.</td>
</tr>
<tr>
<td><strong>Adaptive governance</strong></td>
<td>The enabling environment for adaptive programming and delivery, which includes contractual arrangements with donors and should provide flexibility for adaptation (including adjusting budgets). This level has to navigate the tensions between accountability and delivery of predefined results and enabling learning that might shift priorities and focus.</td>
<td>Annual reporting feeds upwards to the donor through the Accountable Grant mechanism. A close relationship is maintained with DFID through the co-generation phase and continues through implementation. In response to the impact of Covid-19 on the children with whom the programme aims to engage, for example, major programme adaptations were agreed at the adaptive programming level and approved by DFID at this level.</td>
</tr>
</tbody>
</table>

Source: Authors’ own.
in each of these countries. Each group will articulate its own ToC around the actions it designs and will use this to evaluate their effectiveness. The PAR group processes will be documented by programme staff (documenters and facilitators) working closely with PAR group members to track qualitative and quantitative indicators to assess how the groups are working (performance and facilitation) and what they are achieving (innovations and outcomes, based on their ToCs). Periodically (on a biweekly or monthly basis), the documentation teams (programme staff) will review the monitoring data from the PAR groups to synthesise emerging findings that require field-level operations to adapt, for example, by changing the facilitation methods to improve group dynamics or by adding technical support if groups require it, etc. The resulting adaptive delivery is intended to help programme activities remain focused on the goals as defined by the PAR groups – their specific areas of enquiry – and the relationship between those and WFCL.

Adaptive programming is a slower and more structured learning process which will be operationalised through annual cycles of planning, implementation, and reflection. A facilitated learning process has been shown to provide an effective mechanism for changes in both the context and programme implementation to be surfaced, captured, and shared, and for critical areas of learning to be documented and fed into decision-making (e.g. The Asia Foundation 2015; Douthwaite et al. 2015).

Given the nested levels of programme activities in CLARISSA (see Figure 1), the participatory learning cycles are implemented at country level, where the implementation teams working with multiple PAR groups come together in each country of operation as well as at the whole of programme consortium level, where representatives of all partners and all countries come together for a strategic review and decision-making. The cycles are linked intimately with the annual programme cycle and support upward accountability through making the programme accountable not just for delivery of results, but also for learning. This also becomes a natural space for assessing the performance of the partnership and, as explained in Section 2.3, an evaluative rubric is the tool we use for that. The main mechanism that includes partners and enables a ‘programme view’ of actionable learning about changes in both context and programme implementation is the After Action Review (AAR) workshop (see Box 1). AAR is a simple tool that, if applied systematically across programmes and organisations and if facilitated to enable open and honest reflection, has been shown to drive organisational learning in a diversity of contexts including research-for-development programming (Collison and Parcell 2007; Douthwaite et al. 2015; Whiffen 2001).

Critical to the success of any AAR process is to ensure clarity and quality in the monitoring data that are fed into the moment of collective learning. The PAM design of the programme, therefore, takes seriously the need to

---

**Box 1: After Action Review: a versatile reflection and evaluation tool**

The After Action Review (AAR) is a simple structured and facilitated evaluation process used by a group (an implementation team, or a broader stakeholder group) to capture the outcomes and lessons learned from past successes and failures with the goal of improving future performance. It provides an opportunity for the group to reflect on a project, activity, event, or task so that the learning can become actionable and next time they can do it better.

AARs enable teams to step back from day-to-day implementation and take time to look back on what has been achieved. It is a versatile tool that can be used for both small quick post-activity reflections (e.g. a half-hour process by the team after a day in the field) and for longer processes (e.g. to facilitate a three-day workshop including a project implementation team to reflect on a year of programme activities). Three sets of generic questions are the starting point for contextualised AAR design: (i) What was achieved? (ii) What worked well, what didn’t, and why? (iii) What was supposed to happen? What actually happened? Why were there differences?

Source: CLARISSA AAR guidance document.
ensure rigour in the types of data and evidence to form a solid foundation for adaptive decision-making across all levels – from delivery to programmatic to governance levels. A participant feedback system will help to close any gaps in programme monitoring by providing an additional independent means of input for people who are participating in or otherwise targeted by the programme activities that are not collected through the documentation of the PAR activities. Potential methods include interviews and game-based evaluation processes for children at specific moments, as well as random questionnaires to be implemented throughout field work to capture perceptions of non-participating community members (children and others). As discussed in the following sections, both the use of ToC and application of the partnership rubric also determine the types of data and sources of information that should be prioritised to enable adaptive decision-making in such a consortium-driven innovation programme.

2.2 Reflexive use of theory of change

In line with the theory-based evaluation approach used in CLARISSA, ToC is the centrepiece of the evaluation design and, by extension, the programme’s MEL system. In complex programmes and those that focus on generating innovation, using ToC requires high levels of flexibility and needs to be process oriented. It accepts that we cannot determine beforehand exactly how and what change the programme is going to produce (Arkesteijn et al. 2015). Our approach to reflexive ToC follows those by Mayne (2015) and Vogel (2012). In this approach, ToC is not just a product (a diagram with a narrative), but rather, it becomes a facilitated and critical thinking process through which programme assumptions are made explicit, investigated, and evaluated. A ToC is always imperfect, being the result of the imperfect knowledge and aspirations of the stakeholders involved in drafting it (Ton et al. 2019). The ToC evolves over time as the programme takes advantage of opportunities and learns from difficulties that might emerge, rather than being forced to follow a preset programme logic. This is an important element of dealing with complexity in evaluation as well as in the evaluation of innovations (Rogers 2008; Patton 2010).

Terminology relating to ToC can vary widely between evaluations and programmes (e.g. ToC is often used when referring to (a set of) pathways of change); it is therefore important to make explicit how we are using the terminology in CLARISSA (Mayne 2019) – see Box 2. We use a nested model of ToC which combines a programme-level ToC, workstream-level ToC around specific designed interventions, and country-level ToC.

---

**Box 2: Theory of change terminology used in CLARISSA**

**Theory of change (ToC):** The best evidenced guess of how change happens by describing the causal pathways and the assumptions of what needs to happen for the causal links to take place. It is a facilitated process that brings together partners to build a shared vision and ownership of project objectives. It helps to deal with the complexity of the change processes.

**Interventions:** The set of activities implemented to contribute to a set outcome. As a research-for-development programme, CLARISSA has a nested set of ‘interventions’. At the programme level, action research is framed as a response to WFCL. Within the response are the PAR groups who generate actions as well as more traditional interventions (e.g. social protection), and promising PAR actions will be piloted as interventions. (See also Figure 1 and Section 1.)

**Assumptions:** Assumptions about how each of the causal links in the ToC should work. **Contextual assumptions** are assumptions about the external context that explain why the intervention is planned the way it is. **Causal link assumptions** are the events or conditions that almost always need to occur for the causal link to realise.

**Results:** The outputs, outcomes, and impacts in the ToC.

Source: Authors’ own.
Evaluating CLARISSA: Innovation Driven by a Participatory Learning Agenda

Additionally, each of the 18 PAR groups in each country will also develop their own ToC about their actions. The programme-level ToC provides a common framework across the programme that can accommodate local adaptation and change. Country-level ToC, on the other hand, is the adaptation of programme-level ToC to country context. These nested levels of ToC align with the nested levels (country and programme) at which the programme’s PAM approach is operationalised and the nested levels of activities (see Figure 1).

2.2.1 Co-developing and evidencing theory of change (a realist contribution analysis)

Figure 2 sets out the steps to be taken in developing and evidencing a ToC within a contribution analysis approach to impact evaluation (Mayne 2008; Ton et al. 2019). The initial programme-level ToC was co-developed with representatives from all consortium partners during the co-generation phase of CLARISSA. This participatory approach to using ToC means that the outcomes are defined collaboratively by researchers and those that are making the changes (change agents), contributing to co-ownership over the outcomes. For AM, multiple data types and sources are essential to inform programme development and adaptations (Ramalingam et al. 2019). This includes both formal (e.g. collected through quantitative surveys or qualitative interviews or literature searches) and informal (e.g. experiences from the field, expertise from practitioners) data sources. The use of informal data sources is important to make sure that we build on existing processes and potential innovations are not lost. We formalise the integration of informal and formal evidence by using a broad range of evidence (Apgar et al. 2019). In line with CLARISSA’s participatory approach, our starting point is the knowledge generated on the ground. This is supported by a broader typology of evidence that also includes formally published research evidence, the experiences of practitioners (practice-based evidence), evidence produced through participatory research methods (co-produced evidence), and physical or material evidence that prove a fact (physical evidence). The current programme ToC is developed based on multiple types of evidence.

7 The current programme level ToC can be found on the CLARISSA website (https://clarissa.global/about-us/).
The overarching programmatic evaluation question is **How, in what contexts, and for whom can PAR generate effective innovations to tackle the worst forms of child labour and how can they be scaled to reduce the worst forms of child labour?** We are taking a realist approach to our contribution analysis by going beyond identifying whether or not PAR contributes to the reduction of WFCL, and are exploring how, for whom, and in what circumstances (Pawson and Tilley 1997; Rolfe 2019).

To build a theory-driven impact evaluation design, the essential next step (Step 3, Figure 2) is to deepen the evidencing of our underlying assumptions in our ToC (e.g. through literature reviews). In line with our realist approach to contribution analysis, we are undertaking a rapid realist review (Saul et al. 2013) to understand the ways in which PAR produces innovative interventions (to tackle WFCL) and through which key mechanisms it does so. The outcomes of this rapid realist review will enable us to further build the contribution story and identify where more evidence is needed (Step 4, Figure 2). How this evidence will be generated and then combined will depend on what needs to be generated. We intend to use a mix of methods, including using monitoring data that is generated by participants themselves, as well as goal-independent and participatory outcome evaluations in which knowledge from those on the ground will be prioritised (e.g. through using Outcome Harvesting (aligned with the adapted Outcome Evidencing approach; see Paz-Ybarnegaray and Southwaite 2017)) and complemented with process tracing (Beach and Pederson 2019) to verify the contribution claims. This will be conducted once the PAR groups are up and running (in Year 2).

What we have described in this section is the intentional design to use reflexive ToC as the centrepiece of a MEL system that merges theory-based evaluation with PAM. One of the challenges we anticipate with this approach, and which will focus our quality assurance efforts, is how to practically bring together different types of evidence to help CLARISSA learn its way to achieving impact while simultaneously answering the evaluation question and so contributing to the evidence base on how PAR works. The potential contribution to the field of evaluation will be to put contribution analysis at the service of adaptive programming.

### 2.3 Strengthening the consortium partnership using an evaluative rubric

The consortium is key to implementing CLARISSA as a research-for-development programme (see Figure 1) and its key processes, including the ToCs and PAM. The CLARISSA consortium encompasses researchers as well as international and national non-governmental organisations (NGOs). Given the partnership’s key role in making the programme work, the functioning of the partnership is closely linked to the functioning of the programme as a whole. Indeed, it is theorised as a critical mechanism for success of the programme’s innovative MEL design. It is, consequently, important to continually assess how the partnership is developing and functioning and whether it is building a supportive culture for PAM.

An evaluative rubric is used to assist the partnership in reflexively assessing its own functioning and so finding opportunities to strengthen.

Rubrics originated in the field of education to help generate more objective criteria for marking students’ performance and assist teachers to help students acquire the skills to perform (Popham 1997). In the early 2000s, rubrics were introduced as a potentially useful tool for evaluations (Davidson 2005) but uptake beyond a few Australasian evaluators has been slow, and still much practice remains absent – at least formally – from the UK development circles. A rubric is, in its simplest form, a qualitative assessment tool with evaluative descriptions of what performance or quality looks like at various levels. It can help make explicit the judgements about the quality value and the importance of interventions being evaluated.

Generally, a rubric has three components: the key aspects of performance, the level of performance, and the importance of each aspect (Oakden 2018). The components can be combined in different ways and normally the rubric is presented in a table or matrix. Rubrics are preferably co-developed with the stakeholders who identify the different performance aspects as well as the levels of performance, thereby generating a shared understanding and ownership of the evaluative processes (King et al. 2013). This is particularly useful in the context of a learning-oriented evaluation design, as they can assist the partnership in making the evaluation an ongoing conversation for reflection and learning (Apgar et al. 2017; 2013).

---

8 See, for example, recent blogs by Thomas Aston (www.linkedin.com/pulse/rubrics-harness-complexity-thomas-aston) and Maloney and Atkinson (www.artd.com.au/news/riding-the-rubric-wave/).
Wijenberg et al. 2019). There are multiple reasons why we propose using evaluative rubrics for partnership evaluations as a particularly good fit for CLARISSA:

- When partners are involved in the development of the rubric it provides a unique opportunity for the whole partnership to develop a shared understanding of what a successful partnership looks like.
- For each of the aspects of partnership performance in the rubric, all partners are asked to contribute to qualitative descriptions of the levels of performance. This can generate more ownership of partnership as well as over how the partnership is measured, compared to using other standardised surveys that are most commonly used in partnership evaluations.
- The application of the rubric itself promotes discussion between the partners. This is an important element in partnership evaluation, as the evaluation of the partnership is also a tool to improve the workings of the partnership in line with the adaptive management intent of the programme.
- The evaluative descriptors in a rubric will be specific to each programme and context, making rubrics highly adaptive. This adaptability is especially ideal in partnership evaluations given that every partnership is unique.
- Compared to using quantitative indicators, use of an evaluative rubric can actually help improve the functioning of the partnership. The qualitative descriptors in a rubric provide guidance on how to make the partnership function well. When partners are working to merely tick the boxes of these descriptors, they are actually generating a well-functioning partnership – as opposed to when partners are trying to just do the bare minimum to achieve a quantitative indicator.

In line with CLARISSA’s participatory approach and best practice in complexity-aware evaluations (Douthwaite et al. 2017) and evaluative rubrics (King et al. 2013), the partnership rubric was developed with representatives from all consortium partners at the end of the inception phase. In an operational design workshop held in Brighton in April 2019, representatives from each of the partner organisations collectively revisited the partnership principles that had been outlined during the proposal phase. Based on their past experience of consortium working and Institute of Development Studies’ (IDS) experience of participatory programming, it was agreed by the group that the following principles should guide quality adaptive programming: (1) communications; (2) team identity; (3) openness, honesty, and mutual trust; (4) impact orientation; (5) inclusivity and equitability; (6) adaptability and flexibility; and (7) entrepreneurial culture. These seven principles became the partnership’s key performance areas (or elements) around which the evaluative rubric was built. Next, each partner (through their representatives) worked on their own to define what for them would be useful qualitative indicators of the partnership working well, emerging, or needing help, in each of the performance areas. The lead author then combined the input from all partners and presented back to the Strategy and Operations team what became the starting rubric for assessing the performance of the partnership (see Annexe 1).

As a key evaluation component for the consortium, this initial rubric defined by the partners was intended to be used as a self-evaluation tool on a regular basis. In practice, as described below, it has been integrated into the AAR workshops to provide systematic monitoring of the quality of the partnership as it evolves. We expect the rubric itself to evolve through the implementation phases of the programme and it will be periodically updated. To supplement what is necessarily a negotiated collective view from the main self-assessment approach, a partnership survey will be used annually to collect data on individual experiences of the partnership within the same performance areas and will be analysed to illustrate differences by gender and other relevant differentiators within and across the consortium. The two data sets will be brought together at the consortium level AAR to inform adaptive programming and adaptive governance.
Section 3:

EARLY LEARNING THROUGH PRACTISING DESIGN AND OPERATIONALISATION
3 EARLY LEARNING THROUGH PRACTISING DESIGN AND OPERATIONALISATION

In this section, we share our early collective learning in line with the action research approach to our own enquiry through and about the design and implementation of evaluation in CLARISSA. We reflect on the challenges and opportunities we have encountered as we move from clarifying our espoused programme theory to intentional practice. This early period has focused largely on refining the programme design in context and setting up the integrated components of the MEL system as the country teams are being built and evaluation design is deepened around specific activities with defined stakeholders. The learning system and initial adaptive management are now becoming operational. We reflect here on the learning that surfaced through the first round of AARs implemented in all three countries that took place in February–March 2020. Additionally, we also include learning surfaced through scoping missions (to help refine operational ToCs) and reporting. Due to Covid-19 there was a disruption in the consortium-level AAR that was planned as the main adaptive programming moment towards the end of the first year of implementation, which intended to enable a deeper reflection on how the design is taking shape in practice. A virtual engagement will be facilitated instead.

The first round of country-level AARs were three-day workshops (in Nepal and Bangladesh, and a shorter session in Myanmar). The first day included looking back at activities implemented in the programme so far, identifying both challenges and successes as experienced by participants that generated reflection on the programme design and MEL processes. The partnership’s self-assessment was implemented through the first application of the rubric. During the second day, the ToC was revisited, to help surface contextualised assumptions and feed into ongoing ToC evidencing. Participants of the AARs included representatives of all in-country partners and were co-designed and co-facilitated by members of the programme MEL team based in the UK and in-country to ensure consistency across countries while also enabling contextual adaptation. Here we share three areas of learning that illustrate the challenges of moving from theory to practice of PAM and complexity-aware evaluation.

3.1 Allowing time while staying focused on co-ownership as we move from theory to practice

Even though there was no specific session in the AARs that discussed AM, it surfaced repeatedly as an area of significant learning for the teams. When participants were looking back at what they had done under the CLARISSA programme to date and reflected on the challenges and opportunities that arose from it, operationalising the PAM approach surfaced as a challenge across all countries and partners.

Misunderstandings of the participatory nature of decision-making and information that should be used to fuel AM were obvious given that most participants’ experiences with learning to feed into AM has been a more ‘top down’ extraction of learning to feed higher-level decision-making. Further, as the PAR activities with children and stakeholders have not yet begun, nor is the participant feedback mechanism active yet, at this early stage it is not surprising that it is unclear how to practically ensure their learning feeds decision-making by the programme partners – downward accountability is not yet fully operational. One of the known challenges when moving from intentional design to action in a research-for-development programme is the time it takes to co-construct across theoretical and practical dimensions. No funded programme is entirely bottom-up; the starting point is necessarily driven by those responsible for (and funded to) design the programme – CLARISSA is no exception. As ground-level implementation processes start to take shape, and as we ensure quality in the participatory spaces we create, the PAM approach will enable the programme to hear and respond to voices from the ground. Yet at this early stage, as evidenced through the AARs, the lack of practical understanding around how to operationalise the PAM design by the full programme team is resulting in a lack of ownership of the PAM approach in-country.

A key area of confusion and tension around how to set up a programme that is intentionally emergent (and so has a MEL system that is co-constructed in context) is navigating uncertainties with confidence whilst simultaneously learning how to work differently. Partners understood the emergent design to be linked directly to the adaptive intentions of the programme. For some participants, it felt thus far as if ‘adaptive management means that anything goes’. The lack of structured adaptive decision-making in practice – adaptive programming – at country level was experienced...
as unsettling, and some expressed worry that constant change made it difficult to feel like progress was being made. Now that AARs are implemented systematically and decision-making structures are put in place at country level, the challenge that remains is to empower country teams to own the learning and actioning of change while at the same time remaining open to being surprised by learning that is fed upwards from the ground as PAR groups become operational.

Adaptive leadership models and past experience with large participatory programmes (e.g. Apgar et al. 2015) show that embracing ambiguity and practising reflexivity are core competencies that enable quality in AM. Navigating uncertainty and embracing surprise become easier when teams are reflexive and learning oriented in their practice. Yet nurturing such competencies does not happen in programmatic isolation, but rather, must be understood as linked to the institutional enabling environments people work within. The AARs evidenced noticeable differences between individuals and partners from different organisations around comfort with responding to context, enabling honest reflection about and learning from what does and does not work. The differences relate, in part, to the institutional cultures they are embedded within and whether adaptive governance is practised within them. As the partnership rubric includes an element of flexibility, there is scope to reflect on these differences as we move through implementation.

Part of enabling core adaptive competencies as an important prerequisite of PAM to be successful is having organisational structures that are nimble enough to allow for the changes to be implemented in response to the learning that is surfaced. One partner talked of the ‘gymnastics’ that are required for their slow systems to deal with the flexibility that adaptive delivery requires on the ground. We are not the first programme to experience such challenges in implementing AM; indeed, ensuring operational, financial, and contractual processes are flexible enough is a challenge for many (e.g. Prieto Martín et al. 2020). There is also, however, rich experience to build upon within the consortium. Some partners have experience using evidence-based decision-making to guide their practice and implementation of programmes on the ground, but do so mainly in an adaptive delivery mode rather than a structured and more strategic adaptive programming mode. For example, one partner described shifting their interventions around education of indigenous peoples through ongoing monitoring of the context and being directly responsive to how communities were engaging. Another partner used the metaphor of organic growth to reflect on how the programme will take shape as the PAM processes come to life and we naturally develop to remain focused on the impact on WFCL. We have the opportunity to build on teams’ strengths, and to focus ongoing work on operationalising PAM within CLARISSA through a strengths-based approach to team building and capacity development while also keeping an eye out for gaps that need to be filled.

In summary, we have learned that as time enables experiential learning, it is critical that we stay focused on building co-ownership over the processes in-country. An important starting point is to provide further training to help on-the-ground partners to understand the design as we deepen operationalisation. More ought be done in this regard and, as one participant from Nepal suggested in their feedback, we should ‘share more examples so the academic theory/approach can be understood by practitioners’. Due to the different levels of familiarity with the PAM approach across partners, further guidance and structuring of the processes that will enable evidenced adaptation are a priority; in particular, detailing how the PAR documentation system and the participant feedback mechanism, once fully operational, can feed adaptive delivery that is nested within the adaptive programming processes. We expect that fully operationalising the participatory mechanisms will help to reduce the sense that AM means ‘anything goes’ and will thus nurture more ownership.

3.2 Learning about revisiting assumptions

The reflexive use of ToC in CLARISSA means that assumptions should be revisited as we learn more about WFCL and how to address it in each specific sector and country. Often developing a ToC for a country programme is done at the outset, almost as the first organising activity that gives the programme its shape. For a theory-driven evaluation design, it is also paramount to clarify upfront the causal pathways from an intervention to intended outcomes. Yet, when dealing with an emergent participatory design – as we are in CLARISSA with its use of PAR – and acknowledging that the starting evidence base is weak and inconclusive, it means that the right moment to build the detailed causal ToC at country level is not obvious. Indeed, as noted in other large research-for-development programmes (Douthwaite et al. 2017; Apgar and Douthwaite, forthcoming), ToCs should emerge to the right level of detail as and when it is possible to define interventions with specificity. To attempt to detail it in full too quickly could mean that we focus energy on the
wrong aspects of the ToC and so may chase irrelevant data. However, in spite of the rhetoric there is very little evidence of how to practically ‘revisit’ assumptions in a ToC. What does this look like in practice and what have we learned about this so far?

Assumptions about key drivers of WFCL were challenged through research scoping exercises (see Oosterhoff and Hacker 2020; Yunus 2020) as well as during the ToC revisions during the AARs. A CLARISSA Emerging Evidence Review on moneylending as a key driver for WFCL has shown that in all programme countries high-interest informal moneylending is prevalent, and a particularly serious problem in Myanmar (Idris 2020). The scoping work conducted in the programme location of Hlaing Tharyar on the outskirts of Yangon found that many moneylenders are women, illustrating a strong gendered dimension to their role. This evidence is critical to assumptions about interventions linked to shifting social norms.

Another key finding from the scoping on labour intermediaries and labour recruiters relates to how children are recruited into sectors with high levels of WFCL. Counter to assumptions, the trajectories of girls entering the adult entertainment sector in Nepal, for example, is marginally linked to criminal networks that traffic girls, and more commonly is through the invitation of family and friends, highlighting the importance of complex informal recruitment processes linked to kinship ties. This was also found in the scoping of the leather sector in Bangladesh. This complexity in actors influencing a child’s pathway into WFCL should be integrated into our understanding and design of the interventions. Some of CLARISSA’s short- and medium-term outcomes are framed around improving awareness about risks of WFCL amongst parents, children, and other stakeholders. From the ToC revisions during the AARs, it became clear that this requires more nuance. In Nepal, it was highlighted that some programme team members feel girls are aware of the risks of working in the adult entertainment sector and so the lack of awareness is not considered a key driver for WFCL across all partners. In Bangladesh, this picture is further nuanced as children and parents are aware of some of the short-term risks (injuries, getting sick, missing school), but are less aware of longer-term risks (long-term health impacts, impacts of lack of education on future income and opportunities).

Furthermore, national stakeholders (perceived to have the most power) need to have an awareness about the extent to which WFCL is present in different sectors, as this is currently lacking. Based on this, the ToC sections on awareness should be refined to be about extent of awareness rather than merely aware/unaware and potentially contain nested causal pathways around how awareness influences behaviours for different actors.

The AARs also surfaced challenges to the underlying assumptions related to the CLARISSA’s key programme activities of generating evidence and innovative interventions through PAR. In Bangladesh, it was questioned whether PAR will unlock the required creativity that can lead to people developing innovative interventions. It was commented that people are used to only thinking about the things that they see right in front of them and what has happened before and are more likely to repeat this than think of doing things differently. For the ToC, this means that we need to better evidence the mechanisms by which PAR can generate innovative interventions and make sure that these mechanisms are facilitated through how the PAR groups are set up and evaluated. Finally, it was questioned how generating evidence will lead to increased awareness, especially for those stakeholders who are not directly involved in the evidence generation. This requires us to further detail the causal pathway between evidence generation and awareness. This should go hand in hand with the previously mentioned refinements about the extent of awareness among different stakeholders.

We are learning that in such a large, complex, and nested programme, taking time to critically reflect on how our overarching assumptions about PAR at the programme level translate to the local settings with the teams who are programme co-owners means that we will have stronger country-level ToCs that are based on informal evidence. Indeed, revisiting of assumptions in this contextualised approach illustrates the value of the nested approach to ToC. These reflections on the ToC further show us two ways that it will likely lead to improvements in it: (1) by using localised and up-to-date data which means that assumptions can be revisited to better reflect what is happening on the ground; and (2) by those on the ground critically questioning the assumptions of the ToC to help identify where it needs further detailing and evidencing. If we had set the country level ToCs too early and not incorporated emergent evidence to thus refine the ToCs based on critical questions, the CLARISSA programme activities could have focused on the wrong outcomes and risks (e.g. awareness of risks) leading to a reduced impact of the programme.
3.3 Navigating process and content as we ensure co-ownership of the rubric

The partnership rubric was used during the first AARs in Bangladesh and Nepal\(^9\) and at the consortium level for reporting purposes. During the AARs in-country, representatives from each partner organisation were asked to reflect on how they, as an organisation, engaged in the partnership in-country and are enabling quality using the generic CLARISSA rubric. Each partner then presented their self-assessment and the facilitation team brought them together to highlight differences and similarities across the partners. At the consortium level, however, each partner organisation was asked to reflect on how they think the partnership as a whole is functioning, again using the rubric as a guide. During these exercises it was emphasised that each team should provide evidence of why they perceived the partnership as functioning on a specific level and how this linked to

---

\(^9\) This exercise was not completed in Myanmar because there is currently only one in-country partner.
the indicators. Evidencing the ratings in the rubric is an important element to contribute to the rigour of the rubric evaluation (Davidson 2005). The outcome of this first round of applying the rubric is shown in Box 3.

From this first round of application, we have found the rubric to be a promising tool to generate data to inform PAM by identifying spaces for improvements for the individual organisational teams and for the partnership as a whole. For example, in response to the consortium-level evaluation through the rubric, IDS (the lead partner) responded with a desire for greater clarity in lines of accountability and is making more effort to ensure that decisions made at different levels are communicated more broadly. In practice, this means that key decisions are now being shared through a monthly bulletin.

The process of using the rubric received mixed feedback from different partners in the different countries. In Nepal, partners felt it worked well and set the scene for building a culture of trust within which open discussions about the partnership are enabled. Specifically, the exercise highlighted that different levels of communication and different understanding of the impact orientation of the programme were linked to the number of staff that partners have engaged in the programme. The self-assessment mode of application made it possible for partners with less of a stake in the programme to indicate their desire for greater inclusion in a safe way. In Bangladesh, while it was found to promote discussion on how the partnership is working in a safe space, there was also a greater desire to contextualise and specify indicators to the Bangladeshi country programme context. This is potentially due to Bangladesh being the largest country of operation with a diverse team across multiple organisations whose members are learning new ways of working together, so the rubric being more specific to their relationships would make the tool more useful.

However, the evaluative descriptors in the rubric were interpreted by some partners as indicators that all need to be present for the partnership to be rated at the performance level. This illustrates confusion around how to apply the tool in a self-assessment mode, which is intended to stimulate reflection and discussion across partners. The programme in Bangladesh is more complex than in Nepal and as the teams were being built at the time of application, the exercise surfaced areas of confusion around the roles of different partners. This illustrates that reflection on how we work together can lead to insights regarding who we are in the partnership relationship, which is critical for strengthening relationships in a research-for-development programme.

Our learning after this first application of the rubric as a tool to evaluate the partnership is that it requires a balance between, on the one hand, a facilitated process, self-assessment, and co-ownership by partners, and on the other, attention to specificity of the descriptors and evidence that partners bring to bear on the evaluation. This is not an uncommon dilemma with participatory methods and our learning reinforces evidence that refining methodological design through a facilitated process, albeit at times rather messy, can foster greater ownership (Frauenberger et al. 2015). Consequently, the next step is for the country teams to contextualise the rubric to their setting – to make their own version and use of the tool for their own learning and adaptation – and at consortium level to refine the descriptors through application in a collective assessment during the first annual AAR. With new individuals joining the team, it will provide an opportunity to ensure they become familiar with the tool, helping to understand the evaluative descriptors’ ‘meaning’ in an appropriate manner in order to better analyse the partnership, its strengths, and points that need to be improved, and position themselves based on available evidence.

---

10 The Bangladesh programme includes more partners, and a much larger budget due to the piloting of the social protection intervention.
Section 4:

IMPLICATIONS FOR PARTICIPATORY MEL PRACTICE
In this section, we discuss implications of our early learning on the process of designing, operationalising, and implementing a complexity-aware MEL system in a large participatory research-for-development programme. During the early implementation of PAM in the programme it became clear that this is not a way of working that comes naturally to everyone. Consortium partners commented on their organisation being challenged with the level of flexibility that is required for AM to work, when their systems are not set up for this. Our learning in this space reflects other experiences in the field that AM requires organisations to change their systems beyond project management in different areas of their structure, from human resources to finance.

In addition to changing organisational systems to make AM work, it is also important to develop a supportive culture and incentives to make AM work. An important element of AM is reflexivity and to openly discuss mistakes and take risks to make changes to how the programme is developed. During early discussions about consortium partnership in CLARISSA, explicitly creating such a culture was agreed upon and integrated into the partnership evaluation rubric as an important element of the partnership (‘openness, honesty, and mutual trust’ and ‘adaptability and flexibility’). Moreover, using the rubric for partnership evaluation was perceived in this early phase to move towards supporting open and honest discussion, which helps strengthen the partnership and AM. So not only do we see emerging evidence that the evaluative rubric is beneficial to strengthen the partnership, but it could potentially contribute to strengthening a supportive culture for AM.

There is great potential in AM as a tool to strengthen the links between MEL processes and operational aspects by using data generated through MEL in adaptive decision-making. Unfortunately, in the NGO sector MEL and operational activities are too often separate processes. Outputs of MEL activities are primarily used for reporting and accountability purposes, rather than informing internal decision-making processes. This practice is often related to the requested compliance with standards and procedures of the main humanitarian sector donors. The NGO sector is structured taking into consideration both internal and external strategies (like donors’ strategies, national government or international plans regarding development sector). AM can take into consideration these internal and external ‘constraints’ in order to facilitate the integration of MEL processes with operational aspects. AM can form a bridge between MEL and operational aspects as project managers can use MEL-generated data to inform their decisions on budgets and logistics, day-to-day management activities, and also to course correct if need be. AM can introduce a new method of work in the international and national NGO sectors, strengthening the dialogue between MEL staff and project managers, logistics and administrative staff, and in particular, staff working on the front line with the beneficiaries. The PAM we are building in CLARISSA takes this even further by being people driven, rather than problem driven, which means that it has the potential to make even better use of data generated through MEL processes. This is particularly so in CLARISSA where the evidence used to make AM decisions is based on an expanded evidence typology that encompasses evidence that has been formally published, is practice based, has been co-produced, and is physical or material. Including bottom-up, real-life, and on-the-ground experiences as evidence is one of the cornerstones of this innovative approach to AM. As noted already, we will maintain a focus on how we integrate different forms of data in decision-making by project managers to evaluate if and how a research-for-development programme can be radically participatory.

Reflexive use of ToC and PAM are complementary processes. As highlighted in this Working Paper, reflexive use of ToC means that it is further refined when new information emerges, and this information is generated through monitoring and evaluation processes. This, in theory, should lead to programme actions also being adapted. Changes in the ToC, we argue, is evidence that should fuel AM. Vice versa, data used to inform AM can also be used to reflect on the ToC. As described in Section 3.2, in CLARISSA the participatory reflection on the ToCs has uncovered important information about the direction of the programme (e.g. stepping away from awareness raising), which confirms that if a ToC is set too early, the implementers might be chasing the wrong activities and outcomes in their programme (Douthwaite et al. 2017). PAM can facilitate the reflexive use of ToC when there is an explicit culture of reflexivity and adaptability. From our early learning it is clear that for some, working with emergent programme design and embracing uncertainty is uncomfortable. This is a first step in our journey of learning from practice and we expect future learning to be useful to practitioners who are increasingly acknowledging that uncertainty is a key element of complexity.

4 IMPLICATIONS FOR PARTICIPATORY MEL PRACTICE
Section 5:

REFERENCES
REFERENCES


Burns, D. et al. (2020) CLARISSA document, unpublished, IDS


THIS PAGE IS INTENTIONALLY LEFT BLANK
Section 6:

ANNEXE
## ANNEXE 1: CLARISSA PARTNERSHIP RUBRIC AS DEVELOPED IN APRIL 2019

<table>
<thead>
<tr>
<th>Element</th>
<th>Well-functioning</th>
<th>Emerging</th>
<th>Needs help</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Communications</strong></td>
<td>Partners are clear on how the programme is progressing. All partners use teams seamlessly. Regular communication through multiple mechanisms.</td>
<td>Communication is haphazard and sometimes causes confusion. Without regular face-to-face meetings we would not be on the same page about key decisions.</td>
<td>Disagreements due to misinformation leads to conflict. Some partners feel left out or unsure of what is happening. Country-level teams are confused by mixed messages.</td>
</tr>
<tr>
<td><strong>Team identity</strong></td>
<td>Decisions are reached through consensus. Productive and enjoyable working environment. Clear definition of roles helps us work as a team.</td>
<td>There is mutual respect, but this remains formal. People work well together but do not necessarily trust each other.</td>
<td>Each partner focuses only on what is in their contract. There is no mutual support between partners. Partners feel they can make unilateral decisions.</td>
</tr>
<tr>
<td><strong>Openness, honesty, and mutual trust</strong></td>
<td>Problems are identified, shared, and discussed openly. We have positive personal relationships. We handle crises without internal conflict.</td>
<td>Some partners feel apprehensive about sharing honest opinions with the whole group.</td>
<td>There is conflict due to problems not being resolved.</td>
</tr>
<tr>
<td><strong>Impact orientation</strong></td>
<td>Agreed ToC provides clear vision and priorities. The MEL system is co-owned by all partners and it delivers quality information on how we are progressing along impact pathways.</td>
<td>There are frequent conversations between partners about the common vision because it remains unclear.</td>
<td>Activities are not aligned with the programme ToC. Partners are not aware of how their work supports the impact strategy of the consortium as a whole.</td>
</tr>
<tr>
<td><strong>Inclusivity and equitability</strong></td>
<td>Good dialogue that enables all to engage. Smaller organisations feel they have full voice in decision-making processes.</td>
<td>Roles require ongoing clarification.</td>
<td>IDS dominates the consortium decision-making. Smaller partners do not feel valued equally.</td>
</tr>
<tr>
<td><strong>Adaptability and flexibility</strong></td>
<td>Programme stays on track through making evidence-based decisions to adapt. Mistakes are openly discussed.</td>
<td>There is some adaptation along the way, but it is not well documented.</td>
<td>We never deviate from original plans. Budgets never shift throughout the programme.</td>
</tr>
<tr>
<td><strong>Entrepreneurial culture</strong></td>
<td>We find creative practical solutions to problems.</td>
<td>We have lots of new ideas but struggle to find ways to implement them.</td>
<td>We implement the plan without new ideas emerging. There is fear to take any risk.</td>
</tr>
</tbody>
</table>

Source: Apgar et al. (2019).
CLARISSA works by co-developing with stakeholders practical options for children to avoid engagement in the worst forms of child labour in Bangladesh, Myanmar, and Nepal.

The participatory processes which underpin the programme are designed to generate innovation from the ground which can sustainably improve the lives of children and their families.

The programme’s outputs are similarly co-designed and collaboratively produced to enhance local ownership of the knowledge, and to ensure that our research uptake and engagement strategy is rooted in the direct experience of the people most affected on the ground.