The European Journal of Development Research (2023) 35:298–322
https://doi.org/10.1057/s41287-023-00574-0

SPECIAL ISSUE ARTICLE

Theory of Change in Complex Research for Development Programmes: Challenges and Solutions from the Global Challenges Research Fund

Sarah Chapman[1] · Adiilah Boodhoo[1] · Carren Duffy[1] · Suki Goodman[1] · Maria Michalopoulou[2]

Accepted: 13 December 2022 / Published online: 28 February 2023

Abstract

The United Kingdom Research and Innovation (UKRI) Global Challenges Research Fund (GCRF) aimed to address global challenges to achieve the United Nations (UN) Sustainable Development Goals through 12 interdisciplinary research hubs. This research documents key lessons learned around working with Theory of Change (ToC) to guide Monitoring, Evaluation and Learning (MEL) within these complex research for development hubs. Interviews and document reviews were conducted in ten of the research hubs. The results revealed that only one hub invested in an explicit visual system mapping approach, and that funder timelines, budget constraints and issues with capacity and expertise limited the application of these approaches across all hubs. In contrast, many hubs attempted to deal with visual complexity by means of ether constructing multiple, nested ToCs, or a conscious simplification of complexity through reducing their ToC towards a straightforward and uncomplicated chain model or spherical model. While the former approach had some value, most hubs struggled to find capacity to support the full articulation of nested ToCs. In contrast, the latter approach resulted in ToCs which lacked detail or mechanism articulation, but which nevertheless were often 'fit for purpose' in ensuring effective communication and coherence across diverse stakeholders and sub-projects. We conclude that in instances where the reporting, funding and management cycles of complex research for development programmes cannot be adapted to properly support learning-based approaches to ToC development, imposing simplicity in the ToC might be fit for purpose. This might also be preferable to more complex visual approaches that are only partially realised.

Keywords Complex systems · Theory of Change · Research evaluation · Evaluation · Logic model

✉ Sarah Chapman
Sarah.Chapman@uct.ac.za

Extended author information available on the last page of the article

Résumé
Le Fonds de recherche sur les défis mondiaux (GCRF ou Global Challenges Research Fund en anglais) pour la recherche et l'innovation du Royaume-Uni visait à relever les défis mondiaux pour atteindre les objectifs de développement durable des Nations Unies par le biais de 12 centres de recherche interdisciplinaires. Cette étude documente les principales leçons apprises sur l'utilisation de la théorie du changement (TdC) pour guider le suivi, l'évaluation et l'apprentissage (SEA) au sein de ces pôles complexes de recherche pour le développement. Des entretiens et des revues documentaires ont été menés dans dix des 12 pôles de recherche. Les résultats ont révélé qu'un seul pôle a investi dans une approche de cartographie visuelle explicite du système, et que les délais des bailleurs de fonds, les contraintes budgétaires et les problèmes de capacité et d'expertise ont limité l'application de ces approches dans tous les pôles. En revanche, de nombreux pôles ont tenté de gérer la complexité visuelle en construisant de multiples TdC imbriquées, ou une simplification consciente de la complexité en réduisant leur TdC vers un modèle de chaîne simple et pas compliqué ou un modèle sphérique. Bien que la première approche ait une certaine valeur, la plupart des pôles ont eu du mal à trouver les compétences nécessaires pour prendre en charge l'articulation complète des TdC imbriquées. En revanche, cette dernière approche a abouti à des TdC qui manquaient de détails ou d'articulation des mécanismes, mais qui étaient néanmoins souvent « adaptées à l'objectif » pour assurer une communication efficace et la cohérence entre les diverses parties prenantes et sous-projets. En guise de conclusion, nous incluons les cas où les cycles de rapport, de financement et de gestion de programmes complexes de recherche pour le développement ne peuvent pas être adaptés pour appuyer correctement les approches basées sur l'apprentissage pour le développement visuel de la TdC. Dans ces cas-là, il semble être adéquat et préférable en effet d'imposer une TdC simple plutôt que d'utiliser des approches visuellement plus complexes mais qui ne sont que partiellement réalisées.

Introduction

The 12 United Kingdom Research and Innovation (UKRI) Global Challenges Research Fund (GCRF) hubs are interdisciplinary research hubs that collectively represented a £200M planned investment by the UK government to address intractable development challenges faced by developing countries. The hubs' research priorities are closely aligned to the United Nations (UN) Sustainable Development Goals (SDGs). They tackle a wide range of complex global challenges from improving human health and promoting gender equality and social justice to fortifying ecological systems and biodiversity, generating agricultural sustainability, and fostering greater resilience to natural disasters. The hubs comprise up to 400 unique partner organisations in 85 countries, and 550 researchers from a range of disciplines addressing 16 of the SDGs (UKRI 2021).

The sheer scale and ambition of the 12 hubs presents both opportunities and challenges for evaluation. Each hub is complex: it is composed of eclectic thematic, geographical, and operational research streams believed to facilitate the emergence of conditions essential to inform global-level policies relating to intractable

developmental challenges. It is within the above context where a visual programme theory, referred to in this paper as a Theory of Change (ToC), comes in as an almost ubiquitous strategic monitoring, evaluation, and learning (MEL) tool (DuBow and Litzler 2019; Wilkinson et al. 2021). ToC approaches evolved from the perceived limitations of logical framework analysis, which became prevalent from the 1970s onwards in development practice, emphasising the need to involve a wide range of people in its design (Chen 2005); teasing out longer chains of mechanisms between activities and outcomes (Rogers 2000); and using the process of developing the ToC as an opportunity to regularly examine programme assumptions and embed programme learning (Douthwaite et al. 2003).

ToCs are often mandated by funders at different points during the programme cycle and serve to demonstrate planning for implementation and to secure funding. This was the case with the UKRI GCRF interdisciplinary hubs. However, this expectation came in the context of limited documented experience of the 'how' (the process) and 'what' (the product) of ToCs development and application in interdisciplinary research contexts (Deutsch et al. 2021), as attempts to incorporate complexity features into ToCs in practice are often undermined by the context in which these are being used (Wilkinson et al. 2021). Some programmes are *complicated*—they have very many parts occurring either concurrently or sequentially and attempting to depict this intricacy diagrammatically can rapidly deteriorate into a mess of criss-crossing boxes and arrows (Funnell and Rogers 2011). Other programmes are *complex*—they face the added challenge of accounting for elements of emergence, non-linearity and surprise in their ToCs. This can result in endless time-consuming reiterations and revisions of the ToC in an effort to capture the current programmatic reality—an endeavour of questionable value given that truly complex systems can usually "only be made somewhat coherent in retrospect" (Patton 2015, p. 422).

For the evaluation of the GCRF, it is acknowledged that both complex and complicated aspects are likely to be present in hubs (Barr et al. 2019, p. 14), and for these reasons it is suggested that at a high level a ToC diagram might be used to help navigate evaluation and learning in these domains. Our research sought to further our understanding of how hubs went about developing, visually presenting, and ultimately utilising ToCs to guide evaluation and learning within their initiatives. Very few multiple comparative case studies of complexity-aware evaluation approaches exist (e.g. Gates 2017), but to our knowledge none of these analyses have specifically showcased the comparative use of visual ToC. Other examples of complex and complicated visual programme theory presented in the evaluation research are described as illustrative rather than best practice (Rogers 2008). Our contribution, in contrast, is derived from the ongoing experiences of multiple practitioners working across the GCRF research and innovation hubs, including the hub which the authors are formally affiliated with. Our objectives were to categorise and interpret the use of visual ToC in the UKRI GCRF hubs, and to identify factors affecting the development, visual presentation, and successful utilisation of ToC in MEL activities.

Framing Complexity and Situating Visual ToCs

Complexity theory (located within the wider field of system thinking) is not a single coherent body of thought (Walton 2014). Our framing of complexity in this article resonates with that of Shiell et al. (2008). We locate complexity within both interventions as well as the social systems in which they are embedded. We share Hawe et al. (2009)'s view that interventions are 'events', which aim to "disrupt the functioning of complex systems through changing relationships, displacing entrenched practices, and redistributing and transforming resources" (Moore et al. 2019, p. 24). If we were to apply an 'intervention lens' to the GCRF interdisciplinary research hubs, these would be considered complex multi-component interventions requiring coordinated partnerships among multiple actors across multiple sites (including in-country research teams and partner organisations). While each hub has a unique thematic, geographical, and operational configuration, they all consist of multiple research streams embedded in complex social systems, which they seek to transform.

The process of constructing a visual ToC puts into sharp focus the reality of complex interventions embedded in complex systems and often challenges stakeholders' overly simplistic assumptions of change processes. Visual or diagrammatic representations of ToCs vary in shape and form—they can be nuanced and nested (with multiple parallel and intersecting causal pathways) or simplified and linear (Wilkinson et al. 2021). Davies (2018) presents more than 30 visual depictions of ToCs. The visual complexity of a ToC often mirrors the varying expectations imposed by the context in which it was developed and the purpose for which it was developed.

Several specialised visual mapping approaches have been suggested as an extension to traditional ToC approaches to guide evaluation of complex programmes. These include social framework analyses and actor-based change models (Koleros et al. 2020); systems dynamics approaches which centre around the development of causal loop diagrams (Dyehouse et al. 2009; Morecroft 2020); the use of cognitive mapping and cause mapping as part of Strategic Options Development and Analysis (SODA; Ackermann and Eden 2020); soft systems methodologies which draw extensively on visual diagramming to articulate a learning cycle (Checkland and Poulter 2020); and participatory systems mapping (Wilkinson et al. 2021). In our research, we were interested to see the extent to which these visual mapping approaches would be adopted by the hubs to expand on their visual ToCs. We were also interested in exploring the extent to which the development of a good representation of a ToC suffered from issues of representation accuracy—which is one key dimension in Dhillon and Vaca (2018)'s rubric of ToC robustness spectrum. These common deficiencies in published ToCs (e.g. uncaptured or loosely connected impact pathways, assigning the same 'weight' to outcomes, absence of feedback loops and 'aesthetics over substance') have been extensively discussed in literature (e.g. Davies 2018), along with strategies on how to upgrade the visual presentation of ToCs (e.g. Dhillon and Vaca 2018). Yet while these guidance documents exist, there is little published about how practitioners grapple with developing and utilising a ToC (DuBow and Litzler 2019).

The UKRI GCRF Interdisciplinary Hubs: Some Caveats

The GCRF call for proposals was met with about 250 proposals, which underwent a rigorous review process over a period of 16 months. At the end of this highly competitive process, 12 interdisciplinary research hubs were funded for 5 years. In March 2021, the UKRI announced that projects under the GCRF were to be defunded as result of a significant reduction in the 2021/2022 international development budget following the economic impact of COVID-19. This announcement left the hubs with many processes severely disrupted, including evaluation. Nevertheless, we have sought in our article to document the approximately 2.5 years that activities were undertaken for the first half of the project, before funding was disrupted. While it is unfortunate that we are not able to reflect on the completed project duration, the inception period documented is usually a very active time for ToC development and provides important insights into our research area.

Method

Research Design

The GCRF project is ongoing, and the authors were also embedded as practitioners in this initiative. We therefore adopted a research methodology that acknowledges the significance and usefulness of involving ourselves, our research, and our practice partners in the knowledge-production process. We adopted a pragmatic approach which utilised a descriptive case-based research design to describe, explain and validate our findings in order to provide practical solutions and useful applications to real-world issues (Patton 2014).

Our descriptive design drew on both primary and secondary data sources. Secondary data was in the form of existing GCRF hub documentation and reports. Primary data sources were qualitative interviews conducted with Key Informants (KIs) from within the GCRF hubs.

Participants and Sampling

We used purposive sampling to identify participants, with the option to add additional KIs by means of snowball sampling as the study progressed. All KIs comprised the persons tasked with MEL within respective GCRF hub projects where roles are variously performed by programme managers and/or MEL officers, depending on the hub structure. 10 of the 12 hubs agreed to participate in interviews. We conducted a total of 12 interviews with 13 individuals from across the 10 participating hubs (1 interview was a group interview with 2 participants present).

Procedure and Data Collection

We gained permission to access and review available secondary data in the form of existing UKRI GCRF hub MEL frameworks, documentation and reports, including all visual programme theories and ToCs. All 12 hubs agreed to share this documentation.

We worked through our GCRF collaborative networks to invite (via email) persons involved with MEL within the 12 GCRF hubs to participate in a KI interview. The practitioners contacted were aware that several papers were envisioned for a special edition on the collective experiences in the GCRF hubs and that one such paper was based on the challenges and solutions for utilising visual programme theories for complicated and complex interventions. They had also shared secondary data with the authors for the initial document review. The invitation reiterated that the purpose of the interview was to discuss how the ToC and framework was developed with a view to unpack lessons learned in the process.

Interested parties were then sent an informed consent form and invited to partake in interviews as KIs with members of the core research team, who used open-ended interview questions to guide the engagement process. KIs were asked to walk the interviewer through the development and review processes (what activities were followed, who was involved, timing, etc.); to comment on the GCRF guideline documents and their usefulness; to discuss the visual representation of the ToC and how it was utilised for the framework; the challenges encountered and any innovations and strategies they utilised during the process. Lastly, KIs were asked to reflect on their overall experience and perspectives of the project. Interviews were typically one hour in duration, and all the interviews were conducted in English via an online platform such as Zoom, or Microsoft Teams.

Data Analysis

All members of the core research team were evaluation practitioners with expertise in qualitative research methods. NVivo (QSR International) was used to analyse and organise the data, and to coordinate feedback on the data between the research team. The analysis of the primary interview data was supplemented by the inclusion into the NVivo project file of the secondary data in the form of existing GCRF hub documentation. Our research approach was interpretative, and thematic content analysis was used to both deductively or inductively identify and present research themes in the dataset (Braun and Clarke 2006).

For our deductive analysis, we developed an initial coding framework based on the literature, and our experience as evaluation practitioners familiar with the evaluation requirements and processes instituted by GCRF. For example, based on the literature of visual programme theory, we pre-identified several potential visual structures for programme theory that ranged from the simple to complex (box and arrow, onion, spheres of influence, causal loops, nested programme theories, social impact diagrams, network diagrams, etc.). We also predefined codes based on the

key research elements we were interested in exploring, such as the intersectionality between a logical framework (log frame) and the visual programme theory.

Programme documents were coded inductively and deductively. This included visually importing image files of the hubs' programme theories and coding the relevant visual elements. Interview transcripts were then added as primary data into the NVivo project file. The first author conducted the initial coding of all interviews using a combination of inductive and deductive coding, and a system of memos and annotations as initial themes and key excerpts were identified. The third and fourth authors then read all the transcripts, reviewed the coding structure, and made suggestions to improve and revise the codes and add additional memos. A first draft of the thematic map and associated results narrative was then produced, and later iteratively revised following feedback from the third and fourth authors on drafts of the paper.

Reflexivity

Three of the authors served as evaluation and learning practitioners for one of the 12 hubs. Authors one and two were contracted (after the funding stage), to revise the ToC, logical framework and evaluation approach for the reporting and stage-gate review phases of the project. This research was born from their experiences in the project. The authors wanted to investigate how other hub practitioners managed the development and review process and what could be learned from the project.

The two authors' proximity to the project was beneficial in shaping the research agenda, but at the same time, the authors wanted to ensure that the data collection, analyses, and reporting were not limited based on their experiences. To achieve this, several strategies were implemented for data collection and analysis. First, a semi-structured interview schedule was developed. Second, four authors served as interviewers and were randomly assigned KIs from the various hubs. Third, one of the authors interviewed an author who was a hub evaluation practitioner to ensure that their insights were documented independently from the others. This interview transcript was analysed with the others. Fourth, in terms of coding, three of the authors of this paper (two of which had no hub involvement) were involved in the development of the themes from the interviewee transcripts. Fifth, a number of the codes were supplemented with objective evidence from the secondary data utilised. Finally, and most importantly, the authors who also served as practitioners continually monitored that their own experience was not impacting on the way they engaged the data. They did this through intentional reflexive practices, acknowledging and working with their subjectivity throughout the research process.

Ethical Considerations

The research was approved by the University of Cape Town's Commerce Faculty's Ethics in Research Committee. We secured permission from the hubs to access and review the MEL reports and documentation. All sources were anonymised, and respondents also were provided the opportunity to review,

correct, and edit full interview transcripts prior to them being incorporated as primary data in the research process. This allowed respondents to check the accuracy of their transcript, edit or remove harmful or inaccurate content, and protect their institutional and professional reputation considering any interview excerpts being published.

Results

Funder Expectations and Guidelines for ToC

The GCRF mandated the development and presentation of a draft ToC even at the proposal stage, and a fully developed ToC was required within the first six months of inception. Guidance at the proposal stage was a brief two-page document, suggesting a simple chain model of inputs–activities–outputs–outcomes and impacts (Fig. 1).

During the 6 months following project inception, successful applicants were invited in mid-2019 to attend a 2-day training exercise on ToC and evaluation practice. The term 'Monitoring, Evaluation and Learning' (MEL) was used formally for the first time, with the role of ToC emphasised as informing all these components. Also, in mid-2019 came the publication of guidelines for the evaluation of the GCRF initiative (Barr et al. 2019), which were helpful in conceptualising how the GCRF might bring about change, and in what way this change might ultimately be measured. Specifically, Barr et al. (2019) propose that a ToC for the GCRF should visualise change at different spheres of influence, which were understood to help to "represent the complexity of the GCRF change process, and the degree of GCRFs agency to influence change at different levels" (Barr et al. 2019, p. 1). A sphere of direct influence would cover GCRF project activities, outputs, innovation development strategies and intermediate outcomes. Outcomes in this sphere were likely to be 'complicated' to evaluate and would be amenable at least to some extent to 'straightforward' linear evaluation approaches (Barr et al. 2019, pp. 28–34). A second sphere of indirect influence represented the "extent to which evidence and innovation products are replicated and amplified within different policy, practice and market settings" (Barr et al. 2019, p. 10). Outcomes in this sphere were likely to fall in the 'complex' domain, and complexity-aware evaluation and learning approaches would be required for these aspects (Barr et al. 2019, p. 30). At all times, the ToC was indicated to play a critical role in evaluation design and planning in that the stages of a GCRF evaluation "will mirror the progress of the GCRF along its ToC trajectory" (Barr et al. 2019, p. 34).

In the sections which follow, we turn to our primary interview with hub practitioners to reflect on how the ToC process for the GCRF hubs was operationalised in practice. Four main themes emerged relating to factors affecting the development, visual presentation, and utilisation of ToC. Each of these themes had several subthemes, which are presented in Fig. 2.

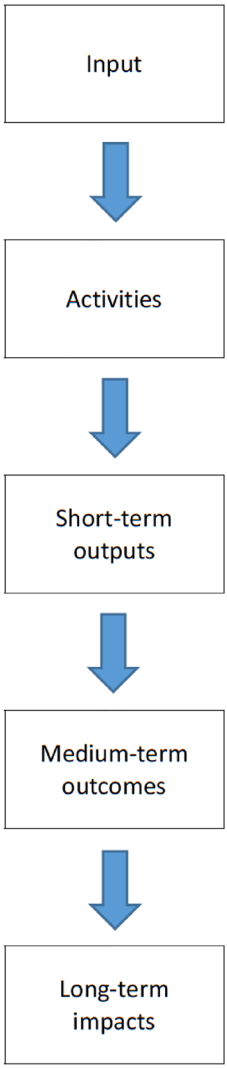

Fig. 1 Simple chain model suggested by the GCRF in their proposal development guidance document

Theme 1: Struggling with MEL Requirements

Time

Six interviewees from four hubs cited not having enough time to properly develop the ToC. Interviewees who cited this concern often referred to the processes of relationship and consensus building, which were enabled by the engagements related to the development of a ToC. For example, this hub manager reflected that:

> I think that should have been a much longer period of time to develop the theory of change because it's not just about a picture, it's about the process.

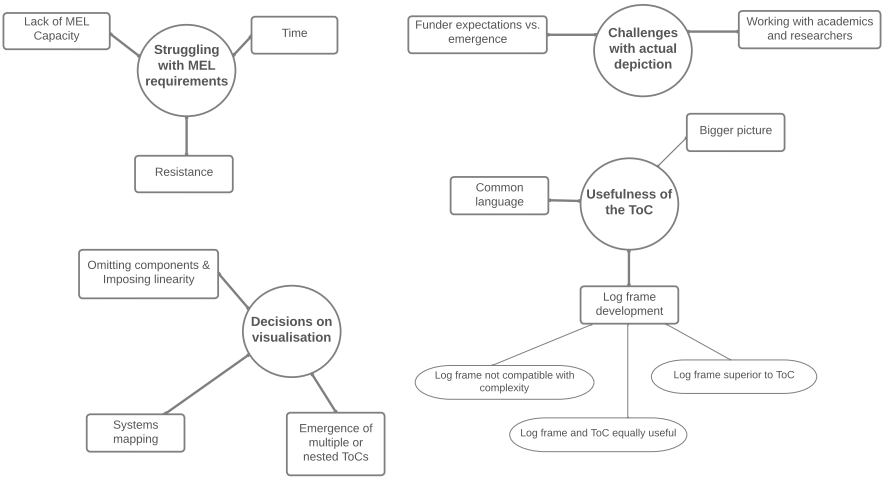

Fig. 2 Thematic map from qualitative data analysis

> It's about engagement, it's about working with people to develop it. [Interviewee #1]

Another hub manager echoed similar sentiments:

> So, people keep saying, 'tell us about interdisciplinary work', and I say, 'people have to know each other'. We spent two years just establishing trust. [Interviewee #2]

This lack of time to establish trust and collaboration was a major source of frustration and reflection. As this hub manager indicated:

> OK, so honestly and candidly. We wish we would have done it differently 1000 times over… it was actually a very unrealistic list of things to do in six months. [Especially] considering that all these hubs are so international with partners all over the place. [Interviewee #3]

Resistance

The complex and interdisciplinary nature of the GCRF projects typically resulted in forms of resistance. The MEL practitioner was often at the centre of this resistance. Many of the hub MEL leads found working with academics, scientists, and researchers on a ToC to be a unique experience that was not without its sources of tension. For example:

> There was divergence because in my view, certainly, the physical scientists within the programme were just, 'what the heck is this, it's a load of rubbish '. So there was a lot of resistance. [Interviewee #1]

And again:

> They [the academics] were a lot more hesitant to accept the mind game of thinking about impact in not just purely technical, but also like imaginative ways. So that was, more of a struggle to bring that across. [Interviewee #4]

Another interviewee went on to explain discipline-specific variations in communication style can render the process difficult.

> I don't know if you've ever worked with [scientists], but they think in boxes and it's very strict, and very rarely is it structured the way that people tend to think. [Interviewee #5]

Another interviewee notes how the benefits of working with ToC came only after some resistance, as the need for a ToC was seen as largely imposed upon the hub by the funders:

> So, I would say, 'OK I get it, you don't want this, this, big brother's eye watching you'. But think of it mostly as a plan for working together. [Interviewee #2]

Other interviewees spoke of the difficulties and frustrations inherent in operationalising a complexity-aware evaluation process more generally. For example, in this hub:

> We talked about it being complexity aware, I think we said that from the get-go. The co-directors at the time accepted the language, but I think had no idea what it really meant. [Interviewee #6]

This hub programme manager puts their frustrations even more succinctly:

> We're not a development project, and the Overseas Development Assistance framework is sort of set us up as if we are, and we're not. Research has a very tenuous relationship with impact, as you know. I think fundamentally, the approach that the GCRF has taken to the evaluation of everything is just, so bizarre. [Interviewee #7]

Lack of Capacity

Ten interviewees from eight hubs cited capacity issues. This was especially pronounced during the inception phase of the projects. In two hubs, interviewees explained this in terms of characteristics of research funding. One interviewee describes how hub resources were:

> Put 99% into what I would call hardcore research and 1% into communications, operations, evaluation learning and monitoring. There was no respect, you know, from my perception of the difference that [adequate investments] could make. [Interviewee #8]

In another hub, the interviewee explained that:

> In the funding call stage, they needed us to outline those resources that will be put aside for MEL, and a ToC is not part of the traditional way of doing

> research projects, which still mostly weren't working with ToC. [Interviewee #1]

Lack of capacity resulted in evaluation and learning strategies being impartially or inadequately implemented. As this interviewee explains:

> If I could redesign the hub, I think I would have somebody whose job it was to worry about theory of change and not add it to my 10%-time allocation - so that's the thing to remember. [Interviewee #6]

In three of the hubs, the programme managers complained about how the expectation to 'do' MEL was thrust as an awkward surprise on them. As one hub manager lamented:

> I thought my research would be feeding into MEL, but upon arriving I realised that all MEL would be done through me. [Interviewee #2]

Another hub manager noted that:

> The donor expects you to perform MEL on top of everything else as well as managing a very complex research programme. I feel we're not really doing a particularly good job on that in the sense that we just don't really have sufficient time and capacity to do it properly. [Interviewee #5]

And analogously:

> We don't have any capacity for evaluation in the budget, because we never realised that was part of the deal. [Interviewee #7]

Theme 2: Challenges with Actual Depiction

Unrealistic Expectations vs. Emergence

Interviewees spoke of tensions between the funder's expectations and the mandate of a ToC, which is supposed to allow exploration into change assumptions and relationship building:

> We had to make it fit on a one-page picture, but we have [a number of] different partners, with incredibly different contexts, political contexts, areas of work… and you're supposed to summarize this in the one-page picture? It makes a mockery of the complexity of what we're dealing with. [Interviewee #7]

Another interviewee goes on to explain how this anxiety around visually simplifying the complicated aspects of the ToC resulted in people shying away from exploring the full depth of innovations within the system:

> People tended to state ambitions which were obvious and achievable as opposed to the things that they actually wanted to do - which was to push the boat a bit further out. And so even though they talked about their very exciting

> ambitions and dreams, that wasn't being reflected in the [visual] documentation. [Interviewee #8]

In this hub, the interviewee explains the pressure points in the GCRF relationship created by the tension between the ToC being 'fixed' (and thus, what the project would ultimately be held accountable to) and at the same time the need to allow for emergence. This culminated to such an extent that the programme manager felt compelled to "get in writing" the agreement from the funder that the ToC would be 'allowed' to evolve:

> There are so many variables that there was a lot of anxiety around the fact that we were stating that we were going to do this. It took a lot of explaining that the ToC and the log frame were going to be revisited on an annual basis, and that we weren't held to this. And actually, it wasn't until we had that from the UKRI in writing that we were really able to persuade some people of that. [Interviewee #8]

In another hub, an interviewee described a similar tension. They explained how the GCRF as funders were putting unrealistic expectations on hub managers, with respects to ToC, in terms of acknowledging complexity on the one hand, but then not allowing for uncertainty and emergence on the other; in other words:

> And another thing having this requirement, you know, the GCRF said they want to know it all: the good, the bad and the ugly. And that was very freeing. However, it was counteracted a bit by saying 'OK give us your ToC and make it rational'. [Interviewee #2]

This tension between programmatic reality vis-à-vis funder accountability was also prevalent in this hub, as this interviewee explains:

> Originally, I was kind of told that because the funder had approved this ToC, this was the ToC that we were going to deal with, and I shouldn't maybe necessarily try to change it too much because it's been approved. [But] from my perspective, maybe it wasn't the most helpful representation of what our ToC could be. [Interviewee #4]

Working with Academics, Scientists and Researchers

Interviewees described how working within research for development programmes is characterised by some tensions between academics, researchers, and MEL practitioners. Some academics were reluctant to outline a ToC before they have commenced the empirical research process:

> With the academics, I mean that particular group that I'm thinking of, their reluctance was sort of, 'well, we can't think about the long-term impact if we haven't done the research yet. Like how are we going to know what is going to come out and what is going to change if we don't know the result of our work yet?' [Interviewee #4]

And again, from another interviewee:

> It was the sense that the UKRI is trying to determine what we do. And because [our collaborators] are academics, they read around this. I was trying to encourage them to think about it more as a plan that we use from time to time rather than a way of policing them. [Interviewee #2]

And relatedly:

> [If I could do it again] I would have worked with the ToC almost in isolation to everything else for many months. I would have actually really worked with a ToC with them [the scientists], and tried to embed their sort of science [in the ToC]. But it was almost like a ToC was there, and the science was here, and I think if we tried to resource [more], we could have done it. [Interviewee #1]

This programme manager described how academics tended to resist the role of the programme manager, which in this interviewee's view was to keep the project's focus on the bigger picture. In the quotation below the interviewee describes how a particular academic tended towards a complicated ToC which might have been accurate in terms of the intricacies of the mechanisms that partners were working with, but was "unusable" from a management and communication perspective:

> The issues is that academics will all over-complicate. So in the [particular partner's] case, there were circles here and circles there, and I looked at it, it's a dog's dinner, its incomprehensible. It could never be used for communication purposes and engagement purposes for trying to describe you know, this is where we're trying to get to, these are the components. [Interviewee #8]

Theme 3: Decisions on Visualisation

Imposing Linearity and Omitting Components

In one hub, this interviewee describes how the first version of the ToC they received from a consultant was "hideous", and "totally incomprehensible – a mishmash of arrows and blocks". The final (revised) version of the ToC was ultimately accompanied by a written narrative that allowed for the elucidation of mechanisms, but which also could allow for a ToC that was:

> Something that was visually arresting. But not overwhelming and not frightening. We thought that was really important for buying in. [Interviewee #8]

One strategy was to explicitly omit components, including feedback loops, from the ToC:

> Our original ToC looks very messy. Which does reflect [the] impact. It actually works better [than the current ToC]. But in our experience with the funder, it's often easier if you have a somewhat streamlined looking document. And so, I think that's why we went in the later iterations more towards this linear

> seeming ToC visualization, which ignores some of the complexity, of course… but yeah. [Interviewee #4]

Interviewees from other hubs also went on to elaborate how the conscious imposition of more simple, linear models and the removal of feedback loops can have some advantages in terms of communication and clarity:

> There's a lot of people that talk about linearity being a very, very bad thing to have in a theory of change. But actually, from my perspective having a simple map to get from A to B can sometimes be what we need [as opposed to] having too many arrows and things. It being circular can sometimes detract from what the purpose of the ToC is about, which at least partly is communication. [Interviewee #9]

In a second hub, this interviewee describes the decision to move towards multiple or nested ToCs through the metaphor of a solar systems within a complicated galaxy, that would otherwise:

> Be an absolute mess of arrows and things whizzing around all the time, and no human is able to actually conceptualize that. [Interviewee #11]

In this hub, this interviewee describes a process of consciously imposing linearity on a complex programme theory with many components and feedback loops. Imposed linearity in contrast provides a sense of overall intention to the initiative:

> It became a very frustrating exercise, because people tend to have quite strong views of how it needs to be and [the] feedback loops, etc. So, we impose that linearity, explained as being in the intention, so it's not that the [hub] system is linear, but you have an intention and that's more or less modelled linearly. [Interviewee #10]

In another hub, similar sentiments were echoed around the need for imposing simplicity on a complicated situation:

> There were quite a lot of different models proposed. And whilst we accept that change is not linear, it is quite helpful to stylize a process of change in a linear manner. We were really concerned that it was simple enough to understand. Logical enough to make sense. But ultimately the sign-off came from the Executive Board. [Interviewee #8]

Nested ToCs

Another strategy to represent complicated components visually was to create multiple ToCs linked to an overall hub-specific ToC in a layered, embedded or otherwise nested manner. When this approach was chosen, the hub-level ToC became a relatively static high-level document, whereas the nested ToCs could be more flexible and iterative tools that more accurately represented the dynamic and co-evolutionary relationships within the hub structures. One interviewee referenced using this

approach to resolve tensions in the differences between research versus more traditional development programmes:

> The programme had initially been organized according to research programs, but that's not helpful for us to plan. You know, if you're going to have impact, research as a unit of analysis does not make sense. So, when we went to the training, that's a natural thing - the nested ToC. [Interviewee #2]

In two other hubs they reached the decision to diversify from an overall ToC to country-specific ToCs that represented initiatives in specific geographic zones. This hub emphasised how the personalities of the individual researchers leading the nested initiatives came through in the visualisation of these diversified ToCs:

> They are an expression of the cultures of the places and the people who were involved in developing them very much. [The researcher] is highly respected as a professional in that country in that context, and his diagram needed to be what helped him explain what they were doing. [Interviewee #6]

However, this process was not an easy solution, as this programme manager described the initiative:

> In theory, that sounds good, but in my experience and practice… disaster. We just didn't have the time or resources to do it [properly]. But it definitely could have worked in my view. I don't think the nested approach was wrong. I think it was very ambitious. [Interviewee #1]

The interviewee later elaborated:

> And [our partners] were like 'yeah, we get this', but they just didn't quite know how to use it, you know? [Interviewee #1]

Another interviewee (Interviewee #11) went so far as to describe the impartially executed attempts to create nested theories of change as "a waste of time", primarily due to lack of capacity. Although the interviewee indicated it would have been "ideal" to develop dynamic and nested ToCs, with hindsight it would have been far more useful given time and capacity constraints to simply re-orientate the work packages back to the original ToC that was developed at the hub level, to make sure that their individual efforts fed into the overall larger vision of the hub's impact.

Systems Mapping

In contrast to developing multiple or nested ToCs, one hub used the concept of visualised impact pathways to flexibly define specific impact mechanism chains broadly embedded in the hub-specific ToC. Starting out with an overall ToC, they then used systems mapping software to create multiple visual impact pathways to expand the ToC with groups of project partners. These visual impact pathways had the advantage of allowing for non-linearity and feedback loops, as well as iterative revisions using a participatory approach and visual mapping software. The hubs found this approach a critical programme management instrument given the project size and its

exploratory nature, as it allowed management to define more specific sets of goals as scoping studies were analysed and the project progressed. The impact pathway work was led by the senior project manager, with assistance from junior staff. However, this approach was noted to be resource intensive, and required the acquisition of an additional MEL staff.

Theme 4: Usefulness of ToC

All ten hubs cited specific advantages of working with a ToC, although some of these benefits were only evident after much investment of time and energy. Other benefits came with stated limitations. Benefits were broadly defined in terms of three sub-themes, elaborated below.

Common Language

Five hub interviewees described the benefits of ToCs in cutting across language and standardising communication. The metaphor of language was particularly pervasive:

> You need to enforce some sort of common language. It's like speaking English instead of all the local languages. It's not the best language at all, but it is a very useful language because everybody speaks it. [Interviewee #10]

The language metaphor was extended further by this interviewee through a biblical metaphor of the Tower of Babel, where communities would have been unable to cooperate on building a complex structure due to speaking different languages. The metaphor evokes barriers in the hubs' interactions with several country partners (different contexts), and across disciplines (different scientific approaches and methodologies) and how a successfully realised ToC process helped resolve this by bringing people together under a common vision:

> [Through the use of ToC] it becomes like a beautiful kind of natural system, and not like a Babel tower where nobody understands each other's language. [Interviewee #3]

One interviewee expressed their opinion that ToC was a suitable tool for broaching the communication divide between academics (research) and practice:

> One of the purposes of developing a ToC with clear points of intervention is to communicate with those [who] are designing interventions – and I'm talking about people working in policy and practice who are primarily outside of academia. [Interviewee #9]

This was supported by another interviewee who also picked up on the 'language' metaphor, emphasising how visual processes allow stakeholders to come together and communicate more effectively:

> The only thing that can allow you to succeed [in developing the TOC] without it being a cynical tick-box exercise is communication. You know those initial workshops were so important. To be able to see face to face where the misun-

> derstandings lie, to clarify the language, to challenge people to say, 'what do you mean by that'? [Interviewee #8]

Bigger Picture

Five hub interviewees described the benefits of ToC in terms to bringing the focus of complex interdisciplinary groups of collaborators back to the bigger picture. Having sight of the bigger picture was important in identifying commonalities and points of collaboration. As this interviewee explained:

> People see it as a tool for saying, 'oh ok, we have these things in common in these three countries, why don't we get together and do something and see what we can learn'? (Interviewee #7).

One hub interviewee described how the ToC helped in terms of "clarity over our mission". Another hub interviewee described how:

> It was really important that we bring focus back to the overall theory of change, because it's a tool to help conceptualize our eyes on the ball. Are we all working together towards the same aim? [Interviewee #11].

An interviewee from another hub saw it specifically as a project-level tool that was referred to as "aspirational", "the timeline", or "the big plan". However, at a more granular level, the ToC was too high level to be of much use as a project management tool. These limitations and benefits were admitted by another interviewee, who indicated that relative to a log frame ToC was only somewhat useful:

> The ToC I find is a picture in my head, and I find that very useful, but [for project management] I will always go to the log frame. [Interviewee #13]

Logical Framework Development

For all hubs, output reporting along predetermined categories was a requirement, and thus, logical frameworks or the like were central para-tools which co-existed with the ToC. The ToC was useful in some of the hubs in informing the logical framework design. However, attitudes towards the usefulness of the logical framework and the ToC in developing a programme monitoring system varied widely.

Logical Framework Superior to TOC In two hubs, logical frameworks were used as the primary MEL tool (more so than the ToC), and explicitly linking work plans to the logical framework was a core deliverable for each work package:

> We have an insanely detailed granular work plan for [all the] work packages that we have tied to the log frame really well. [Interviewee #5]

In one of the two hubs, the logical framework was used more than the ToC. Although the ToC provided the overarching narrative, the interviewee admitted of the ToC, "I don't particularly find it useful", given that:

> I primarily worked with a log frame, [as] I could see [then] which activities were taking place and which targets had been reached. But the TOC did tell the story. [Interviewee #13]

Logical Framework and TOC Equally Useful The majority (five) of the hubs had a complimentary approach, where little conflict was reported in the development and application of the logical framework, and the ToC and logical framework components of the MEL system worked quite well together:

> The way I see it is that the ToC is our top line ambition. Our impact narratives are how we're going to get that there, and our log frame is our check and balance tool - so it's the way that we assess our progress against our ambitions. [Interviewee #8]

These interviewees typically indicated that the logical framework was used to generate routine progress reports for hub management and funders.

Logical Framework Not Compatible with Complexity In three hubs, the logical framework was referred to as a source of frustration, which brought to question its relevance when dealing with a situation of adaptation and emergence. In the first hub, due to the complex manner in which policy change unfolded in the hubs:

> It quickly became clear to the teams this was kind of a waste of time. [Interviewee #6]

In the second hub, the interviewee described how the logical framework was essentially shelved once it was developed:

> I didn't detect any desire to go back to the log frame and touch it again. [Similar to] the ToC [when it's been approved] Now leave it, sort of thing. [Interviewee #4]

The final hub rejected the notion of a logical framework altogether; refusing to develop one despite the funder requirement. This was due, in their view, to the logical framework being completely with a complex systems approach:

> We can't use a log frame. That's how we see it, and we've made that case in our documentation to do so. Because change is not linear, it's complex and not unidirectional; because there's so much diversity in the context because of the political negotiations around a lot of these things. [Interviewee #7]

Discussion

The development of a ToC frequently requires practitioners to manage at least two competing, yet valid requirements: simplicity (to ensure readability and usability) and sufficient detail (to capture the complexity and ensure that the ToC is evaluable). In other words, ToCs are faced with the challenge of reducing complexities by creating complexities (Van Tulder and Keen 2018).

In their evaluation framework for the GCRF, Barr et al. (2019) provide some assistance in resolving these tensions by differentiating the ToC (and corresponding evaluation approach) into complicated and complex domains. The challenge came in that for complex domains, no explicit recommendations were provided by GCRF regarding suitable complexity-aware approaches to visually mapping the ToC. In addition, there was a tension around how these approaches, if adopted, would play out within the GCRF funding and reporting cycles. Given that these systems-based visual mapping approaches tend to use visual models as a source of questions to pose around the emerging real-world situation, the expectation is that ToC become a tool that "will surface worldviews and generate ideas for change and improvement" (Checkland and Poulter 2020, p. 202). For these reasons, tips for visual mapping typically urge practitioners to: "[not] worry about appearance. One of the most common concerns is that 'it looks so messy', but mapping is supposed to be a way to facilitate free thinking and conversation. This is contrary to the reporting or deliverable style requirements of the funders" (Ackermann and Eden 2020, p. 161).

Our research shed some light on how these tensions were experienced over a 2.5-year period in 10 complex and interdisciplinary research for development hubs. For some hubs, our research revealed how ToC was seen as an opportunity to engage more honestly with the complexity of change processes, and ToC was used as a way of facilitating thinking around complexity emerging in the local contexts. For others, ToC was seen more narrowly as extending the concept of a simple linear causal chain or logical framework. Others understood ToC primarily as a compliance, communication, and high-level planning tool that needed to speak elegantly and coherently to a simple vision of their hub's hoped-for impact, at the risk of communicating incoherence to funders and donors. For others, the ToC was simply a tool upon which to scaffold the logical framework, which itself was embraced with varying degrees of enthusiasm by hub managers. Almost all the hubs experienced tensions and frustrations around managing expectations around how these multiple roles for ToC would be reconciled given limited resources and donor reporting requirements. With these nuances in mind, we have several recommendations and reflections to strengthen the use of ToC in complex research for development programmes moving forward.

Resource Appropriately

The GCRF interdisciplinary research hubs were not typical complex development programmes. They were driven by researchers, who come with unique views around how impact should be conceptualised, and who may not have much familiarity with the language of MEL. Our research indicated that in this environment, complexity sensitive ToC processes cannot be realistically sustained over time without intensely active brokerage by MEL leads. Complexity-aware approaches to evaluation can be intimidating to even confident evaluators (Hargreaves and Podems 2012), and yet many hubs' MEL functions were delegated to the hub manager—some of whom had only limited experience in MEL and limited time to dedicate, considering all the other aspects of their role. The resources, experience and expertise required to drive

these processes were almost uniformly underestimated by the funder, and the hub leads.

Provide Guidelines for Visual Depiction

While it is understandable that GCRF did not wish to be overly prescriptive in their issuing of MEL guidelines, it is our opinion, based on the data that emerged in the interviews, that much time and frustration could have been saved by providing hubs with a clearer set of options for visual mapping of ToCs along with a realistic sense of the timelines and resources for each option. We discuss the benefits, limitations, and potential applications of these options in the sections which follow.

Systems-Based Visual Modelling

Systems-based visual mapping approaches have been very useful in guiding MEL in complex system environments, yet this approach to visual mapping was only attempted in one hub. For this hub, the approach was certainly useful in helping researchers think through their work and how it contributes to impact in a participatory way. However, the visual models were resource intensive to maintain and use, and the approach was ultimately complemented by a simpler 'big wins' linear impact mapping approach which was more in keeping with the reporting or deliverable style requirements of the funders.

Ideally systems mapping should be done from the inception phase of a project, whereafter they might be used to build a more 'traditional' linear ToC (Wilkinson et al. 2021). If this is to be realised, the GCRF funder guidelines would need to dramatically extend the inception phase of research for development projects. Perhaps more importantly, the GCRF would need to create meaningful scope for hubs to engage in a potentially messy phase of ToC exploration that provides the basis for generating ideas for change and improvement. Many of our informants indicated how the GCRF funding proposal didn't make it clear enough to the extent to which MEL was going to be an important component of the project programme design and management cycle, and the GCRF structures were certainly not supportive enough to allow for this process.

Nested ToCs

The concept of nested hierarchies is integral to complex systems theory (Moore et al. 2019), and a few practitioners have extended this concept to visual programme theory (Koleros et al. 2020; Stein and Valters 2012). In theory, nested ToCs help with showing how different parts of a complex intervention fit together to facilitate an effective monitoring and evaluation plan (Mayne and Johnson 2015), and notable case studies exist which document the usefulness of this approach. (e.g. Thornton et al. 2017). Several of the interviewees believed that multiple or nested ToCs were a useful and promising move towards accommodating complexity. However, most interviewees felt this approach could have been realised more fully only with more

resources. In the absence of adequate time and resources, some went so far as to describe the process as a "disaster" and a "waste of time". Thus, as with our observations regarding visual system mapping approaches, we would caution the move towards nested ToCs as an obvious solution to working with ToC in complex programmes unless concerns regarding resources and expertise can be meaningfully addressed.

Simple Impact Models with Linear Elements

From our research, it is apparent that both systems-based visual mapping approaches and nested ToCs will only be possible with dramatic shifts in how projects like GCRF hubs are resourced and managed. A key question is, thus, whether this is feasible? Almost two decades ago, Barnes et al. (2003) concluded that the difficulties of applying a ToC approach in the context of complex systems are *not* solely methodological problems that could be resolved with only sufficient resources and commitment. While 'framing and revisiting' a ToC is central to this process of learning through accountability (Guijt 2010, p. 349), in our own research, we saw how frequently programme managers resorted to simple, linear impact theories that could demonstrate clear pathways to impact and serve as a 'big win' public relations, communication and accountability tool. While it is true that viable concerns can be raised about these kinds of ToCs 'taming' the complexities of the system dynamics—the most important consideration is surely for these models to be useful and achievable given the realities of the funding and time constraints these projects frequently face. While more complex systems-based and nested models certainly have their role to play, when resources and time are constrained, our informants suggested that more simple models that might necessarily impose some linearity and visual simplicity still offer considerable benefits—which brings us to our final recommendation, below.

Take Advantage of the Benefits

The framework developed by Barr et al (2019) for the evaluation of the GCRF envision a two-domain approach to MEL, where the ToC is divided into two portions: complicated aspects (in the sphere of direct influence) and complex aspects (in the sphere of indirect influence) (Barr et al. 2019). Contrary to this, our research showed that hubs did not find linear ToC models merely useful for those aspects of the programme in their direct sphere of influence. Even in the complex system 'domain', simple, often linear ToCs were used as high-level navigation tools as opposed to a detailed planning tool. Indeed, having a broad idea of where the hub was going (as opposed to how they are getting there) was frequently in tune with the management reality of complex programmes (Ramalingam et al. 2014). In our research, these were the ToCs that were also most closely aligned to the project logical framework, which were in keeping with funder expectations and requirements, and which many programme managers also found useful. We should also bear in mind that the evaluation of nonlinear processes involves multiple stakeholders and relations and often

demands multiple ToCs (Barnes et al. 2003). Having a simple ToC to work with (and to revise) can greatly simplify this process where learning and quick adaptation to new insights are to be prioritised.

Conclusion

Simplicity has been framed as a characteristic of weaker ToCs, for example 'uncaptured or loosely connected' impact pathways, absence of feedback loops and 'aesthetics over substance'—issues extensively discussed in Davies (2018). Perhaps however, this is not in and of itself problematic if those diagrams are fit for purpose, while at the same time allowed to be flexibly redefined. As Stein and Valters (2012) have contended, there is a continuum of categories of ToC purpose (strategic planning, monitoring and evaluation, description, and learning). Where the ToC is applied it should be with due consideration of stakeholders' ability to meaningfully engage in the ToC process given their technical capabilities, inclination, and the political and bureaucratic forces both within and beyond their programmes (Davies 2018). Like Barnes et al., we do feel that "there are points of complexity that stretch the application of ToC to a point at which it becomes both methodologically and theoretically suspect" (Barnes et al. 2003, p. 272). In such a context, it might be pragmatic to embrace more simple models, which may include linear elements, as a way of thinking about how to get from point A to point B, or as means to develop 'political literacy' (Davies 2018) rather than a complex and nuanced understanding of how systemic change processes are activated. When used in this way, the potential benefits of ToC can be leveraged meaningfully to create spaces for positive engagements around simple visual tools.

Data Availability The data used in this article are not publically available.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Ackermann, F., and C. Eden. 2020. Strategic Options Development and Analysis (SODA). In *Systems approaches to making change: A practical guide*, ed. M. Reynolds and S. Holwell, 139–200. London: Springer.

Barnes, M., E. Matka, and H. Sullivan. 2003. Evidence, understanding and complexity: Evaluation in non-linear systems. *Evaluation* 9 (3): 265–284.

Barr, J., B. Bryan, P. Kolarz, X. Potau, M. Punton, P. Simmonds, and I. Vogel. 2019. *GCRF Evaluation Foundation Stage—Final Report*. London: Global Challenges Research Fund.

Braun, V., and V. Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3 (2): 77–101.

Checkland, P., and J. Poulter. 2020. Soft systems methodology. In *Systems approaches to making change: A practical guide*, ed. M. Reynolds and S. Holwell, 201–253. London: Springer.

Chen, H.T. 2005. *Practical programme evaluation: Assessing and improving planning, implementation and effectiveness*. Thousand Oaks: SAGE.

Davies, R. 2018. Representing theories of change: Technical challenges with evaluation consequences. *Journal of Development Effectiveness* 10 (4): 438–461.

Deutsch, L., B. Belcher, R. Claus, and S. Hoffmann. 2021. Leading inter- and transdisciplinary research: Lessons from applying theories of change to a strategic research programme. *Environmental Science and Policy* 120: 29–41.

Douthwaite, B., T. Kuby, E. van de Fliert, and S. Schulz. 2003. Impact pathway evaluation: An approach for achieving and attributing impact in complex systems. *Agricultural Systems* 78 (2): 243–265.

DuBow, W.M., and E. Litzler. 2019. The development and use of a theory of change to align programs and evaluation in a complex, national initiative. *American Journal of Evaluation* 40 (2): 231–248.

Dyehouse, M., D. Bennett, J. Harbor, A. Childress, and M. Dark. 2009. A comparison of linear and systems thinking approaches for programme evaluation illustrated using the Indiana Interdisciplinary GK-12. *Evaluation and Programme Planning* 32 (3): 187–196.

Funnell, S., and P. Rogers. 2011. *Purposeful programme theory: Effective use of theories of change and logic models*. San Francisco: Jossey-Bass.

Gates, E.F. 2017. Learning from seasoned evaluators: Implications of systems approaches for evaluation practice. *Evaluation* 23 (2): 152–171.

Guijt, I. 2010. Accountability and learning: Exploding the myth of incompatibility between accountability and learning. In *NGO management: The Earthscan companion*, ed. A. Fowler and C. Malunga, 339–352. Milton: Taylor & Francis Group.

Hargreaves, M.B., and D. Podems. 2012. Advancing systems thinking in evaluation: A review of four publications. *American Journal of Evaluation* 33 (3): 462–470.

Hawe, P., A. Shiell, and T. Riley. 2009. Theorising interventions as events in systems. *American Journal of Community Psychology* 43 (3): 267–276.

Koleros, A.S., M. Oldenbeuving. Mulkerne, and D. Stein. 2020. The actor-based change framework: A pragmatic approach to developing programme theory for interventions in complex systems. *American Journal of Evaluation* 41 (1): 34–53.

Mayne, J., and N. Johnson. 2015. Using theories of change in the CGIAR Research Programme on Agriculture for Nutrition and Health. *Evaluation* 21 (4): 407–428.

Moore, M., B. Parsons, and P. Jessup. 2019. How evaluators can use a complex systems lens to get 'untrapped' from limiting beliefs and assumptions. *American Journal of Evaluation* 40 (1): 75–91.

Morecroft, J. 2020. System dynamics. In *Systems approaches to making change: A practical guide*, ed. M. Reynolds and S. Holwell, 25–88. London: Springer.

Patton, M. 2014. *Qualitative research and evaluation methods: Integrating theory and practice*, 4th ed. Thousand Oaks: SAGE Publications, Inc.

Patton, M.Q. 2015. Book review: Evaluating the complex: Attribution, contribution, and beyond. *American Journal of Evaluation* 36 (3): 419–429.

Ramalingam, B., M. Laric, and J. Primrose. 2014 *From best practice to best fit: Understanding and navigating wicked problems in international development.* Overseas Development Institute Working Paper. London: Overseas Development Institution.

Rogers, P. 2000. Programme theory: Not whether programs work but how they work. In *Evaluation models*, ed. D. Stufflebeam, G. Madaus, and T. Kellaghan, 209–232. Dordrecht: Springer.

Rogers, P.J. 2008. Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation* 14 (1): 29–48.

Shiell, A., P. Hawe, and L. Gold. 2008. Complex interventions or complex systems? Implications for health economic evaluation. *BMJ* 336 (7656): 1281–1283.

Stein, D., and C. Valters. 2012. *Understanding theory of change in international development.* Justice and Security Research Programme Paper 1. London: London School of Economics and Political Science.

Thornton, P.K., T. Schuetz, W. Förch, L. Cramer, D. Abreu, S. Vermeulen, and B.M. Campbell. 2017. Responding to global change: A theory of change approach to making agricultural research for development outcome-based. *Agricultural Systems* 152: 145–153.

UKRI. 2021. *Global Challenges Research Fund.* UKRI, Accessed December 16, 2022. https://www.ukri.org/our-work/collaborating-internationally/global-challenges-research-fund/.

Walton, M. 2014. Applying complexity theory: A review to inform evaluation design. *Evaluation and Programme Planning* 45: 119–126.

Wilkinson, H., D. Hills, A. Penn, and P. Barbrook-Johnson. 2021. Building a system-based theory of change using participatory systems mapping. *Evaluation* 27 (1): 80–101.

Authors and Affiliations

Sarah Chapman[1] · Adiilah Boodhoo[1] · Carren Duffy[1] · Suki Goodman[1] · Maria Michalopoulou[2]

Adiilah Boodhoo
Adiilah.Boodhoo@uct.ac.za

Carren Duffy
Carren.Duffy@uct.ac.za

Suki Goodman
Suki.Goodman@uct.ac.za

Maria Michalopoulou
maria.michalopoulou@spi.ox.ac.uk

[1] School of Management Studies, University of Cape Town, Leslie Commerce Building, Upper Campus, Rondebosch, Cape Town, Western Cape, South Africa

[2] The UKRI GCRF Accelerating Achievement for Africa's Adolescents Hub, Centre for Evidence-Based Intervention, Department of Social Policy & Intervention, University of Oxford, Oxford, UK