



Article

# Applying a realist evaluation to an intervention in children's social care: A worked example from the Safeguarding Family Group Conference study

Evaluation

2025, Vol. 31(1) 49–69

© The Author(s) 2025



Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/13563890241309643

[journals.sagepub.com/home/evi](https://journals.sagepub.com/home/evi)



**Bekkah Bernheim** 

University of Exeter Medical School, UK

**Lorna Stabler** 

Cardiff University, UK

**Jennie Hayes**

University of Exeter Medical School, UK

**Alankrita Singh**

Cardiff University, UK

**Katrina Wyatt**

University of Exeter Medical School, UK

## Abstract

Children's social care researchers are increasingly drawing on realist evaluation to understand the complexity within their field by identifying underlying contexts and mechanisms that lead to outcomes of interest. However, there are few published worked examples of realist evaluations of interventions in children's social care. This makes it challenging to understand how to put this approach to best use in practice. To address this gap, we share how we conducted a realist evaluation of Safeguarding Family Group Conferencing, a family-led decision-making process. In doing this, we highlight several challenges and opportunities of conducting a realist evaluation in a children's social care setting. We conclude that realist evaluation is adaptable and generative and benefits from a team-based approach to retroductive theorising and analysis. By outlining our

## Corresponding author:

Bekkah Bernheim, Department of Health and Community Sciences, University of Exeter, St. Luke's Campus, Heavitree Road, Exeter EX1 2LU, UK.

Email: [r.bernheim@exeter.ac.uk](mailto:r.bernheim@exeter.ac.uk)

process, we aim to provide a resource for children's social care researchers wishing to use realist evaluation in the United Kingdom and beyond.

### **Keywords**

children's social care, Family Group Conferencing, realist approach, realist evaluation, social work

## **Introduction**

### *Children's social care in the United Kingdom*

The children's social care system in the United Kingdom is described as being in 'crisis' (Mackley et al., 2018). Over the past two decades, austerity measures introduced by the UK government have significantly impacted local authority budgets and spending on children's social care. While funding for local government fell by 55 per cent between 2010 and 2020 (Ogden and Phillips, 2024), these cuts were not spread equally across children's services. Instead, spending on keeping families together fell significantly (Bennett et al., 2021), while overall spending in children's social care has more than doubled in the last decade (Webb, 2023), with spending increasingly focused on statutory intervention such as child protection investigations, removing children from their parents and keeping children in care (MacAllister, 2022; Webb et al., 2022). This trend has been perpetuated by cuts in preventive services driving 'demand' for statutory intervention (Bennett et al., 2021), further intensifying challenges, including high case-loads, workforce burnout and high staff turnover (Family Rights Group, 2018).

In England, local authorities have a statutory responsibility to step in when there are concerns about a child's safety specifically when professionals believe 'the child is suffering, or likely to suffer, significant harm' (Children Act, 1989; HM Government, 2023: 8). An Initial Child Protection Conference is held to assess concerns. At the Initial Child Protection Conference, social workers and partner agencies (i.e. health, police and schools) share information and take the lead to create a plan to keep the child safe. Evidence suggests that parents and children find the Initial Child Protection Conference proceedings exclusionary, adversarial and shaming (Gibson, 2015). However, in the United Kingdom, the need is being recognised for social work practice to move away from a top-down 'expert' culture towards one that is relational and addresses children's safety concerns through a process of co-production with families (Ferguson et al., 2022; Ingram and Smith, 2018). Family Group Conferencing (FGC) is one way to involve families and their social support networks in making decisions in partnership with professionals about the best way to protect their children (Family Rights Group, 2021; Pennell et al., 2011).

Evaluations of FGC have sought to identify and measure outcomes such as participant satisfaction, family ties, improved relations between families and professionals, children's safety, children remaining out of care and cost savings (Holland et al., 2005; Marsh and Crow, 1997; Mason et al., 2017; Merkel-Holguin, 2004; Munro et al., 2017; Pennell and Burford, 2000; Taylor et al., 2023). These evaluations include systematic reviews and large-scale randomised controlled trials (see Mason et al., 2017; Taylor et al., 2023) with a focus on the quantitative outcome of child placement (i.e. how many children were removed from their family or not), the impact on public money and projected cost savings. Capturing the detail of family-led processes using traditional evaluation methods is challenging because interventions are rarely

clear-cut, and implementation is often dependent on local contextual factors (Mezey et al., 2015). In addition, there can be challenges in recruitment and limited access to data (Mezey et al., 2015; Moody et al., 2021), impacting on how useful the findings might be. Without a substantial evidence base of what works, it is difficult to determine where limited resources should be allocated to have the greatest impact. Moreover, randomised controlled trials alone fail to capture *how* and *why* family-led decision-making processes such as FGC give rise to specific outcomes in different contexts.

The settings in which social care practices take place are open, dynamic systems which evolve and adapt with the wider eco-systems of geographical place and the relations between practitioners and families. As such, interventions in children's social care are typically adapted to their setting rather than tightly pre-specified, which is more common in health settings. Evaluative approaches which can capture the dynamic conditions that support or hinder an intervention achieving outcomes of interest in different settings are needed; realist evaluation is one such approach.

### *Using realist evaluation to understand interventions in children's social care*

Unlike traditional forms of evaluation, realist evaluation is less interested in answering the question *does an intervention or programme work?* but instead seeks to address how and why an intervention or programme works in different settings with different populations (Pawson and Tilley, 1997: 141). The task of the realist evaluator is to develop programme theories that hypothesise 'how a program is expected to work, given contextual influences and underlying mechanisms of action' (Jagosh, 2019: 363). Programme theories are the 'units of analysis within realist evaluation' (Dalkin et al., 2015: 3). Through the identification, articulation, testing and refining of these theories, realist evaluators undertake an iterative process of honing their hypotheses (Pawson and Tilley, 1997). Pawson and Tilley (1997) introduced the heuristic Context-Mechanism-Outcome configurations, or CMOs for short, as an analytical tool to operationalise programme theories. Context refers to key 'elements in the backdrop environment of a program that have an impact on outcomes (e.g. demographics, legislation, cultural norms)' (Jagosh, 2019: 363). Mechanisms are the 'resources offered through a program and the way people respond to those resources (e.g. information, advice, trust, engagement, motivation)' (Jagosh, 2019: 363). Disaggregating mechanisms into mechanism resources and mechanism responses is useful as it helps differentiate what is a mechanism and what is a context which adds clarity to the programme theory (Dalkin et al., 2015). Outcomes in realist evaluation are the 'intended or unintended effects based on context-mechanism interactions (i.e. changed outlook, service update, decision making, resiliency, health outcomes, self-efficacy, social connections)' (Jagosh, 2019: 363).

A central part of undertaking realist evaluation is grappling with the ontological and epistemological assumptions underpinning the programme or intervention that is being evaluated (see Jagosh, 2019, 2020; Pawson and Tilley, 1997). Four core ideas provide a useful starting point: (1) ontological depth, (2) generative causation, (3) abduction and (4) retroduction. Ontological depth is the notion that 'reality is stratified into layers' (Jagosh, 2019: 364). According to realists, what is real is not necessarily observable, and there is a need to go beyond empirical evidence to begin to understand reality. Following on from this, generative causation, in contrast to successionist causation, is 'the idea that underpinning hidden mechanisms generate outcomes' (Jagosh, 2019: 364) – meaning realists must look beyond empirical

evidence to theorise possible reasons behind what they observe. This requires both abduction and retroduction (Emmel, 2021). Abduction is ‘pragmatic theorizing with a focus on creativity as a logic of inference’ (Tavory and Timmermans, 2014). In lay terms, it is a ‘gut feeling, hunch or informed imagination that leads to new ideas for generating theories and testing possible mechanisms’ (Jagosh, 2020: 122). Retroduction is the foundational mode of inference used in realist evaluation and is the process of *laying claim* to causal mechanisms (see Emmel, 2021; Jagosh, 2019). Through abduction and retroduction, realist researchers enter into ‘a dialogue between evidence and ideas’ (Emmel, 2021: 103) about CMOs, how the programme or intervention works and for whom under which circumstances. The philosophical principles outlined here are the foundation for conducting realist research.

Reporting standards seek to demystify how to conduct realist evaluations (Wong et al., 2016). Others have followed this up with helpful ‘how to’ guidance. For example, Gilmore and colleagues (2019) published a detailed paper demonstrating how they conducted two realist evaluations of community health interventions in three case study sites. By providing in-depth descriptions of what they did and how they did it, they brought much-needed transparency to the analytical process involved in realist evaluations conducted in complex health settings. Although there is some overlap, the children’s social care setting poses unique challenges to researchers. For example, within children’s social care, the social worker is usually assigned to the child, but the ‘target’ of an intervention is typically a parent/carer. Understanding outcomes from these interventions therefore necessitates explicit theorising to understand how the intervention can lead to outcomes for children themselves who may not be directly engaged by the intervention.

In this article, we provide a worked example of a realist evaluation of Safeguarding Family Group Conferencing (hereafter Safeguarding FGC) in England. We draw on headings used by Gilmore and colleagues (2019) and Mukumbang and colleagues’ (2016a) in their realist evaluations to structure the paper. After a brief description of the Safeguarding FGC study, we introduce phase 1 of the research – building the initial programme theory. In this section, we show how we consulted with stakeholders to build an initial programme theory. In phase 2, we share how we designed interview guides and observation templates based on the initial programme theory and generated data. In phase 3, we demonstrate how we analysed the data and used the data to refine, refute and test our programme theory within and across cases. In phase 4, we outline how we represented our programme theory at the end of the study. In the final section, we highlight the key challenges and opportunities offered by realist evaluation in a children’s social care setting. We conclude with reflections on how realist evaluation is a useful approach that benefits from the inclusion of different types of evidence from literature, stakeholders and research data, adaptive recruitment strategies, and a team-based approach to retroductive theorising and analysis.

## **Safeguarding Family Group Conferencing: A worked example of a realist evaluation**

### *Introduction to the safeguarding FGC study*

Safeguarding FGC is a family-led decision-making process involving families and professionals when there is a child protection concern. The Safeguarding FGC study is a research-practice partnership between the University of Exeter Medical School; the Children’s Social Care Research and Development Centre (CASCADE) at Cardiff University; The Royal Borough of Kensington and Chelsea; Westminster City Council and the London Borough of Hammersmith

and Fulham. The aims of the research were to identify which families were most likely to benefit from Safeguarding FGC and in what way; to identify and understand enablers and barriers to implementing Safeguarding FGC; and to develop a detailed implementation package for national roll-out in local authorities in England. The research questions were as follows:

1. What works and in what way to enable uptake and embedding of Safeguarding Family Group Conferencing into the Child Protection pathway as an alternative to an Initial Child Protection Conference?
2. What outcomes are deemed most appropriate by families, including children and young people, and professionals?
3. For which families under which circumstances do Safeguarding Family Group Conferences enable a more positive experience of the Child Protection pathway and promote outcomes appropriate for families and professionals?

The Safeguarding FGC Study is funded by the National Institute for Health Research (Ref: NIHR131922).

This article focuses on the methodology we used to evaluate Safeguarding FGC (research questions 2 and 3). For more on research question 1, a publication is under review to report the qualitative findings from the implementation evaluation (Day et al., 2024). The full study protocol is also available (Stabler, 2024a). Ethical approval for the research was granted by the Research Ethics Committee at the University of Exeter Medical School (Reference Number 493165) and the participating local authorities. Participants were given information sheets about what taking part in the research would entail, including the possibility of publishing findings. They gave written consent or, where this was not possible, verbal consent was audio recorded in keeping with the parameters of our ethics approval.

The following phases were not a linear progression of tasks but happened in iterative cycles of data collection, team discussions and theory building and testing with stakeholders and participants. Dalkin and colleagues note that this process of developing and testing programme theories is ‘often convoluted’ (Dalkin et al., 2021: 124). By outlining our approach, we seek to bring transparency to the process and illustrate one way of doing a realist evaluation in practice.

### *Phase 1: Building the initial programme theory*

The initial programme theory is used to ‘hypothesise how, why, and for whom the intervention may work, based on literature and document reviews, and key informant interviews with programme architects or implementers’ (Gilmore et al., 2019: 3; Mukumbang et al., 2016b; Pawson and Tilley, 1997). The initial programme theory functions as a ‘first draft’ explanation of how the programme works. We use the term *initial programme theory* to refer to our first attempt to explain how Safeguarding FGC works based on stakeholder engagement, literature and past research. We intentionally then use the term *programme theory* to refer to the theory we tested and refined with data. We use the singular term ‘programme theory’ to refer to our overarching idea of how Safeguarding FGC works and ‘programme theories’ (plural) to refer to the CMO configurations that make up the wider theory. We expressed our programme theories at every stage of the realist evaluation as ‘if . . . then’ statements (see also Brand et al., 2019). ‘If . . . then’ statements are one way of articulating hypotheses about ‘*how* programme outcomes are manifested through programme mechanisms and corresponding contexts’

(Jagosh et al., 2022: 2). The ‘if . . . then’ format helped the research team to keep the identification of generative mechanisms in mind from the start. The terms ‘if . . . then’ statements, CMOs and programme theories are used interchangeably.

The Safeguarding FGC study is based on a rapid realist review (Stabler et al., 2019) and the findings of a 2019 pilot study in London (Bowring and Daly, 2021). The programme theory about how Safeguarding FGCs work and which families appeared to benefit informed the development of Safeguarding FGC in practice in the 2019 pilot site and was the impetus for the current study. The programme theory from the review and the 2019 pilot formed the basis of our initial programme theory and included a range of ‘intermediate interpersonal mechanisms’ about how and why the Safeguarding FGC gave rise to outcomes in particular contexts (Stabler et al., 2019). These explanations included (1) reducing shame and blame in meetings for families and professionals; (2) parents participating more in decisions about how to keep their child safe; (3) parents and their wider support group feeling empowered; (4) the child’s voice being central to decision-making; and (5) professionals feeling less concerned about risk by knowing a fuller picture of the family’s life. Given the small scale of the 2019 pilot study and the time that had passed since it was conducted, we took this programme theory as a starting point to develop a programme theory for the study, rather than test this programme theory directly in the study.

**Stakeholder engagement.** To update the 2019 programme theory, we conducted 13 stakeholder interviews with 14 social care and partner agency professionals who had experience of implementing and/or delivering Safeguarding FGCs, observed three meetings where Safeguarding FGC was discussed by professionals from the pilot site and engaged with partner local authority team members. Questions from the stakeholder engagement interviews were prioritised and agreed upon by the study team, some of whom had worked on the 2019 programme theory. Questions focused on implementing and delivering Safeguarding FGC and related to the following areas: risk management, engagement of children and young people, family engagement and ownership, which families might benefit from the intervention, core components of the Safeguarding FGC and outcomes.

**Developing ‘if . . . then’ statements.** The following bullet points outline the steps we took to develop a set of ‘if . . . then’ statements capturing the study team’s ideas about how Safeguarding FGC could work and informed by initial, extensive stakeholder engagement.

- Created 240 ‘if . . . then’ statements based on researcher notes from stakeholder engagement and the research team’s informed hunches and theorising about possible causal pathways.
- Organised and grouped the ‘if . . . then’ statements in Excel by theme.
- Consolidated the 240 ‘if . . . then’ statements down to 18 through discussions with the study team and key stakeholders. One of the challenges of realist evaluation is managing the expansion of the initial programme theory as new hunches and ideas are discussed, and then reducing the number of ideas that will be taken forward in the evaluation. This inevitably meant that we had to let go of many of our initial hunches.
- Deprioritised nine consolidated ‘if . . . then’ statements based on realist prioritisation criteria (Jagosh, 2022b).
- Used the remaining nine ‘if . . . then’ statements to develop the interview guides and the observation template. The observation template included headings corresponding to the



stages of an FGC (see observation template in Supplementary Table S1). Under each heading, prompt questions based on the initial programme theory were included to guide the researcher's notetaking.

At this stage, we felt we had articulated as much as possible how we thought Safeguarding FGC might work and were ready to move on to data collection.

### *Phase 2: Data collection to test and refine the programme theories*

In this section, we outline how the initial programme theory informed the interview guides and observation template. We then describe and justify the data collection methods used and the sample size.

Qualitative data were collected over 9 months in three case study sites in England. The three local authorities participating in the research opted to pilot Safeguarding FGC before deciding to scale up, or not. The reasons for this vary and are in part due to the scrutiny faced by children's services by regulators such as Ofsted, limited resources and a desire to first see evidence of success (for more on implementation see Day et al., 2024). This meant only a small number of families in each case study site had a Safeguarding FGC during the study's recruitment period. Of these, three families agree to take part in the research. Emmel asserts that when collecting qualitative data as part of a realist study, 'reporting that 1 or 200 cases were collected is not as important as the ways in which insights into events and experiences are used for interpretation, explanation, and claims from research' (Emmel, 2013: 141). Manzano echoes this, arguing 'the importance is not on "how many" people we talk to but on "who", "why" and "how"' (Manzano, 2016: 349). The limited number of families recruited to the Safeguarding FGC study masks the richness and depth of the qualitative data collected. For every participating family, a researcher observed and took anonymised notes at up to three meetings between the family and professionals (including the Safeguarding FGC and review meetings) using a notetaking template. Everyone who attended the Safeguarding FGC was then invited by the researchers to take part in two interviews, the first within 1 month of the Safeguarding FGC and the second within 3 months. Qualitative data also included researcher reflective notes from meetings and conversations with professionals. For the first family who consented to the research, a researcher observed and took notes at 3 meetings each lasting at least 1.5 hours and conducted 12 interviews with 8 individuals, including family members and professionals, at two time points over the course of 6 months. Despite the limitations, the researchers were able to collect in-depth data that shed light on family's and professionals' experiences and perceptions of the Safeguarding FGC process and how it worked.

*Developing interview guides and observation templates to test and refine the programme theory.* The initial programme theory was used to create interview guides and an observation template for data collection. The aim of the realist interview is to shed light on how participants 'understand and have experienced the programme and compare those experiences with our hypotheses about how the programme is working' (Manzano, 2016: 349). We created different interview guides for parents/carers, extended family/friends, children/young people, social care practitioners and partner agencies based on which aspects of the programme they were best positioned to comment on. We also consulted with a parents' group with lived experience

**Table 1.** Example of how the initial programme theory was used to develop interview questions.

'If . . . then' title: Social care practitioners collaborate with families	
<b>'If . . . then' statement</b>	If practitioners put aside existing assumptions and show they are interested in the family perspective and wider social, cultural and structural context of their lives (context) AND demonstrate appropriate and proportional trust through genuine and transparent collaboration with families THEN practitioners will be able to have a conversation with families about the purpose and benefits of Safeguarding FGC (mechanism resource) AND THEN families will realise the potential of Safeguarding FGC (mechanism response) and be more willing to say yes to the offer (outcome)
<b>Interview questions for parents/carers</b>	<ol style="list-style-type: none"> <li>1. Can you tell me how the idea of your family having a Safeguarding FGC was introduced to you?           <ul style="list-style-type: none"> <li>- Who first told you about the Safeguarding FGC?</li> <li>- How did they explain the purpose of the Safeguarding FGC to you?</li> <li>- How did they explain the benefits?</li> </ul> </li> <li>2. Can you tell me why you decided to accept the offer of a Safeguarding FGC?           <ul style="list-style-type: none"> <li>- Was there anything about how practitioners explained the Safeguarding FGC offer that led you to agree to it?</li> </ul> </li> <li>3. There is this idea that if social workers show they are genuinely interested in families' points of view and are willing to work collaboratively with families, then families will be more open to say yes to the Safeguarding FGC offer.           <ul style="list-style-type: none"> <li>- What do you think about that?</li> <li>- What made you feel like they were genuinely interested in your point of view?</li> <li>- What does it look like for practitioners to work collaboratively with families?</li> <li>- Is there anything else you think would help families say yes to a Safeguarding FGC?</li> </ul> </li> </ol>

of children's social care to get their input on interview questions and the observation template. One part of the Safeguarding FGC we were interested in was understanding the underlying mechanism(s) that led families to say yes to the offer of a Safeguarding FGC as an alternative to an initial child protection conference (the usual way of working between families and professionals when there is a safeguarding concern). Table 1 shows how we translated an 'if . . . then' statement from our initial programme theory into interview questions for our interview guide for parents/carers.

The first two questions developed based on the 'if . . . then' statement in Table 1 were left intentionally open to try and understand what it was that led parents/carers to say yes to having a Safeguarding FGC. Open questions and probes that explored participants' experiences of how different parts of the intervention worked for them ensured that the research captured a range of ideas beyond what we had theorised. We combined these open questions with a direct interviewing technique used in realist evaluation where aspects of the initial programme theory are presented to interviewees (Manzano, 2016). In this approach, the researcher explains all or part of the CMO to an interviewee and invites them to agree, disagree or add to it (see question 3 in Table 1) (Manzano, 2016). The benefit of this type of questioning is that participants have the opportunity to directly comment on the logic of the CMOs, clarify them, add to them or refute them. In practice, we found that this direct approach had its limitations



depending on how confident participants were being presented with ideas, disagreeing with the researcher and posing alternative explanations. Griffiths and colleagues note this approach to interviewing is ‘likely to be more effective with policy makers and some practitioners’ (Griffiths et al., 2022: 2). Given these limitations and because we wanted to capture a range of participant responses, most questions asked during interviews were open (like questions 1 and 2 in Table 1), but still directly based on the programme theory, and developing an understanding of how the programme worked in different settings for different people. Posing questions in a variety of ways was helpful and generated data that furthered our understanding of how the interviewees experienced the Safeguarding FGC. We then compared and contrasted those experiences with other data and with our programme theories (Manzano, 2016).

### *Phase 3: Data analysis*

Realist data analysis is iterative and as data were collected, it was analysed in a cyclical process of testing and refining (Wong et al., 2016). A key difference between realist evaluation and other types of evaluation is the use of retroductive theorising. Retroduction goes beyond inductive and deductive reasoning and involves, ‘going back from, below, or behind observed patterns or regularities to discover what produces them’ (Lewis-Beck et al., 2004). Retroduction is ‘the activity of theorizing and testing for hidden causal mechanisms responsible for manifesting the empirical’ (Jagosh, 2020: 121). As data were collected and fed into the analysis, the programme theories evolved, and the researcher teams’ understandings of the nuances of how the programme worked for different families and in different contexts came into focus.

Interview data were recorded and transcribed verbatim. Transcription was done externally due to time constraints. Notes from observations were typed up using observation templates. Researcher’s reflective notes were also typed up and included in the analysis. All participants, places and organisations in the data were pseudonymised to protect anonymity.

*Step 1: Familiarisation.* Realist evaluation requires the researcher to look for causal chains within the data that lead to intended and unintended outcomes. Causal insight may not be fully evident in a particular quote from an interview because, according to realist methodology, generative mechanisms exist beneath the empirical (i.e. ontological depth) and through retroduction realist researchers theorise what these might be. This is why we first considered the data holistically; reading over each interview transcript and reflecting on what insight it provided into how the programme works (Jagosh, 2022a). This helped the researchers to get a sense of the main causal insights of the data. To capture these insights, the researchers wrote a short summary of each transcript with the programme theory in mind and identified key themes emerging from the data. The summary and themes were added to the top of each piece of data before it was prepared for more detailed coding.

*Step 2: Coding and annotating in NVivo.* Data were uploaded to the QRS International software programme NVivo 12. NVivo is a useful tool to organise the analysis of a large and diverse qualitative data set. Dalkin et al. (2021) write about the use of NVivo and other computer-assisted qualitative data analysis software in realist evaluation. They note the need for researchers to use NVivo flexibility in a way that fits with the specifics of their team and study. The following section outlines our approach to using NVivo.

We used NVivo as a tool to organise and code data. We did not use NVivo as an exhaustive record of all decision-making related to theory development. Instead, we tracked changes to the

**Table 2.** Example parent code and child codes.**Example parent code:** Social care practitioners collaborate with families

**Code description:** If practitioners put aside existing assumptions and show they are interested in the family perspective and wider social, cultural and structural context of their lives (context) AND demonstrate appropriate and proportional trust through genuine and transparent collaboration with families THEN practitioners will be able to have a conversation with families about the purpose and benefits of Safeguarding FGC (mechanism resource) AND THEN families will realise the potential of Safeguarding FGC (mechanism response) and be more willing to say yes to the offer (outcome).

Child codes	Child code descriptions	Fictional example to guide coding
Supports	Does the data support or reflect the existing 'if . . . then' statement, in that it is sufficiently similar to the 'if . . . then' statement to be integrated without further changes or nuance?	Practitioner says in an interview, 'when I visited the mother, her friend was there, and it challenged my preconceptions. I realised that she lives in a supportive community. We then had a conversation about Safeguarding FGC and both walked away thinking, yes, this could work!'
Refines	Does the data add important refinements to the understanding of contexts, mechanisms or outcomes, provide nuance, further detail or understanding to the existing 'if . . . then' statement? Should the 'if . . . then' statement, therefore, be changed to reflect this new data, and the increased explanatory power it offers?	The parent says in an interview, 'When the FGC Coordinator visited me, she explained the process, benefits, and everything. She said she's independent from the Local Authority and she'd be facilitating the process. That put my mind at ease, and I said yes to the offer straight away'. [Refinement –change 'practitioner' to 'FGC Coordinator' in the 'if . . . then'.]
Challenges	Does the data challenge, contradict or disconfirm any aspect of the existing 'if . . . then' statement, either in full or in part? Does it, therefore, require a reworking of the 'if . . . then' statement, or is it a rival theory? Or is it adding important contextual info about 'for whom'?	The parent says in an interview, 'When the FGC Coordinator visited me, she was nice and told me all about Safeguarding FGC. I decided not to go for it though because I don't want anyone knowing what I'm going through.' [Potentially a more important context – parent's willingness to involve their network]
Connects	How does the data connect 'if . . . then' statements to each other?	Saying yes to the offer of Safeguarding FGC (the outcome) could be the context for the 'if . . . then' statement about families being receptive to professionals' concerns.

programme theory using Excel. A coding framework was developed in NVivo using nine CMOs from the initial programme theory expressed as 'if . . . then' statements. The short title of each CMO was entered into NVivo as a parent code. Each of these nine CMOs (parent codes) were assigned four child codes. Child codes are sub-codes which further categorise the data to support the analysis. For each parent code, the child codes were (1) Supports, (2) Refines, (3) Challenges and (4) Connects (see Table 2 for descriptions of each child code). Qualitative data from interviews, observations and researchers' reflective notes were coded using this framework.

We coded each selected data segment to a parent code (one of the nine programme theories), and to one of the four child codes – 'supports', 'refines', 'challenges' or 'connects'.

**Table 3.** Example of how data were coded and annotated.

Data	Parent code	Child code	Annotation
<p><b>Interviewer:</b> Yeah, yeah. And did he [the social care practitioner] explain any of like, [the] possible benefits?</p> <p><b>Parent:</b> Yeah, everything, yeah.</p> <p><b>Interviewer:</b> Yeah, what did he kind of have to say?</p> <p><b>Parent:</b> Well, like, we . . . obviously we didn't know the benefits until we come up with the plan, but obviously the benefits were that we were not going to child protection, to avoid that. And more about getting support for my family, my family being there, and recognising that I'm not on my own, not doing it on my own.</p> <p><b>Interviewer:</b> Yeah.</p> <p><b>Parent:</b> Whereas I think I would have if I'd been on the child protection. It would have been a lot different for us.</p>	<p>Social care practitioners demonstrate trust in families</p>	<p>Supports</p>	<p>This data shows that the parent/carer understood Safeguarding FGC as a process that relies on family involvement and something that is different to the 'usual' ways of working (i.e. the Initial Child Protection Conference). Understanding these two benefits of Safeguarding FGC informed her decision to agree to it (researcher's initials and date).</p>

This way, we were able to see which data related to which of our programme theories, and if that data supported, refined or challenged our hypotheses or connected to other programme theories.

The interview guides and observation template were developed based on the updated programme theory and the questions and prompts corresponded to specific programme theories; this helped provide a sense of which programme theories to code the data to. We used our judgement to look for causal insight across the whole programme theory when coding the data. For each coded data segment, we also created an annotation in NVivo. The annotation function in NVivo allows the researcher to create a note attached to a specific segment of data. For each piece of data we coded, we created an annotation where we wrote the code title, the child code, and one or two sentences explaining the reasoning behind the coding, along with the researchers' initials and the date. This helped us to track the rationale, decision-making and further justification of coding specific data segments to different codes and child codes (see Table 3 for an example from our data).

We coded data as soon as possible after it was collected and transcribed. Towards the start of the data collection period, we coded a significant portion of data to the 'refines' child codes. This is because we were moving beyond our initial hunches and honing our understanding of how the programme was working by speaking to people who were experiencing it. As the study progressed, we coded much more of the data to 'supports' as our 'if . . . then' statements became clearer and were tested. It should be noted that the coding process was an analytical tool, and we did not go back to recode data when our thinking about a specific CMO evolved. Instead, we conferred with others about our hypothesis.

**Step 3: Consolidating.** To gain a better understanding of the data, the research team developed a spreadsheet to capture key insights. The spreadsheet had the name of each data file along the Y axis and the parent codes along the X axis. The researcher who collected the

data then included a short summary of the key insights related to each parent code for every interview transcript, observation and reflective note. This included the number of data segments coded to the parent code, and then a few sentences summarising how the data supported, refined or challenged the ‘if . . . then’ statement. This spreadsheet provided a way to share insights across the team and formed the basis for further discussions in data sharing meetings.

*Step 4: Conferring.* Much of the analysis was done through conversation between the researcher team, and between the research team and key stakeholders in data sharing meetings, theory meetings and expert stakeholder group meetings.

*Data sharing meetings.* Data sharing meetings were held weekly between the members of the study team primarily responsible for collecting data. The aim of data sharing meeting was for the researchers working in the case study sites to talk about the data they wanted to discuss the meaning of, or to highlight data that they felt may affect the existing logic of the ‘if . . . then’ statements. In each meeting, with reference to the key insights spreadsheet, we usually looked at data coded to one ‘if . . . then’ statement from different sources and across different case study sites. We discussed details of the data collected and checked our interpretations against that of other researchers. All potential suggested changes to the programme theory were recorded and then brought to the wider research team at theory meetings.

*Theory meetings.* Theory meetings were held roughly every 4–6 weeks. Dates where the team could be together in person were prioritised due to the complex nature of the discussions. The agenda for the theory meetings changed depending on what new data had been collected. In total, 15 theory meetings lasting at least 2 hours each were held throughout the project. Sometimes only one ‘if . . . then’ statement was discussed, during other meetings, the programme theory was considered as a whole. The overarching aim of these meetings was to lay claim to and test generative mechanisms with data and through discussion, and to identify key contexts and outcomes.

In most theory meetings, coded data were presented and considered related to one or two ‘if . . . then’ statements. Based on the insights from the data, and discussions in data sharing meetings, suggested changes with excerpts from the data were presented. At other times, the researcher shared the ‘story’ of the data, and the research team had an open discussion of what might be happening, and which ‘if . . . then’ statements this connected to. Theory meetings acted as a space for the research team to make joint decisions to improve the articulation of causal chains within the ‘if . . . then’ statements and agree on changes.

Theory meetings also functioned as a way of ensuring the theory was capturing the most important contexts, mechanisms and outcomes and deciding which CMOs to prioritise or deprioritise. Periodically, theory meetings also provided an opportunity to ‘sense check’ how all the programme theories fit together, particularly towards the end of the study. The team used the following questions as a guide:

1. Does the causal chain within the ‘if . . . then’ statement make sense logically?
2. Are we capturing what is most important to explain how Safeguarding FGC does, or does not, work?

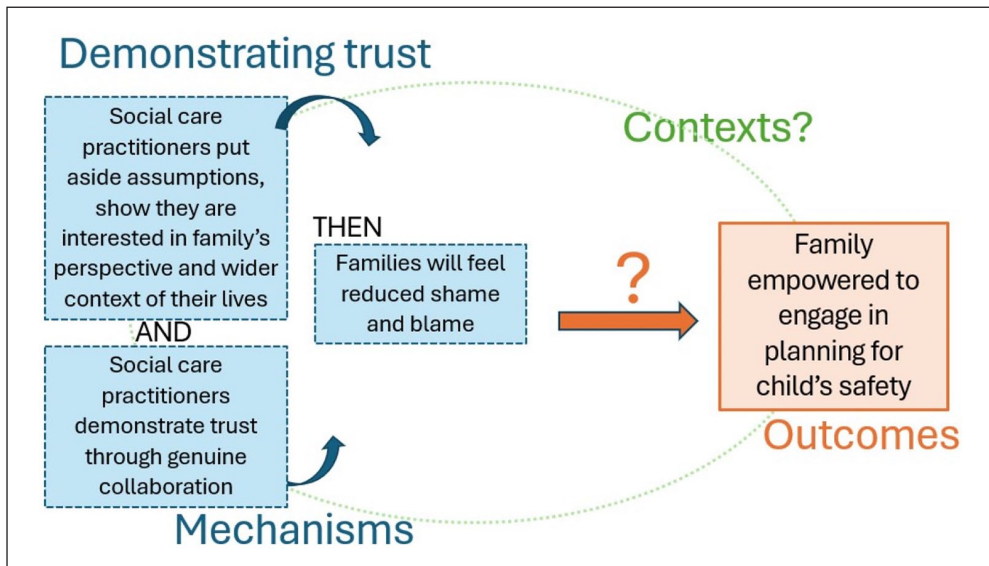
3. Are we missing anything, or have we lost anything as we have refined the ‘if . . . then’ statements that we want to add back in?
4. Have we captured the most relevant context? (with a prompt to consider contexts that may sit on different levels of the socio-ecological model – individual, interpersonal, institutional, community and policy)
5. What is at the core of this ‘if . . . then’ statement and what keywords could we use to summarise it?
6. Are we making any claims in the ‘if . . . then’ statements that we feel go too far beyond our evidence?

Theory meetings were audio recorded, and after each meeting a member of the research team listened back to the recording and made changes to the programme theories based on the discussion to improve their logic, causal pathway and articulation. The new version of the programme theory was then circulated to the team and agreed. The updated programme theory was then used as the basis for the next round of coding and analysis.

*Expert stakeholder group.* To provide additional sense checking, insight and testing of the programme theory, the researchers periodically discussed CMOs with the study’s Expert Stakeholder Group. The Expert Stakeholder Group included individuals with lived experience of children’s services, as well as practitioners and academics working in the field of children’s social care. Different configurations of the Expert Stakeholder Group met six times between July 2022 and June 2024. The study team presented the programme theory, or parts of programme theory, to the group in the form of ‘if . . . then’ statements or shared topics and themes related to the programme theory. The Expert Stakeholder Group then suggested nuances, refinements and new insights that helped to clarify the CMOs and overall logic of the programme theory through discussion and by responding to questions based on the programme theory using the interactive presentation software Mentimeter. Like Griffiths et al. (2022), we adapted our approach to gathering feedback on the programme theory to the particular group of individuals we were speaking to. All proposed changes and ideas were then discussed by the study team and recorded in an Excel spreadsheet. Of the original 18 CMOs, the nine that we prioritised and refined are the ones that the Expert Stakeholder Group members told us were most important.

#### *Phase 4: Representing the programme theory*

Realist researchers take different approaches to representing their programme theories. Many of these approaches involve written descriptions or narratives for each CMO that explain how it worked (Greenhalgh et al., 2009; Jagosh et al., 2015). Sometimes these are accompanied by visual representations of individual CMOs or the programme theory as a whole (see, for example, Mukumbang et al., 2016b). At times, we used visual representations of our programme theories in the form of diagrams when presenting to stakeholders and at theory meetings. An overarching diagram of the initial programme theory was used to theorise how CMOs fit together and to identify gaps in our knowledge (see study protocol Stabler, 2024a: 12). Later in the study, visuals of individual ‘if . . . then’ statements were used to help simplify the programme theories for discussion (for an example, see Figure 1 – an early draft ‘if . . . then’ statement diagram about demonstrating trust). We found that these visuals were useful tools



**Figure 1.** A diagram version of an early draft 'if . . . then' statement used to generate discussion with stakeholders.

for building and refining the programme theories, especially with stakeholders. While diagrams worked well to prompt discussion and theorising, we decided a narrative description was necessary to capture the nuance and complexity when sharing our findings more widely.

At the end of the study, we chose to represent our programme theory by writing explanatory narratives with examples from the data for each 'if . . . then' statement. We developed the following questions and prompts to guide the writing process:

For each 'if . . . then' statement, consider the following questions:

- What do we mean by this CMO?
- What might this look like?
- What might be a barrier to this CMO?
- What might make this CMO more likely to happen?
- How does this CMO connect to other CMOs (the outcome for one may become the context for another)?

Prompts to help write the narratives:

- Some examples of this are . . .
- This seems important: (a) across the whole programme theory or (b) in this circumstance.
- In area A, we see X happening in these ways, whereas in area B it might look like . . .
- We did not see this, but it might be that if X is in place, then Y will happen.
- This is evidenced in this literature . . .

The final resulting document was a comprehensive and long-form version of our programme theory which we were able to edit and adapt when presenting it to different audiences, for



example, to feedback to the participating local authorities, to write up in a report for our funders and to submit to academic publications.

In this article, we have provided a worked example of how we conducted a realist evaluation of Safeguarding FGC. In the following section, we reflect on the specific challenges and opportunities we encountered doing realist research within a children's social care setting.

## **Key challenges and opportunities afforded by realist evaluation in children's social care setting**

There are several key challenges to carrying out evaluations within children's social care that are not all unique to realist evaluation. We have drawn out some challenges we experienced associated with each of the phases articulated above. We have also attempted to show how the realist evaluation methodology allowed us to overcome some of these challenges.

### *The quality of existing evidence*

The starting point for a realist evaluation is an initial programme theory which draws on published literature and stakeholder consultation to articulate an initial understanding of how a programme works which can then be tested. Evaluative literature within children's social care is relatively underdeveloped when compared with disciplines such as health and education (Proctor, 2017), due in part to disagreements about the right approach to evaluation in this setting (Frost and Dolan, 2021). Within the literature that does exist, there is often a lack of explicit theorising of how interventions are thought to work (Stabler et al., 2022). This limits the pre-existing data available to draw upon in an evaluation in this setting. However, realists are encouraged to seek theoretically rich data from different sources to develop an initial programme theory. This includes looking outside of the discipline of the study. Social work has many allied disciplines such as psychology and counselling which can be drawn on to build an understanding of social work interventions. Moreover, realist evaluations can build on other work that has been carried out in the same area, or of similar interventions, without them explicitly having been designed as a programme of work. This study drew on other studies carried out by the team focused on interventions to reduce the need for children to enter care, including a scoping review (Stabler et al., 2022), a rapid realist review with a specific focus on interventions similar to the one evaluated (Stabler et al., 2019; Stabler, 2024b) and a pilot evaluation of the intervention with three local authorities (Bowring and Daly, 2021). This provided a strong starting point for understanding how the intervention might work. Understanding can then be further strengthened through stakeholder engagement, where people with experience of parts of the programme theory can share their expertise which can be synthesised alongside data from literature and other studies to produce an initial programme theory.

### *Participant recruitment*

Another enduring challenge in children's social care research is participant recruitment. Recruitment is hard because people who are experiencing social work interventions are likely to be going through a difficult time in their lives (Flaherty and Bromfield, 2021), and practitioners are often managing high workloads and other barriers to participation in research

(Pulman and Fenge, 2023; Yoon et al., 2022). This can mean that participation in research can feel like an additional burden on top of expectations on family members and practitioners to take part in a lot of meetings with numerous professionals. It can also be difficult for researchers at this time to explain the independence of their role, and the confidential nature of taking part, when families are at a point where a lot of information about their lives might be being shared in a way they are not comfortable with (Yoon et al., 2022). Moreover, local authorities are not equipped with the infrastructure and resources to facilitate research in the same way that health organisations are (Mezey et al., 2015). However, within realist evaluation, there is not a reliance on ‘high’ participant numbers as ‘the unit of analysis is not the person, but the events and processes around them’ (Greenhalgh et al., 2017: 2). Participants are selected based on what knowledge they may have about different parts of the programme theory. For example, frontline practitioners might have important perspectives on who interventions do and do not work for, whereas people who receive the intervention have insight into how it worked for them. Therefore, rich qualitative data from case studies can use a sampling frame whereby ‘sampling’ is based on contexts rather than pre-proscribed characteristics of families (Greenhalgh et al., 2017).

### *Data analysis*

Interventions in children’s social care may well be loosely defined rather than clearly manualised. Often, interventions in this setting are not just about ‘what’ is done, but how it is done, drawing on relationships and ways of interacting as a core intervention component (Frost and Dolan, 2021). In addition, the remit of children’s social care work is broad, meaning that outcomes may not be clear, may be difficult to measure and may be contested between different stakeholders (Forrester, 2017). Within this setting, a realist approach can feel complicated as it may be difficult to clearly articulate the programme architecture in relation to the mechanisms. The process of developing an initial programme theory drawing on the expertise of stakeholders is one way to address this, as noted above. Another opportunity afforded is a team-based approach to analysis which allowed the data to be considered through discussion, drawing also on the expertise and knowledge of the team and their understanding of how the programme was working. This is particularly helpful in a setting where data is seeking to uncover different knowledges, with different languages and jargon attached to them. Team-based analysis can allow for clarifying understanding but also highlighting which parts still feel uncertain. These can then be prioritised for further testing using the iterative ‘teacher-learner’ (Greenhalgh et al., 2017) model of data generation to interview participants, or by seeking further literature.

### *Developing relevant and accessible evidence*

The work of the realist evaluator is to elicit and bring together what different people think that they know about how a programme works. In this way, a realist evaluation within a children’s social care setting can generate evidence that speaks to practitioners, policy makers and families. Rich qualitative data about how programmes work can give nuances to ideas about causal chains, and useful real-world examples. A well-articulated programme theory should bring together this data to reflect the experiences of those familiar with the programme. In a sense, it is presenting back to participants what they already ‘know’. However, it is not possible to present the level of detail that is present in the data. The art is ensuring there is the level of detail

that can give helpful information to guide practitioners, policy makers, families, other researchers and anyone for whom the research findings are relevant, without including so much information that it is overwhelming. The ‘CMO’ heuristic helps to theorise ontologically deep generative mechanism beyond the empirical, while also maintaining a focus on causality, allowing for more or less information to be included depending on the audience, helping to keep the findings as relevant as possible to the audience.

## Conclusion

In this article, we showed how we conducted a realist evaluation of the Safeguarding Family Group Conference study in three local authorities in England. Through an iterative process of collecting, coding, analysing and discussing the data, the research team tested and refined the initial programme theory with significant input from key stakeholders. Mukumbang and colleagues note that ‘realist evaluation starts and ends with a theory’ (Mukumbang et al., 2016a: 6). In this article, we described how we went through a process of stakeholder engagement, data collection, analysis and writing and developed, tested and refined our programme theory. Realist evaluation is based on the underlying philosophical argument that there is no ‘final truth or knowledge’ rather the desired end result of a realist evaluation is a ‘*better* understanding of whether, how and why programmes work’ (Greenhalgh et al., 2019: 1). In the pursuit of better explanations, realist evaluators could continue to refine their programme theories indefinitely (Manzano, 2016: 356). The reality is that research, including the Safeguarding FGC study, is constrained by funding and project timescales. We invite other researchers to pick up where we left off and continue to test and refine the theory we developed. Finally, it should be noted that this article presents one way of conducting a realist evaluation. We recognise that this is not the only way to do so and would encourage others embarking on realist evaluation to adapt the approach to fit their specific project, team and timeline. Realist evaluation is flexible and benefits from the inclusion of different types of evidence from literature, stakeholders and research data, adaptive recruitment strategies, and a team-based approach to retroductive theorising and analysis.

## Acknowledgements

This study received the support of the Clinical Research Network (Reference CPMS54314). Katrina Wyatt is partially supported by the National Institute for Health Research Applied Research Collaboration South West Peninsula. The views expressed in this publication are those of the author(s) and not necessarily those of the National Institute for Health Research or the Department of Health and Social Care. Special thanks to Anna Rockhill for providing valuable comments on an early version of the 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: This study was funded by the NIHR Health Services and Delivery Research Programme (Grant reference number NIHR131922). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.


## Ethical approval

The study was approved by the University of Exeter Research Ethics Committee (Reference Number 493165) and participating local authorities.

## Consent to participate and publish

Participants were given information sheets about what participating in the research would entail, including the possibility of academic publication. Participants gave written consent or, where this was not possible, verbal consent was audio recorded in keeping with the parameters of the ethics approval. Participants, places and organisations were pseudonymised.

## ORCID iDs

Bekkah Bernheim  <https://orcid.org/0000-0002-6483-7622>

Lorna Stabler  <https://orcid.org/0000-0003-0742-4271>

## Data availability

Due to the sensitive nature of the data collected, the small sample size and in keeping with ethics requirements, data were not made publicly available.

## Open access

For the purpose of open access, the author has applied a ‘Creative Commons Attribution’ (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

## Supplemental material

Supplemental material for this article is available online.

## Note

1. In England, a concern is deemed to be a child protection concern ‘when there is reasonable cause to suspect that a child is suffering, or is likely to suffer, significant harm, whether the harm is from inside or outside the home including online’ (HM Government: 14).

## References

- Bennett DL, Webb CJR, Mason KE, et al. (2021) Funding for preventative Children’s Services and rates of children becoming looked after: A natural experiment using longitudinal area-level data in England. *Children and Youth Services Review* 131: 106289.
- Bowring H and Daly K (2021) Alternative Pathways-Safeguarding Family Group Conference (SFGC)/ Initial Child Protection Conference (ICPC) Pilot Project 2019-2021 Unpublished report.
- Brand SL, Quinn C, Pearson M, et al. (2019) Building programme theory to develop more adaptable and scalable complex interventions: Realist formative process evaluation prior to full trial. *Evaluation* 25(2): 149–70.
- Children Act (1989). Available at: <https://www.legislation.gov.uk/ukpga/1989/41/contents> (accessed 3 June 2024).
- Dalkin SM, Greenhalgh J, Jones D, et al. (2015) What’s in a mechanism? Development of a key concept in realist evaluation. *Implementation Science: IS* 10(1): 49.
- Dalkin SM, Forster N, Hodgson P, et al. (2021) Using computer assisted qualitative data analysis software (CAQDAS; NVivo) to assist in the complex process of realist theory generation, refinement and testing. *International Journal of Social Research Methodology* 24(1): 123–34.

- Day J, O'Rourke S, Stabler L, et al. (2024) Professional views of the contextual factors influencing the implementation of an innovation in child protection services in England: A qualitative analysis using the Consolidated Framework for Implementation Research. *Manuscript under review*.
- Emmel N (2013) *Sampling and Choosing Cases in Qualitative Research: A Realist Approach*. London: Sage.
- Emmel N (2021) Post-disciplinary realism. *International Journal of Social Research Methodology* 24(1): 95–108.
- Family Rights Group (2018) Care crisis review: Options for change. Available at: <https://frg.org.uk/wp-content/uploads/2021/04/Care-Crisis-Review-Options-for-change-report.pdf>
- Family Rights Group (2021) What is a family group conference? Available at: <https://frg.org.uk/family-group-conferences/what-is-a-family-group-conference/>
- Ferguson H, Warwick L, Disney T, et al. (2022) Relationship-based practice and the creation of therapeutic change in long-term work: Social work as a holding relationship. *Social Work Education* 41(2): 209–27.
- Flaherty R and Bromfield L (2021) Pregnant women involved with statutory child protection services: The impact of difficult-to-reach on recruiting a non-biased sample. *Social Work & Social Sciences Review* 22(1): 42–63.
- Forrester D (2017) Outcomes in children's social care. *Journal of Children's Services* 12(2-3): 144–57.
- Frost N and Dolan P (2021) Theory, research and practice in child welfare: The current state of the art in social work. *Child & Family Social Work* 26(3): 498–506.
- Gibson M (2015) Shame and guilt in child protection social work: New interpretations and opportunities for practice. *Child & Family Social Work* 20(3): 333–43.
- Gilmore B, McAuliffe E, Power J, et al. (2019) Data analysis and synthesis within a realist evaluation: Toward more transparent methodological approaches. *International Journal of Qualitative Methods* 18: 160940691985975.
- Greenhalgh T, Humphrey C, Hughes J, et al. (2009) How do you modernize a health service? A realist evaluation of whole-scale transformation in London. *The Milbank Quarterly* 87(2): 391–416.
- Greenhalgh T, Pawson R, Wong G, et al. (2017) The realist interview. *The RAMESES II Project*. Available at: [https://www.ramesesproject.org/media/RAMESES\\_II\\_Realist\\_interviewing.pdf](https://www.ramesesproject.org/media/RAMESES_II_Realist_interviewing.pdf)
- Greenhalgh T, Pawson R, Wong G, et al. (2019) Philosophies and evaluation design. *The RAMESES II Project*. Available at: [https://www.ramesesproject.org/media/RAMESES\\_II\\_Philosophies\\_and\\_evaluation\\_design.pdf](https://www.ramesesproject.org/media/RAMESES_II_Philosophies_and_evaluation_design.pdf)
- Griffiths S, Weston L, Morgan-Trimmer S, et al. (2022) Engaging stakeholders in realist programme theory building: Insights from the prospective phase of a primary care dementia support study. *International Journal of Qualitative Methods* 21: 160940692210775.
- HM Government (2023) Working together to safeguard children: A guide to multi-agency working to help, protect and promote the welfare of children. Available at: <https://www.gov.uk/government/publications/working-together-to-safeguard-children--2>
- Holland S, Scourfield J, O'Neill S, et al. (2005) Democratising the family and the state? The case of family group conferences in child welfare. *Journal of Social Policy* 34(1): 59–77.
- Ingram R and Smith M (2018) Relationship-based practice: Emergent themes in social work literature (Insight 41). *The Institute for Research and Innovation in Social Services (IRISS)*. Available at: <https://www.iriss.org.uk/resources/insights/relationship-based-practice-emergent-themes-social-work-literature>
- Jagosh J (2019) Realist synthesis for public health: Building an ontologically deep understanding of how programs work, for whom, and in which contexts. *Annual Review of Public Health* 40(1): 361–72.
- Jagosh J (2020) Retroductive theorizing in Pawson and Tilley's applied scientific realism. *Journal of Critical Realism* 19(2): 121–30.

- Jagosh J (2022a) Coding, configuring and conveying in realist analysis. In: *Centre for Advancement in Realist Evaluation and Synthesis (CARES)*, 3–8 March 2022. Online workshop.
- Jagosh J (2022b) Constructing excellent initial programme theories. In: *Centre for Advancement in Realist Evaluation and Synthesis (CARES)*, 10–15 February 2022. Online workshop.
- Jagosh J, Bush PL, Salsberg J, et al. (2015) A realist evaluation of community-based participatory research: Partnership synergy, trust building and related ripple effects. *BMC Public Health* 15(1): 725.
- Jagosh J, Stott H, Halls S, et al. (2022) Benefits of realist evaluation for rapidly changing health service delivery. *BMJ Open* 12(7): e060347.
- Lewis-Beck M, Bryman A and Futing Liao T (2004) *The SAGE Encyclopedia of Social Science Research Methods*. Thousand Oaks, CA: Sage.
- MacAllister J (2022) *The Independent Review of Children's Social Care: Final Report*. London: Department for Education.
- Mackley A, Jarrett T and Long R (2018) *Findings of the Care Crisis Review. Research Briefing*. London: House of Commons Library.
- Manzano A (2016) The craft of interviewing in realist evaluation. *Evaluation* 22(3): 342–60.
- Marsh P and Crow G (1997) *Family Group Conferences in Child Welfare*. Oxford: Wiley.
- Mason P, Ferguson H, Morris K, et al. (2017) *Leeds Family Valued Evaluation Report*. Children's Social Care Innovation Programme Evaluation Report 43. London: Department for Education.
- Merkel-Holguin L (2004) Sharing power with the people: Family group conferencing as a democratic experiment: Restorative justice and responsive regulation. *Journal of Sociology and Social Welfare* 31(1): 155–73.
- Mezey G, Robinson F, Campbell R, et al. (2015) Challenges to undertaking randomised trials with looked after children in social care settings. *Current Controlled Trials in Cardiovascular Medicine* 16(1): 206.
- Moody G, Brookes-Howell L, Cannings-John R, et al. (2021) What are the challenges when recruiting to a trial in children's social care? A qualitative evaluation of a trial of foster carer training. *Current Controlled Trials in Cardiovascular Medicine* 22(1): 241.
- Mukumbang FC, Van Belle S, Marchal B, et al. (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(4): e009977.
- Mukumbang FC, Van Belle S, Marchal B, et al. (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(8): e0161790.
- Munro ER, Meeto V, Quy K, et al. (2017) *Daybreak family group conferencing: Children on the edge of care*. Evaluation report, DFE-RR582. London: Department for Education.
- Ogden K and Phillips D (2024) What is the outlook for English councils' funding? Available at: <https://ifs.org.uk/publications/what-outlook-english-councils-funding#:~:text=As%20highlighted%20in%20Ogden%20and,social%20care%20services%3B%20and%20temporary>
- Pawson R and Tilley N (1997) *Realistic Evaluation*. London: Sage.
- Pennell J and Burford G (2000) Family group decision making: Protecting children and women. *Child Welfare* LXXIX(2): 131–58.
- Pennell J, Burford G, Connolly M, et al. (2011) Taking child and family rights seriously: family engagement and its evidence in child welfare. *Child Welfare* 90(4): 9–18.
- Proctor E (2017) The pursuit of quality for social work practice: Three generations and counting. *Journal of the Society for Social Work and Research* 8(3): 335–53.
- Pulman A and Fenge L-A (2023) Building capacity for social care research – Individual-level and organisational barriers facing practitioners. *The British Journal of Social Work* 53(8): 3923–42.
- Stabler L (2024a) How can Family Group Conferences be embedded as an alternative to Initial Child Protection Conferences to improve outcomes for families on the Child Protection pathway and under



- which circumstances: A realist evaluation by a research-practice partnership (Research protocol). Available at: <https://fundingawards.nihr.ac.uk/award/NIHR131922> (accessed 12 August 2024).
- Stabler L (2024b) “They Finally See Me, They Trust Me, My Brother’s Coming Home” Recognising the Motivations and Role of Siblings Who Become Kinship Carers. *Societies (Basel, Switzerland)* 14(2): 24.
- Stabler L, Evans R, Scourfield J, et al. (2022) A scoping review of system-level mechanisms to prevent children being in out-of-home care. *The British journal of social work* 52(5): 2515–2536.
- Stabler L, O’Donnell C, Forrester D, et al. (2019) *Shared decision-making: What is good practice in delivering meetings? Involving families meaningfully in decision making to keep children safely at home: A rapid realist review*. Department for Education, England What Works Centre for Children’s Social Care.
- Tavory I and Timmermans S (2014) *Abductive Analysis: Theorizing Qualitative Research*. Chicago, IL: University of Chicago Press.
- Taylor S, Blackshaw E, Lawrence H, et al. (2023) Randomised controlled trial of family group conferencing at pre-proceedings stage. *Foundations What Works Centre for Children & Families*. Available at: <https://www.coram.org.uk/resource/randomised-controlled-trial-of-family-group-conferencing-at-pre-proceedings-stage/>
- Webb C (2023) Children’s social care spending doubled in decade. *Local Government Chronicle*, 12 October. Available at: <https://www.lgcplus.com/finance/childrens-social-care-spending-doubled-in-decade-12-10-2023/>
- Webb CJR, Bennett DL and Bywaters P (2022) Austerity, poverty, and children’s services quality in England: Consequences for child welfare and public services. *Social Policy and Society: A Journal of the Social Policy Association* 23: 278–99.
- Wong G, Westhorp G, Manzano A, et al. (2016) RAMESES II reporting standards for realist evaluations. *BMC Medicine* 14(1): 96.
- Yoon S, Coxe K, Bunger A, et al. (2022) Feasibility of engaging child welfare-involved parents with substance use disorders in research: Key challenges and lessons learned. *Journal of Public Child Welfare* 16(4): 451–69.

**Bekkah Bernheim** is a Research Fellow at Exeter Medical School. She was a research-in residence throughout and acted as theory lead during the latter part of this study.

**Lorna Stabler** is a Research Fellow at CASCADE, Cardiff University, specialising in child and family social work research. She was the Principal Investigator of the study reported in this paper.

**Jennie Hayes** is a Research Fellow in the Relational Health Team at Exeter Medical School specialising in community research and health inequalities. She was theory lead during the initial part of the study.

**Alankrita Singh** is Research Assistant at CASCADE, Cardiff University and was researcher in residence for two study sites for this study.

**Katrina Wyatt** is Professor of Relational Health at Exeter Medical School and Public Health theme lead for NIHR SW PenARC. She was a Co-Investigator of the study reported in this paper.