Data Analysis and Synthesis Within a Realist Evaluation: Toward More Transparent Methodological Approaches

Brynne Gilmore¹,², Eilish McAuliffe³, Jessica Power¹,²,⁴, and Frédérique Vallières¹

Abstract
Realist evaluations are increasingly used in the study of complex health interventions. The methodological procedures applied within realist evaluations however are often inexplicit, prompting scholars to call for increased transparency and more detailed description within realist studies. This publication details the data analysis and synthesis process used within two realist evaluation studies of community health interventions taking place across Uganda, Tanzania, and Kenya. Using data from several case studies across all three countries and the data analysis software NVivo, we describe in detail how data were analyzed and subsequently synthesized to refine middle-range theories. We conclude by discussing the strengths and weaknesses of the approach taken, providing novel methodological recommendations. The aim of providing this detailed descriptive account of the analysis and synthesis in these two studies is to promote transparency and contribute to the advancement of realist evaluation methodologies.

Keywords
realist evaluation, realist data analysis, data synthesis, NVivo, research methods

Introduction

Realist Evaluations
Recognizing the importance of understanding how contextual factors influence health interventions, realist evaluations are increasingly used by academics, practitioners, and policy makers within the health field (Byng, 2005; Goicolea, Coe, Hurtig, & San Sebastian, 2012; Maluka et al., 2011; Marchal, Dedzo, & Kegels, 2010; Prashanth, Marchal, Devadasan, Kegels, & Criel, 2014; Ranmuthugala et al., 2011; Rycroft-Malone et al., 2016; Rycroft-Malone, Fontena, Bick, & Seers, 2010; Tolson, McIntosh, Loftus, & Cormie, 2007; Van Belle, Marchal, Dubourg, & Kegels, 2010). A form of theory-driven evaluation, realist evaluations acknowledge that interventions and their outcomes are subject to contextual influence. As such, a realist evaluator’s duty is to understand “how, why, for whom, and under which conditions” interventions work (Pawson & Tilley, 1997). Incorporating theory throughout the study, realist evaluations aim to be pragmatic by producing policy-relevant findings at a level of abstraction that can be transferred across settings (Fletcher et al., 2016; Salter & Kothari, 2014). To do so, the context-mechanism-outcome configuration (CMOC) is central to analysis and the theory building/refining process for realist studies (Power et al., 2019). Developed as an analytical tool by Pawson and Tilley (1997), CMOCs describe how specific contextual factors (C) work to trigger particular mechanisms (M), and how this combination generates various outcomes (O) (Pawson & Tilley, 1997). By exploring these mechanisms of change, realist evaluations aim to understand how a programme works or is expected to work within specific contexts, and what conditions may hinder or promote successful outcomes (Jagosh et al., 2011; Pawson, 2006). Realist evaluations therefore seek to explain generative causation within the social world by identifying particular patterns of interactions.

A requirement for realist evaluation research is that data analysis takes a “retroductive” approach. An advancement on the more common reasoning techniques of induction or deduction, retroduction refers to “the identification of hidden causal forces that lie behind identified patterns or changes in those patterns” (The RAMESES II Project, 2017, p. 1). Retroduction therefore uses both inductive and deductive reasoning and includes researcher insights to understand generative causation, by exploring the underlying social and psychological drivers identified as

¹ Trinity Centre for Global Health, School of Psychology, Trinity College Dublin, Dublin, Ireland
² Trinity Centre for Global Health, School of Medicine, Trinity College Dublin, Dublin, Ireland
³ School of Nursing, Midwifery and Health Systems, College of Health and Agricultural Sciences, University College Dublin, Dublin, Ireland
⁴ Health Research Board Trials Methodology Research Network, Ireland

Corresponding Author:
Brynne Gilmore, Trinity Centre for Global Health, Trinity College Dublin, 2 College Green, Dublin, Ireland.
Email: gilmorb@tcd.ie
influencing programme outcomes. For retroduction to occur, it is important to have multiple data sources and incorporate one’s common sense to test and refine programme theories (PTs; The RAMESES II Project, 2017, p. 1).

To obtain these multiple data sources, a number of methodological steps are common across realist evaluations. First, realist evaluations typically begin through identifying initial programme theories (IPTs) that hypothesize how, why, and for whom the intervention may work, based on literature and document reviews, and key informant interviews with programme architects or implementers (Mukumbang, Van Belle, Marchal, & van Wyk, 2016b; Pawson & Tilley, 1997, 2004). These IPTs are subsequently used to inform the next steps of the realist evaluation, the field study, and its design, with the aim that contextually relevant data can further inform and refine the IPTs.

Although recognized as being methods neutral, in that a variety of methods and tools can be used that best refine theories, realist evaluations most commonly employ the recommended mixed-method case study approach (Pawson & Tilley, 1997; Westhorp, 2008). During qualitative data collection within the evaluation (i.e., focus group discussions, in-depth or semistructured interviews, and key informant interviews), it is important that evaluators also employ the realist interview technique, a collaborative form of theory refinement in which the interview is guided by the theories you are aiming to refine (Manzano, 2016; Pawson & Sridharan, 2010; Pawson & Tilley, 1997). For further reference on the realist interview and realist data collection, Manzano’s (2016) paper provides an overview of interviewing within a realist evaluation. During the data collection and analysis stages, the case studies are used to refine or further generate CMOCs (Marchal, Van Belle, van Olmen, Hoereee, & Kegels, 2012). The aforementioned realist interview techniques are considered an especially important tool to uncover and refine CMOs and PTs (Manzano, 2016).

Aligned to retroductive approaches, CMOCs are identified and coded within the data (Jagosh et al., 2012). This process is often presented as “equations” within a table to show individual components and their relationships (Priest, 2006). Semipredictable patterns of CMOCs, or semi-rules, can be identified and collated, which show CMOCs with more evidence support. During (or following) the analysis, the CMOC findings are then synthesized back into the IPT and/or PTs. This step in a realist evaluation therefore involves using patterns of generative causation to support and justify theory refinement. Finally, the above steps can be repeated multiple times in a form of continued theory refinement and specification. Over time, and across multiple iterations, the ultimate goal of realist evaluations is to produce a refined middle-range theory (MRT) of how a programme works by identifying common patterns within reality.

The MRT, defined by Merton (1968) as the “theory that lies between the minor but necessary working hypotheses . . . and the all-inclusive systematic efforts to develop a unified theory that will explain all the observed uniformities of social behavior, social organization and social change” (p. 39), is a result of programme specification. The MRT is therefore a refinement of the IPT which has been tested through a series of case studies, culminating to a contextually specific and relevant policy theory that describes the explanatory pathway of how change occurs.

Analysis Within Realist Evaluation

Despite the increasing number of published realist evaluations within the field of health, few papers provide sufficient detail on the methodological processes used. Consequently, and as highlighted by several authors (Byng, 2005; Marchal et al., 2012; Salter & Kothari, 2014; Westhorp, 2008), there is little guidance on specific analysis approaches and steps to use within a realist evaluation. While most realist studies would employ CMOCs as their primary analytical tool, the techniques for how one identifies “CMOCs,” and how these are subsequently incorporated into theory refinement, varies within the extant literature. While some propose analytical induction (Byng, 2005) or thematic analysis (Moore, Moore, & Murphy, 2012), others have conducted specific analysis techniques including “realist qualitative analysis” and the study of “enabling, disabling, and generating mechanisms” (Kazi, 2003; Westhorp, 2008). Consistent with other realist evaluation studies, however, these analysis techniques often lack details on the actual process/steps used.

Paper Objective

The lack of methodological guidance within realist evaluations has prompted the Realist And Meta-narrative Evidence Synthesis: Evolving Standards (RAMESES) project (Wong, Greenhalgh, Westhorp, Buckingham, & Pawson, 2013; Wong et al., 2016), currently the leading methodological group of realist syntheses and evaluations, to call for “review authors to provide detail on what they have done and how—in particular with respect to the analytic processes used” (pp. 13, 29). Our article responds to this call by documenting the analytical process used as part of realist evaluations of two different community health interventions taking place across three countries. By detailing our analytical process, we aim to contribute to the methodological advancement of realist evaluation by providing guidance on CMOC elicitation and theory refinement. Specifically, this article puts forward more transparent methods for data analysis and synthesis within a realist evaluation by explaining (i) how CMOCs were extracted and elicited, (ii) how CMOCs were collated, (iii) how collated CMOCs were synthesized to develop and/or refine PTs, and finally, (iv) how findings were synthesized across case studies to arrive at MRTs. Figure 1 provides an overview of the analysis phases used across our realist evaluations and will be referred to throughout the remainder of the article, highlighting Phases 3–5 (more details on Phases 1 and 2 can be found here; Gilmore et al., 2016).

Study Background

The analytical process presented within this article has been developed throughout two realist evaluations. While the findings are not presented here, a description of the studies is introduced to provide greater context for the methodological discussion within. Both realist evaluations took place within community health
interventions being implemented by nongovernmental organizations operating within sub-Saharan Africa. The first realist evaluation sought to understand “how, why, and for whom community health committees (CHCs) contribute to community capacity building” within North Rukiga, Uganda, and Mundembo, Tanzania. These CHCs were implemented as a standard component of a Maternal, Newborn, and Child Health intervention led by World Vision Ireland. A total of six case studies, three case studies within each country, each following a different CHC group, were conducted using the recommended mixed-methods approach within the first study in Uganda and Tanzania. Further details on the study are described elsewhere (Gilmore et al., 2016). The second realist evaluation study explore: How, why, and for whom a community engagement intervention (titled Community Conversations [CCs]) influences behavior change communication for health within Marsabit, Kenya. This second evaluation involves conducting four case studies, each one involving one CC group, from initiation to completion of a behavior change intervention being implemented by Concern Worldwide Kenya. Thus, this article has been developed through a total of 10 case studies, across three countries, from two realist evaluation studies. The first study was completed in mid-2017, and the second study is currently ongoing, having completed two of three rounds of data collection and theory refinement via analysis.

In both realist evaluations, data collection occurred at the case level, with the exception of several Key Informants Interviews (KIIIs) whose knowledge was relevant to the type of programmes, or intervention, being implemented across all sites. Participants were principal stakeholder groups, each with varying degrees of involvement with the intervention including community members (men and women), community health workers/volunteers, CHC or CC members, community KIIIs (health staff, local leaders), and more strategic-level KIIIs (programme staff and Ministry of Health). The majority of the data collection consisted of qualitative interviews and focus group discussions, all incorporating the realist interview technique. In both evaluations, and in line with mixed-method approaches, quantitative surveys were conducted to better understand programme outcomes. Demographics were taken for all participants. Programme documentation review, meeting minutes, available health facility data, and observations/field notes were also used as data sources. For logistical reasons, case study data were collected concurrently rather than sequentially. In the first evaluation, a theory iteration phase was conducted approximately 4 months after data collection was completed. This consisted of presenting preliminary PTs to participant groups for continued refinement (or refuting) and clarification on any questions arising within the data analysis. The second study includes three iterative rounds of primary data collection.

**Method**

All case studies were first analyzed separately, in order to refine PTs specific to each case. Consistent with realist evaluation procedures, the resulting refined PTs were synthesized across the cases to identify MRTs for each research question. Given the lack of methodological guidance, however, a transparent and rigorous process for analysis within the case studies was first developed, based on previous work from the IPT elicitation. Specifically, a blog post written by Dalkin and Foster (2015) on using NVivo software for realist analysis served as the
foundation for this process, which was then expanded and adapted. Figure 2 presents an overview scheme of the data analysis process described below.

**Data Sourcing**

Consistent with realist approaches, data analysis was retroductive in that it applied both inductive and deductive logic to multiple data sources, while also incorporating the lead author’s own understanding to uncover generative causation (The RAMESES II Project, 2017). The different types of data, however, played different roles within this process. Specifically, the following multiple data sources were used throughout the analysis to identify the generative mechanisms: interview and focus group transcripts, surveys related to intervention outcomes, intervention documentation and meeting minutes from groups, programme and affiliated health facility monitoring data and records, and observations and field notes. Qualitative interviews with participants were featured most heavily and were the primary source for CMOC coding. They were thus the primary source for theory testing and refinement, as they were the only data source to contain any extractable CMOCs in their entirety. Observations as field notes were used to assist in the decision-making processes of refinement resulting from the qualitative interview sources. The surveys were used to triangulate and inform the testing and revision of the theory and were helpful in understanding components of specific PTs and for guiding the conceptual theory refinement. Other documentation including the capacity survey, meeting minutes, and reports were used to triangulate findings arising from the aforementioned sources and to identify relevant outcomes related to the intervention. Finally, KIs were predominantly used during the synthesis and MRT development phase.

**Phase 3, Step 1: Data Preparation**

Prior to starting the coding, all transcripts, additional data sources, and field notes were read to gain a better contextual understanding of the interview. Quantitative data were also analyzed in Excel (Version 14.7.1). Once completed, quantitative results were imported into NVivo for Mac (Version 11.4.0) to be coded alongside qualitative data. Once all data were imported into NVivo, the analysis of the case studies consisted of completing the following steps, undertaken separately for each case (see Box 1 for description of terms):

- Each piece of data was stored as an individual source (i.e., one interview or focus group transcript was an individual source).
- A “node” was created for each IPT.
- Child nodes were created if the node/PT underwent revision. Any coding to the new revised PT would thus occur under the new child node. This process occurred...
any time there was substantial revision and allowed for a clear tracking of theory refinement.

- A memo was linked to each node (and any new child nodes) to record decision-making processes and rationales for refinement of theories. This is where the majority of CMOC elicitation occurred.
- All case studies also included a node for further context/outcomes that arose to ensure that any contextual information participants were discussing, even if not directly linked CMOCs were captured.

Phase 3, Step 2: CMOC Extraction and Elicitation

Coding occurred when an observable “context–mechanism–outcome” was found in the data. Once a CMO was found, it was coded to an appropriate node—that is, CMOs were linked to the relevant IPT/PTs. Once a whole data source was thoroughly read and coded, all the nodes that had new coded data were reviewed. Coded material within specific nodes was subjected to an in-depth exploration, one code (CMO) at a time. This included rereviewing the data source to further understand the codes’ context and relevance and then adding it to the memo that was linked to the IPT/PT (i.e., the same CMO was found within the node and then the memo that was specific to that node). From here, the code (CMO) added to the appropriate memo allowed for a more in-depth exploration. Within the memo, each new CMO was subjected to review using the template depicted in Box 2, whereby the contents of the memo were placed under the following headings: context, mechanism, outcome, potential CMOC, supports/refutes/refines, how/why/decision-making processes, supports/refutes/refines, how/why/decision-making processes, supports/refutes/refines, how/why/decision-making processes, supports/refutes/refines, how/why/decision-making processes.

Box 1. Terms in NVivo. Adapted from QSR International, 2019.

Sources refer to the research material. Within this study, each source contained a separate interview or focus group discussion. Minute meetings, documents, photos, and field notes were also stored as separate sources.

Nodes allow one to gather relevant data under a common theme. These may also be understood as codes. When a data source contains an important piece of information, researchers code this information to a relevant existing node or create a new node depending on the data. After which, these data segment are stored within the node along with any other data coded to that node.

Child nodes arise when a new node is created within an existing node, creating a hierarchy. The creation of child nodes within this study was particularly important for transparency of refinement throughout the analysis as the adult node can contain all aggregates from child nodes.

Memos are blank documents that can be linked to data and nodes. This allows users to document processes/ideas/thoughts on the analysis.

Box 2. Example of Memo for CMOC.

Memo for Initial Programme Theory 1

IPT - Individuals within the CHC are likely to provide supportive and consistent engagement for activities if they have strong motivation, a desire for volunteering for their community, and are committed to the group and its objectives. This may be influenced by the individual members’ specific attributes (such as availability of time and knowledge), previous experience and incentives, training, and supervision provided to them. These factors together result in increased collaboration within the committee, increased respect by community members, and an overall committed committee.

Code: “The programmes and the members continue because they see that there are some changes because of what they do, for example women now come to the hospital for delivery.”

Source: FGD women.

C: Visible positive changes to health/behavior from women and children.

M: Positive reinforcement (through individual and/or community).

M2: Motivation and work satisfaction.

O: Sustainability and commitment.

CMOC: When CHC members are able to recognize positive contributions they made to their communities, it acts as positive reinforcement which encourages motivation, continued work and commitment to their duties.

Support/Refute/Refine: Refine.

(Continued)
Box 2. (Continued)

How/Decision-Making: Places more emphasis on the individual members’ internal motivation for working, highlighting the need to see/be informed of etc. (get feedback on) their work and how it is impacting on the community. Several potential mechanisms for this. Interesting that this is from a community member, so it is the service users that recognize their positive work, which could also be an important mechanism and/or context.

Links/Ripple Effects: Several CHC interviews highlight this specific CMOC. Potential link to IPT 2 and as a group.

Notes: Linked to motivation of members

Other Codes: Links to IPT2—CHCs place within communities

Note. FDG = focus groups discussion; CHC = Community Health Committees; CMOC = context–mechanism–outcome configuration; IPT = initial programme theories.

links to other IPTs, and additional notes. This process was done for each new CMO, across each of the memos (that are linked to the nodes). This memo therefore served as a tool for analysis, allowing for greater transparency for how the CMOCs were generated.

In the next stage of analysis, any CMOs resembling already refined CMOCs were combined, and if appropriate, further refinement occurred. For example, if the CMOC “1.2” was coded and refined from an interview with CHC1, and the same (or similar) CMOC was present during an interview with CHC4, the “coded reference” from CHC4 was added to the CMOC from CHC1. This resulted in only one CMOC (1.2) with data derived from two sources (CHC1 and CHC4). The consolidated CMOCs were then used to support, refute, or refine the IPTs or PTs, depending on the stage of analysis. Figure 3 provides an example of the process described in this section.

Phase 4, Step 1: Using CMOCs to Refine PTs

The IPT refinement process occurred continuously throughout the data analysis, so that multiple PTs were simultaneously refined during the analysis. Refinement occurred where there was sufficient CMOCs support for refinement within each PT. Where the refinement of an IPT/PT occurred, a new child node was created for that refined theory in order to document the refinement processes throughout the analysis. For instance, if two extracted CMOCs worked to refine IPT 1 after one interview (source) was analyzed, the node for this (IPT 1) was split into a new child node, depending on the number of resulting PTs after refinement. For example, as illustrated in Figure 4, IPT 1 was refined into three PTs after two data sources were reviewed, resulting in child nodes for PT 1.1, PT 1.2, and PT 1.3. Thus, any new relevant CMOCs elicited from the subsequent sources were coded directly to the most relevant child nodes.

Phase 4, Step 2: Collating Evidence and Refinement Verification

Continuing with the example from Case Study 1 in Uganda from the first evaluation, a total of 23 CMOCs (from 191 codes) were generated across all the data sources, excluding data from the strategic level KIIIs. As noted, refinement was ongoing throughout the analysis. However, an important step was to ensure this process was reflective of all the data by collating all CMOCs to verify theory refinement. Once coding was complete, all CMOCs and their supporting evidence were collated into tables; organized under their resulting PT, they contributed to refining. That is, if resulting PT 1.1 had six CMOCs that contributed to its development/refinement, the first row in the table would be IPT 1 and the last row 1.1. All rows in between would be comprised of the CMOCs and their evidence, to show the progression between IPT 1 and PT 1.1. Each CMOC relevant to that IPT/PT listed in the table had a strong supporting quote, additional notes on the thought process, and a list of all the sources of evidence that supported the CMOC’s development. This is where other data sources’ support for the CMOC was able to be consolidated (see Table 1, for an example of how this was documented). In this way, multiple CMOCs worked to refine the majority of PTs.

Once this process was complete, all the resulting PTs were then also collated and examined. In some instances, further refinement occurred due to similarities and overlaps or one PTs’ ability to provide explanatory power to another. This process became more of a “tidying up” of the resulting PTs, for more clarity and better linkages after a long refinement process. Figure 4 shows the refinement process for Case Study 1, from IPTs to the resulting PTs, demonstrating the multiple revisions from IPTs to finalized PTs for the case study. From the figure three IPTs can be seen, which after analysis was complete resulted in 10 refined PTs. Emerging thin lines denote a new PT (a derivative of an earlier theory, which was separated to be more specific), and double slash denotes when a theory underwent refinement. In this example, we can see how IPT 3 was refined into PT 3.3 through separation after source “men” and undergoing refinement two subsequent times (after CHC and after data review and triangulation).

Phase 5: Synthesis Across Case Studies for MRT

As noted, and for logistical reasons, the case studies conducted as part of these evaluations were not run consecutively. As such, the process of refining PTs through the iterative cycles between each case study, as described above, was instead simultaneously conducted (refer to Figure 2). This section reports only on actions done within the first study from Uganda and Tanzania. Based on the previous phases’ results, three methods were used to assist in the synthesis of the PTs with
the aim of identifying MRTs, with each method building from the product of the last. First, a search for demi-regularities (semipredictable patterns occurring in the data) across the case studies was conducted. Second, KIIs with strategic-level stakeholders providing an “overall” view of the CHCs were analyzed and incorporated into the demi-regularities. Lastly, the refined theories were reviewed in relation to existing relevant literature.

This process was done manually, instead of continuing the synthesis within NVivo. Initially, efforts were made to conduct...
Individuals within CHCs are likely to be committed and have continued an active engagement with responsibilities if they are motivated by refined programme theory 1.1:

Refined programme theory 1.1:

I liked it because people’s lives where changed for example I had seen people’s lives were getting in danger but later when we got the health workers, supporting quote:

Supporting quote:

thought processes:

Thought processes:

Community members recognize the importance of providing incentives for CHC members so that they are able to continue their work. With this interview (quoted) also called for better/more/sustained incentives. Difficult to decide how to code, but ended up being more about the community perceptions and their want for CHCs to have recognition and motivation. Considered C and Ms around “support motivates CHCs,” but this was not a CHC speaking, so decided to consider it more from an “appreciation/respect/acknowledgement of work” angle. Interviews were quite positive about CHCs, so had in mind while reviewing that the CHCs were well received and their work was supported.

Supporting quote:

I liked it because people’s lives where changed for example I had seen people’s lives were getting in danger but later when we got the health workers, they were able to attend to all the patients and now people’s lives and health has been improved. So I loved it so much. (CHC 4).

Refined programme theory 1.1:

Individuals within CHCs are likely to be committed and have continued an active engagement with responsibilities if they are motivated by both intrinsic and extrinsic factors. These may consist of financial compensation and/or reimbursement (extrinsic) and positive reinforcement (intrinsic) such as feeling respected, which can increase pride in ones’ work. Community recognition of their services is a likely condition in order for this to occur.

Note. CMOC = context–mechanism–outcome configuration; CHC = Community Health Committees.

the synthesis in NVivo, but the need for more tactile engagement with the data and the various steps and tools used were better suited to do by hand. It is feasible to do this within NVivo however, and others are encouraged to find the system that best suits them.

Searching for demi-regularities occurring across the case study findings involved several steps (see Supplementary File 1, for an example):

- Findings from each case study (including PTs and their supporting CMOCs) were separated on different colored paper to ease in recognition of case study.
- PTs and CMOCs from all cases were combined.
- Commonalities within the combined PTs and CMOCs were searched for and grouped onto a single piece of paper.
- Demi-regularities within the grouped PTs/CMOCs were highlighted.
- When PT demi-regularities were identified, all CMOCs were reviewed to see whether any additional elicited CMOCs offered explanatory information.

Once this was complete, the documents were reviewed and resulting demi-regularities were synthesized to inform the proposed MRTs. These were then examined in light of KIIs that were not specific to one case study but that were conducted with individuals with knowledge on the overall workings of the intervention. As such, their interviews were focused on operations at this level. This was done to support and refine the current theories but not to refute those that had arisen from the case studies. Data were also used to help manage any discrepancies between the case study findings and were specifically useful in understanding generative causality between the different intervention levels, as the KIIs could offer a more holistic “systems thinking” perspective on the intervention. This process also worked to conjecture theories at the “societal” level, as the KIIs had knowledge of functioning within this socioecological domain. The KIIs were analyzed using NVivo in a similar process to that described within the analysis section for Phases 3 and 4.

These generated MRTs were then compared with existing literature, namely any relevant formal theories. The aim was to identify any such theories that report on related causal chains or moderating factors as a type of “plausibility check” (Marchal et al., 2010). This also worked to expand the explanatory mechanisms and situate PTs within other existing formal theories. The completion of this process led to MRTs for the research question being put forward and the commencement of this research.

Conclusion

The importance of documenting methodological processes within realist evaluation is especially apparent in light of the lack of guidance currently available to guide the analysis within realist evaluation. In response to the demand for more transparent procedures within realist methodologies, the above
outlines a number of concrete, practical methods that can be used to support the identification and extraction of CMOCs, the refinement and generation of PTs, and the synthesis of findings across multiple case studies as key components of realist evaluations. In doing so, the current article makes an important contribution to advancing the methodological approaches that can be applied to this emerging methodology. To the best of our knowledge, this is the first article that attempts to document this process for the specific use of realist evaluations.

Adhering to the realist principles of retroduction, using case study approaches, and the importance of iteration within theory refinement, the above methods further contribute ideas for more transparent procedures that can be used throughout the analysis and synthesis process. The analytical and synthesis approaches proposed above also adhere to the RAMSES quality standards (Greenhalgh et al., 2016). Notably, the quality guidelines focus on eight main components: the evaluation purpose with realist methodology being appropriate, principle of generative causation is applied throughout the study, there is construction and refinement of PTs, the evaluation design is justified and appropriate ethical guidance is followed, data collection methods are suitable for the evaluation, the participants are appropriate to provide sufficient data, the data analysis approach is retровective and all consistent with principles of generative causation to develop and refine theory, and lastly that reporting is clear and in line with realist assumptions (Greenhalgh et al., 2016).

It is worth noting, however, that the proposed approaches are not without their limitations. Notably, the procedures outlined above are time-consuming, largely owing to the numerous layers of documentation and explanations/justifications for decision-making. Early stages of the analysis are particularly time-consuming, as new CMOCs elicitation and theory refinement is at a high. Furthermore, this process, as is the case within all realist research, assumes that the research team has some innate knowledge on the functioning and contextual conditions of the intervention of the study. As elicitation of CMOCs and the refinement of theories are often dependent on the researchers’ judgment and existing knowledge, documenting detailed “decision-making” processes within the analysis, albeit time consuming, helps to ensure transparency and rigor across this component of a realist evaluation. Future work could also consider identifying appropriate processes for synthesizing PTs into MRTs using NVivo software, something which was not done within this presented work.

Overall, it is recommended that realist research expands analysis documentation by moving beyond thematic analysis guided by the CMOC analytical framework as a sole analytical and synthesis approach. Like others before, we echo the need for clear and published documentation within realist methods to contribute to its methodological advancement and its perceived rigor within the scientific community. This article provides support for realist evaluations using rigorous and transparent methodological processes, where others are able to follow the data and logical flow of retroduction and subsequent theory refinement. Applying similar approaches to upcoming realist research can support replication and study verification, while continuing to grow this important methodology, allowing for more usable and purposeful research moving forward.

Authors’ Note
Jessica Power is now affiliated with Health Research Board Trials Methodology Research Network, Ireland.

Acknowledgments
The authors would like to thank World Vision Ireland, World Vision Uganda, World Vision Tanzania, and Irish Aid for the support of the first study. They would also like to acknowledge Concern Worldwide and Concern Worldwide Kenya, and the Irish Research Council, Marie Skłodowska-Curie Actions and the European Union, for the support of the second study. The authors would also like to acknowledge all the participants and programme staff who supported this work.

Author Contributions
B.G. led both study designs, data collection, analysis, and synthesis. F.V. and E.M. supervised the first study, providing substantial guidance on the methodological process. F.V. contributed to the design of the second study and provided support throughout this research. J.P. assisted in the design of the analysis procedure and provided invaluable contribution and feedback on methodological processes throughout the studies. B.G. drafted and revised the manuscript. F.V., E.M., and J.P. provided valuable contributions to the manuscript draft. All authors reviewed the final manuscript.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: B.G. is currently a postdoctoral fellow, conducting the study within Kenya. This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 713279.

ORCID iD
Brynne Gilmore https://orcid.org/0000-0003-4496-9254

Supplemental Material
Supplemental material for this article is available online.

References
Byng, R. (2005). Using realistic evaluation to evaluate a practice-level intervention to improve primary healthcare for patients with long-term mental illness. *Evaluation, 11*, 69–93. doi:10.1177/1356389005053198
Dalkin, S. M., & Foster, N. (2015). Using NVivo to enhance transparency in a realist evaluation. Retrieved from http://www.qsrinternational.com/blog/how-researchers-use-nvivo-to-enhance-transparency
Fletcher, A., Jamal, F., Moore, G., Evans, R. E., Murphy, S., & Bonell, C. (2016). Realist complex intervention science: Applying realist principles across all phases of the medical research council
framework for developing and evaluating complex interventions. *Evaluation*, 22, 286–303.

Gilmore, B., McAuliffe, E., Larkan, F., Conteh, M., Dunne, N., Gaudreau, M., . . . Vallières, F. (2016). How do community health committees contribute to capacity building for maternal and child health? A realist evaluation protocol. *BMJ Open*, 6. doi:10.1136/bmjopen-2016-011885

Goicoeia, I., Coe, A. B., Hurtig, A. K., & San Sebastian, M. (2012). Mechanisms for achieving adolescent-friendly services in Ecuador: A realist evaluation approach. *Global Health Action*, 5. doi:10.3402/gha.v5i0.18748

Greenhalgh, T., Dawson, W., Wong, G., Westhorp, G., Greenhalgh, J., Manzano, A., & Jagosh, J. (2016). *Quality standards for realist evaluation. For evaluators and peer-reviewers*. Retrieved from http://ramesesproject.org/media/RE_Quality_Standards_for_evaluators_and_peer_reviewers.pdf

Jagosh, J., Macaulay, A. C., Playe, P., Salsberg, J., Bush, P. L., Henderson, J., . . . Herbert, C. P. (2012). Uncovering the benefits of participatory research: Implications of a realist review for health research and practice. *The Milbank Quarterly*, 90, 311–346.

Jagosh, J., Playe, P., Macaulay, A. C., Salsberg, J., Henderson, J., Sirett, E., . . . Greenhalgh, T. (2011). Assessing the outcomes of participatory research: Protocol for identifying, selecting, appraising and synthesizing the literature for realist review. *Implementation Science*, 6, 24.

Kazi, M. A. F. (2003). Realist evaluation for practice. *British Journal of Social Work*, 33, 803–818.

Maluka, S., Kamuzora, P., Sansebastian, M., Byskov, J., Ndawi, B., Olsen, O. E., & Hurtig, A. K. (2011). Implementing accountability for reasonableness framework at district level in Tanzania: A realist evaluation. *Implementation Science*, 6, 11. doi:10.1186/1748-9098-6-11

Manzano, A. (2016). The craft of interviewing in realist evaluation. *Evaluation*, 22, 342–360. doi:10.1177/1358326X16638615

Marchal, B., Dedzo, M., & Kegels, G. (2010). Turning around an ailing district hospital: A realist evaluation of strategic changes at Ho Municipal Hospital (Ghana). *BMC Public Health*, 10, 787. doi:10.1186/1471-2458-10-787

Marchal, B., Van Belle, S., van Olmen, J., Hoeree, T., & Kegels, G. (2012). Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research. *Evaluation*, 18, 192–212. doi:10.1177/1356389012442444

Merton, R. K. (1968). *Social theory and social structure*. New York, NY: Free Press.

Moore, G. F., Moore, L., & Murphy, S. (2012). Integration of motivational interviewing into practice in the national exercise referral scheme in Wales: A mixed methods study. *Behavioural and Cognitive Psychotherapy*, 40, 313–330.

Mukumbang, F. C., Van Belle, S., Marchal, B., & Van Wyk, B. (2016a). Realist evaluation of the antiretroviral treatment adherence club programme in selected primary healthcare facilities in the metropolitan area of Western Cape Province, South Africa: A study protocol. *BMJ Open*, 6, e009977. doi:10.1136/bmjopen-2015-009977

Mukumbang, F. C., Van Belle, S., Marchal, B., & Van Wyk, B. (2016b). Towards developing an initial programme theory: Programme designers and managers assumptions on the antiretroviral treatment adherence club programme in primary health care facilities in the metropolitan area of western cape province, South Africa. *PLoS One*, 11, e0161790.

Pawson, R. (2006). *Evidence-based policy. A realist perspective*. London, England: Sage.

Pawson, R., & Sridharan, S. (2010). Theory-driven evaluation of public health programmes. In A. Killoran & M. P. Kelly (Eds.), *Evidence-based public health: Effectiveness and efficiency* (Vol. 49, pp. 43–112). Oxford, England: Oxford University Press.

Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. London, England: Sage.

Pawson, R., & Tilley, N. (2004). *Realist evaluation*. England, London: Sage. Retrieved from http://www.communitymatters.com.au/gpage1.html.

Power, J., Gilmore, B., Vallières, F., Toomey, E., Mannan, H., & McAuliffe, E. (2019). Adapting health interventions for local fit when scaling-up: A realist review protocol. *BMJ Open*, 9, e022084.

Prashanth, N. S., Marchal, B., Devadasan, N., Kegels, G., & Criel, B. (2014). Advancing the application of systems thinking in health: A realist evaluation of a capacity building programme for district managers in Tumkur, India. *Health Research Policy and Systems*, 12, 42.

Priest, N. (2006). ‘Motor Magic’: Evaluation of a community capacity-building approach to supporting the development of preschool children. *Australian Occupational Therapy Journal*, 53, 220–232.

QSR International (2019). Get started: Glossary of key terms. *NVivo for Mac Help*. Retrieved from http://help-nv11mac.qsrinternational.com/desktop/concepts/Key_terms.htm

Rammuthugala, G., Cunningham, F. C., Plumb, J. J., Long, J., Georgiou, A., Westbrook, J. I., & Braithwaite, J. (2011). A realist evaluation of the role of communities of practice in changing healthcare practice. *Implementation Science*, 6, 49. doi:10.1186/1748-9098-6-49

Rycroft-Malone, J., Burton, C. R., Wilkinson, J., Harvey, G., McCormack, B., Baker, R., . . . Thompson, C. (2016). Collective action for implementation: A realist evaluation of organisational collaboration in healthcare. *Implementation Science*, 11, 17.

Rycroft-Malone, J., Fontenla, M., Bick, D., & Seers, K. (2010). A realistic evaluation: The case of protocol-based care. *Implementation Science*, 5, 38. doi:10.1186/1748-9098-5-38

Salter, K. L., & Kothari, A. (2014). Using realist evaluation to open the black box of knowledge translation: A state-of-the-art review. *Implementation Science*, 9, 115.

The RAMESES II Project. (2017). Retroduction in realist evaluation. The RAMESES II Project—Resources and training materials for realist evaluation. Retrieved from http://ramesesproject.org/standards_and_Training_materials.php-RAMESES_II

Tolson, D., McIntosh, J., Loftus, L., & Cormie, P. (2007). Developing a managed clinical network in palliative care: A realistic evaluation. *International Journal of Nursing Studies*, 44, 183–195.

Van Belle, S. B., Marchal, B., Dubourg, D., & Kegels, G. (2010). How to develop a theory-driven evaluation design? Lessons learned from an
adolescent sexual and reproductive health programme in West Africa. 
*BMC Public Health*, 10, 741. doi:10.1186/1471-2458-10-741
Westhorp, G. (2008). Development of realist evaluation models and 
methods for use in small-scale community based settings (Doctor 
of philosophy). Nottingham, England: Nottingham Trent 
University.

Wong, G., Greenhalgh, T., Westhorp, G., Buckingham, J., & Pawson, 
R. (2013). RAMESES publication standards: Realist syntheses. 
*BMC Medicine*, 11, 21.
Wong, G., Westhorp, G., Manzano, A., Greenhalgh, J., Jagosh, J.,
& Greenhalgh, T. (2016). RAMESES II reporting standards for 
realist evaluations. *BMC Medicine*, 14, 96.