

Combining Theories of Change and Realist Evaluation in practice: Lessons from a research on evaluation study

Author: Dr Steve Rolfe

Faculty of Social Sciences, University of Stirling, Stirling, Scotland

(+44)1786-467749

steve.rolfe1@stir.ac.uk

ORCID: 0000-0003-1465-7401

Abstract

A rapidly growing range of studies apply Theories of Change or Realist Evaluation approaches to get to grips with complex causal processes. Each methodology has been criticised in terms of practicality, usefulness and theoretical difficulties. The hypothesis that combining the two approaches could overcome some of these problems and generate deeper learning has been put forward, but there have been no published assessments of the combined methodology. This paper provides findings from an evaluation of community participation policy and practice, which specifically aimed to utilize and evaluate the application of the two approaches within one study. It suggests that there are still challenges in applying Theories of Change and Realist Evaluation approaches, but they can be practically employed together, and that this synthesis can partially overcome the critiques of each individual methodology.

Keywords

Theory-based evaluation; Theories of Change; Realist Evaluation; Community participation

La combinaison des Théories du Changement et de l'Évaluation réaliste en pratique : les leçons de recherches sur une étude d'évaluation

Résumé

Un nombre croissant d'études appliquent les approches de la Théorie du changement ou de l'Évaluation réaliste pour aborder des processus causaux complexes. Chaque méthodologie a été critiquée en matière de mise en pratique, d'utilité et de difficultés théoriques. L'hypothèse selon laquelle la combinaison des deux approches pourrait résoudre certains de ces problèmes et générer une connaissance plus approfondie a été avancée, mais aucune évaluation de cette méthodologie combinée n'a été publiée. Cet

article présente les résultats d'une évaluation des politiques et des pratiques de participation communautaire, qui visait spécifiquement à utiliser et à évaluer l'application des deux approches au sein d'une même étude. Cette étude suggère que l'application d'approches de théories du changement et d'évaluation réaliste présente toujours des difficultés mais celles-ci peuvent être, en pratique, utilisées ensemble et que cette synthèse peut surmonter en partie les critiques de chaque méthodologie.

Mots-clés

Évaluation basée sur la théorie ; Théories du changement ; Évaluation réaliste ; Participation communautaire

Introduction

Since the 'crisis' in experimental evaluation methodology encapsulated by Martinson's (1974) conclusion that 'nothing works', theory-based evaluation approaches have gained significant currency as a means of understanding the differential outcomes of social policies and programmes. In particular, theory-based approaches are seen as providing a key to unlock complex processes between policy intent and policy outcome, by examining implementation, the causal processes that generate outcomes and contextual factors that influence them (Weiss, 1998, Chen and Rossi, 1980). Thus, they aim to overcome some of the limitations of experimental designs, such as a limited explanation of exactly how/why outcomes are delivered and the decontextualisation of net effects, which undermine the potential for formative learning and policy transfer (Mackenzie et al., 2010; Davies et al., 2000).

Two of the most influential theory-based evaluation schools have been 'Theories of Change' (ToC) approaches and 'Realist Evaluation' (RE). Whilst both methodologies are being applied across diverse fields, including community change initiatives, international development, health and social work, they each face a number of critiques and challenges. In particular, ToC approaches are often seen as too descriptive, linear and non-critical, focused on implementation rather than underlying causal theory (Mackenzie and Blamey, 2005; Coryn et al., 2011; Breuer et al., 2016), whilst RE faces considerable conceptual and accessibility challenges (Porter, 2015b; Adams et al., 2016; Marchal et al., 2012), and both methodologies raise time and resource challenges (Vogel, 2012; Blamey and Mackenzie, 2007).

Partly as a response to these challenges, the notion of combining the two methodologies has been hypothesised, applying the strengths of each approach to different elements within a single evaluation (Blamey and Mackenzie, 2007; Dickinson, 2006). However, the handful of published studies which have attempted to combine ToC and RE approaches have either been unable to apply them in practice (Barnes et al., 1999; Benzeval, 2003) or blurred them together, rather than employing them in a complementary fashion (Secker et al., 2005; Macfarlane et al., 2011; Clapham et al., 2017). Moreover, whilst there are more recent evaluations using the two approaches,

there are as yet no detailed assessments of the combined methodology within the growing body of ‘Research on Evaluation’ (RoE) literature (Vallin et al., 2015, Coryn et al., 2017).

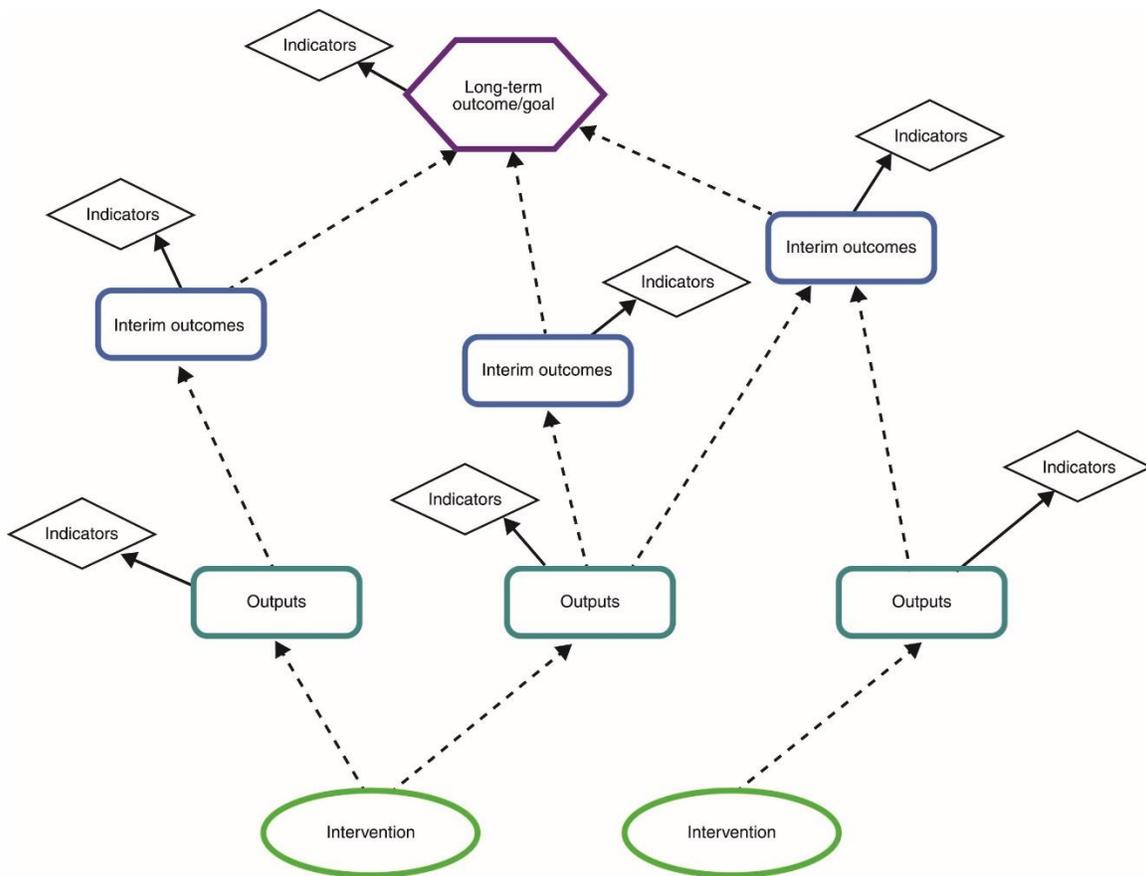
This paper attempts to contribute to the RoE literature by considering the feasibility, applicability and potential impact of the combined methodology in a study of community participation policy and practice (Miller, 2010). Whilst there are other ‘varieties’ of theory-based evaluation available, this combination was selected for this study to meet the requirements for collaborative, formative evaluation within the practice case studies (ToC) and more generalisable findings for policy evaluation (RE) (see [author], 2016a for more details). The methodological findings from this study provide some support for the hypothesis that ToC and RE approaches can be productively combined, showing that the combination is practically possible in some circumstances, and goes some way to overcoming the challenges and limitations of each individual methodology.

Overview of Theories of Change and Realist Evaluation

The Theories of Change (ToC) approach to evaluation has been primarily developed by the Aspen Institute (Connell et al., 1995; Fulbright-Anderson et al., 1998; Kubisch et al., 2002; Kubisch et al., 2010), building on earlier theory-based evaluation work (Weiss, 1998; Weiss, 1972; Chen, 1990; Chen and Rossi, 1980). The approach attempts to develop a visual, narrative model of the intervention being evaluated, setting out how inputs and activities are expected to create particular outputs and how these subsequently generate the interim and long-term outcome targets, along with success criteria at each stage to assist with attribution claims. Whilst the terminology is not used consistently by different authors, the programme’s ‘theory of change’ is generally understood as this complete model, comprised of ‘implementation theory’, which sets out how the programme should practically work, and ‘programme theory’, which consists of the ideas about how the activities should generate change (Weiss, 1998).

Essentially the approach involves four main stages (Anderson, 2005; Connell and Kubisch, 1998). Firstly, the ToC is ‘surfaced’ through collaborative work with programme staff and commissioners/funders, through a process of ‘backwards mapping’ from the ultimate goals towards inputs and activities. Whilst most ToC authors emphasise the centrality of implicit stakeholder beliefs as the key source of ‘theory’ (Weiss, 1998), others suggest that these can be complemented by existing research evidence and project documentation (Mason and Barnes, 2007). Secondly, three tests are applied to the assumptions underlying the ToC: ‘plausibility’, to see whether the posited links between inputs, activities, outputs and outcomes are logical and realistic, based on prior experience and theory; ‘doability’, to assess whether the necessary resources are available and can be employed as the model suggests; and ‘testability’, to explore the extent to which the different elements of the ToC can be practically measured. Thirdly, as the programme is implemented, a range of indicators at each stage of the ToC are measured and used to assess possible attribution. And lastly, the data regarding inputs, activities, outputs and outcomes is reviewed to assess impact, again through a collaborative process with programme stakeholders. Figure 1 provides a generic example of a ToC model.

Figure 1 – Example of Theory of Change model



Source: Adapted from Anderson (2005: 6)

At the core of the ToC approach is the idea that explicitly surfacing the complex processes involved in social interventions can get to grips with causality:

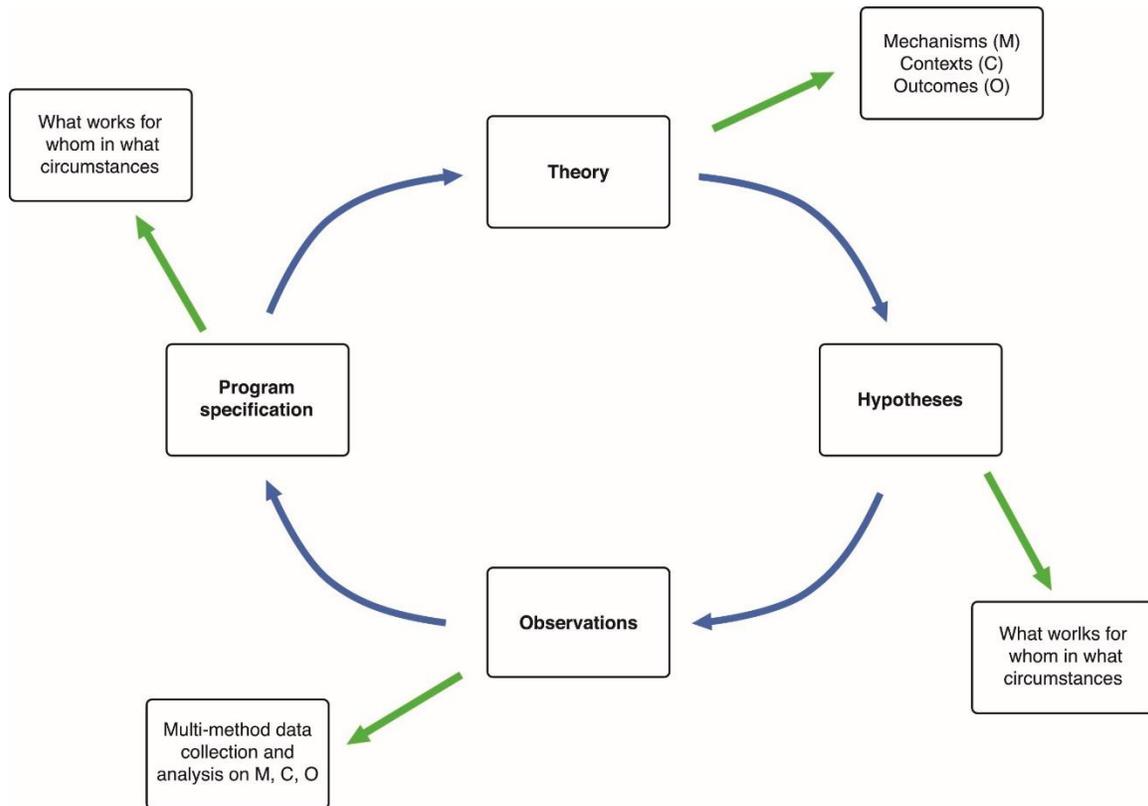
"If the evaluation can show the series of micro-steps that lead from inputs to outcomes, then causal attribution *for all practical purposes* seems to be within reach. Although such an evaluation cannot rule out all the threats to validity we have come to know and love, it has the advantage (if things go well) of showing what processes lead to the outcomes observed; if some of the posited steps are not borne out by the data, then the study can show where the expected sequence of steps breaks down." (Weiss, 2007: 70, emphasis in original)

Thus the contention is that that a ToC approach can potentially cope with long and complicated chains of causality within a programme, establishing plausible contribution claims on the basis of a range of evidence.

ToC approaches have been widely used, particularly in community change initiatives in the US (Connell et al., 1995; Fulbright-Anderson et al., 1998; Kubisch et al., 2002; Kubisch et al., 2010), international development (Vogel, 2012), and public health (Breuer et al., 2016). Looking across this experience, however, there are a number of commonly cited challenges and limitations. Firstly, despite the claim that ToC approaches help to deal with complexity, there are concerns that many ToC evaluations tend towards linear models (Mackenzie and Blamey, 2005; Dickinson, 2006), with limited detail (Breuer et al., 2016), resulting in rather descriptive studies which neglect issues of power (Coryn et al., 2011; Mackenzie and Blamey, 2005; Barnes et al., 2003). Secondly, whilst ToC approaches emphasise the importance of stakeholders' own theories, the emphasis on consensus can create difficulties in dealing with multiple perspectives (Barnes et al., 2003; Mason and Barnes, 2007). Thirdly, the collaborative nature of ToC techniques, combined with the challenges of managing complexity, can limit critical analysis, failing to question underlying assumptions and contextual issues, or to assess unexpected outcomes and causality (Breuer et al., 2016; Vogel, 2012; Coryn et al., 2011). And lastly, ToC approaches are often seen as excessively time and resource intensive, because of the requirements for collaboration and examination of entire programmes (Vogel, 2012; Mackenzie and Blamey, 2005).

Realist Evaluation (RE), developed over the last two decades by Ray Pawson and colleagues (Pawson and Tilley, 1997; Pawson, 2013; Pawson, 2006; Pawson and Manzano-Santaella, 2012), has some significant similarities to ToC approaches, particularly in focusing on underlying theory and understanding causal processes. The essential elements of RE are set out in the realist evaluation cycle (Pawson and Tilley, 1997: 85), illustrated in Figure 2.

Figure 2 – The realist evaluation cycle



Source: Pawson and Tilley (1997: 85)

As with ToC approaches, RE generally starts through discussions or ‘realist interviews’ between the evaluator and practitioners, in order to understand the nature of the programme. This is followed by the development of tentative theories, drawing on the information from practitioners together with existing empirically-supported theory from other studies across a range of fields. These hypotheses take the form of ‘Context-Mechanism-Outcome configurations’ (CMO-Cs), identifying the ways in which particular causal mechanisms may operate in particular contexts to generate particular outcomes for particular groups of people. Thus, the RE mantra is to examine, “what works for whom in which circumstances” (Pawson and Tilley, 1997: 77), often now extended to include questions of ‘how and why?’. These hypothesised CMO-Cs are then tested through mixed method data collection and realist interviews, to refine the theories and improve programme specification.

The application of RE approaches has grown significantly in recent years, primarily in health contexts (e.g. Macfarlane et al., 2011; Wong et al., 2015; Adams et al., 2016; McDermott et al., 2017), although their use is also growing in other fields, including social work (Kazi et al., 2011) and sports studies (Chen and Henry, 2016). In addition, the ‘Realist Synthesis’ approach developed by Pawson (2006) is emerging as a significant alternative to conventional meta-analysis, to combine findings across studies with

complex causal processes (e.g. Daykin et al., 2007; Silvaggi and Miraglia, 2017) and, in the form of ‘Rapid Realist Review’, as a practical approach to summarising complex evidence for policy-making purposes (Saul et al., 2013). As with ToC approaches, RE methodology has faced a number of significant critiques. Firstly, there is concern about the difficulty of conceptualising and identifying mechanisms, with many studies confusing programme interventions with the ‘underlying’ causal mechanisms (Porter, 2015a; Dalkin et al., 2015; Lacouture et al., 2015). Secondly, there are difficulties regarding the separation of contexts from mechanisms and identifying which elements of context are relevant in any given situation (Marchal et al., 2012; Dickinson, 2006; Barnes et al., 2003). Thirdly, some studies have struggled to identify appropriate theories, either because practitioners are unclear about their approach (as with ToC), or because there is insufficient or conflicting academic theory (Adams et al., 2016; Marchal et al., 2012). Fourthly, as with ToC approaches, the time and resources required for RE studies are often seen as a challenge (Blamey and Mackenzie, 2007; Marchal et al., 2012). And lastly, there is concern that the emphasis on theory and the jargon of ‘CMO-Cs’ may make findings from RE studies inaccessible and therefore unusable for practitioners (Adams et al., 2016).

Drawing the critiques together, Table 1 summarises the main challenges identified in the literature with regard to the application of ToC and RE approaches.

Table 1 – Summary of challenges associated with ToC and RE approaches

Theories of Change	Realist Evaluation
Models tend to be too linear and lacking detail, particularly around causality and success criteria	Difficult to identify and conceptualise mechanisms
Models can be too descriptive and ignore issues of power and conflicting theories	Difficult to separate contexts from mechanisms and to identify relevant contextual factors
Difficulties with multiple perspectives	Lack of substantive theory in some fields, multiple and conflicting theories in others
Limited critical analysis, ignoring unexpected outcomes	Inaccessible jargon
Time and resource intensive	Time and resource intensive

Clearly these critiques need to be considered in relation to the application of these approaches, either separately or in combination.

The hypothesis – combining ToC and RE approaches

Since both approaches have been developed in response to perceived failings in experimental evaluations, there are significant similarities between them. In particular, both approaches attempt to understand the role of context (albeit in slightly different ways), rather than attempting to exclude contextual influences through controls. As Pawson and Tilley express it:

"Our argument is that precisely what needs to be understood is what it is about given communities which will facilitate the effectiveness of a programme! And this is what is written out." (Pawson and Tilley, 1997:52, emphasis in original)

Alongside this, both approaches explicitly aim to explore processes, arguing that a narrow focus on inputs and outcomes fails to grasp the complexity of social programmes and therefore cannot provide useful attribution claims. Thus, rather than trying to experimentally restrict complexity, they deliberately 'stare it in the face' (Pawson, 2003).

However, there are also notable differences between ToC and RE approaches. Firstly, whilst ToC methodology emphasises the collaborative nature of the process, in RE approaches stakeholder views are just one source of evidence for possible theory, whilst the researcher sits atop the 'hierarchy of expertise' (Pawson and Tilley, 1997: 163-164). Secondly, the schools are epistemologically distinct, as manifested in their understandings of causality. Whilst ToC approaches are not wedded to a specific conception of causality, RE draws on the realist notion of 'generative causation', whereby interventions alter the context and resources available, triggering causal mechanism within subjects' internal capacity and reasoning (Pawson, 2013). Thirdly, the approaches differ in their approach to generalisation. For ToC approaches, it is the strength of the overall model and the level of detail, which help to identify how it may apply or differ in a new context:

"Armed with a strong theory, evaluators are better prepared to anticipate and then examine how between-site variations may shape effects." (Granger, 1998: 240)

For RE, by contrast, the power of the approach is not in the detailed understanding of entire programmes, but rather the opportunity to 'cumulate' specific CMO-Cs across different programmes, since similar processes of generative causation may apply across different policy areas (Pawson, 2006).

These differences suggest that ToC and RE tend to focus on distinct conceptions of 'theory'. Whilst ToC approaches emphasise the importance of building complete theoretical models of interventions and anticipated effects, RE approaches stress the need to 'concentrate your fire' (Pawson, 2003) and focus evaluative effort on narrower aspects of causality within social programmes and processes, "learning more and more about less and less" (Pawson and Tilley, 1997: 198).

This perspective leads to the hypothesis that they can be fruitfully combined to examine different aspects of social programmes, expressed most cogently by Blamey and Mackenzie (see also Dickinson, 2006; Pawson, 2003):

"One implication from the level at which the two approaches operate is that there is no obvious reason for believing that Theories of Change and Realistic Evaluation could not coexist within the one programme evaluation, with the former providing broad strategic learning about implementation theory and the

latter bearing down on smaller and more promising elements of embedded programme theory.” (Blamey and Mackenzie, 2007: 451)

The suggestion is that evaluations of complex social programmes might productively employ ToC techniques to assess how the programme is being delivered on a broad scale, whilst using RE approaches to examine the specific causal processes which may (or may not) be generating change in different parts of the target population. Thus the differences between the two methodologies are envisaged as complementary rather than contradictory.

Moreover, combining the two approaches could offer solutions to at least some of the critiques. Thus, RE may provide the detail that can be missing from ToC models, ensuring that programme theory is included alongside implementation theory. RE might also help to deal with multiple stakeholder perspectives and facilitate a more critical analysis, by giving primacy to the researcher as an ‘external’ adjudicator. Equally ToC approaches may provide a framework to identify, frame and prioritise contexts and mechanisms by developing a complete model of an intervention. And ToC models may also offer a more accessible means of communicating with practitioners than jargon-heavy CMO-Cs.

However combining the two approaches may fail to deliver on this promise and their differences may throw up additional challenges. The RE prioritization of the evaluator’s viewpoint may undermine the collaborative approach of ToC, exacerbating existing questions about whose ToC gains primacy (Sullivan and Stewart, 2006), whilst ToC models may prove too complex for the identification of CMO-Cs. Indeed, amongst the few published studies which explicitly combine ToC and RE approaches, the evaluation of the English Health Action Zones encountered significant difficulties in specifying ToC models due to the complexity and evolving nature of the programmes. These studies also encountered time and resource constraints (Barnes et al., 1999; Benzeval, 2003) and clearly found it difficult to identify realist mechanisms, focusing instead on interventions (Judge, 2000). Similarly, the handful of other published studies which ostensibly attempt to combine ToC and RE approaches tend to blur ‘theories of change’ and ‘mechanisms’, mixing elements of the terminology rather than rigorously applying the methodologies (Secker et al., 2005; Macfarlane et al., 2011; Clapham et al., 2017).

This paper attempts to directly address the hypothesis that ToC and RE approaches can be productively combined, through a study which aimed not merely to utilise the two methodologies, but also to evaluate the experience.

Outline of the study

The research aimed to evaluate the implementation of community participation (CP) policy in Scotland and England – Community Empowerment and Localism respectively. Six detailed case studies of community organisations were used to explore the implementation of different policy elements in varied contexts. A brief outline of the cases is provided in Table 2, whilst the process through which ToC and RE

methodologies were employed in five phases of the project is set out in Table 3 (more detail including findings from phases 1, 3 and 4 has been published elsewhere – see [author], 2016b; 2017).

As Table 3 indicates, the study used ToC and RE approaches at different stages, attempting a challenging assessment of both national policy and local practice, as well as contributing to the evidence base regarding impacts of CP. However, for the purposes of this paper, the focus is on the ways in which the two methodologies were combined, rather than the detail of how they were applied individually. The approaches were primarily used along the lines posited by Blamey and Mackenzie (2007), with ToC methodology being employed to explore and assess the overall policies and programmes, particularly in terms of implementation theory, whilst RE methodology was applied to the examination of more specific, contextualised programme theories within the broader ToC models.

Table 2 – Outline of case studies

	Organisation	Type of organisation	Socio-economic status of community	Main focus of organisation's work	Community participation process(es) examined through the research
England	A	Parish Council	Affluent	Influencing services – planning	Development of Neighbourhood Plan (one of the national 'frontrunners')
	B	Development Trust	Middling/mixed	Community self-help – facilities and services	Taking over control of community centre from local authority, following removal of grant funding.
	C	Engagement meetings organised by local authority	Disadvantaged	Influencing services – crime and grime	Groups run by local authority to enable community members to identify and address local issues
Scotland	D	Community Council	Affluent	Influencing services – planning, crime and grime	Monitoring and lobbying re planning and 'crime and grime' issues
	E	Development Trust	Middling /mixed	Community self-help – facilities and activities	Funding and installation of play park
	F	Non-profit limited company	Disadvantaged	Community self-help – wellbeing	Delivery of wellbeing service, providing psychological and alternative therapies

Table 3 – Application of ToC and RE methodologies

Phase	Application of methodologies	Outputs
1	ToC employed to structure review of literature on history and impacts of CP policy, and to analyse policy documentation for Community Empowerment and Localism.	Generic ToC model of CP policy Specific ToC models for Community Empowerment and Localism
2	Building on generic ToC model from Phase 1, RE approach used to identify the range of mechanisms which might be hypothesised to operate at different points within the model. Rapid Realist Review undertaken to examine existing evidence for selected range of mechanisms (relating to the wider social impacts of CP) and relevant contextual factors.	Map of possible mechanisms within generic ToC model Overview of evidence relating to selected mechanisms and the contextual factors that may affect their operation
3	ToC approaches used with each participant community organisation to develop local theories of change through collaborative workshops. Data collected on agreed indicators within each model and reviewed with each organisation to assess their impact and reflect on processes.	ToC models for each participant organisation Evidence of impact and process issues for each organisation
4	Data from case studies used to apply key ToC tests of plausibility and doability to policy models developed in Phase 1.	Assessment of likely impacts of national policy
5	Data from case studies used in RE analysis to develop and refine context-mechanism-outcome configurations.	Refined CMO-Cs for elements of CP policy and practice

The Research on Evaluation (RoE) aspect of the study utilised Miller's (2010) criteria for evaluating the theory-practice relationship in evaluation. This particularly focuses on the feasibility of combining ToC and RE methodologies, alongside discernible impact for different stakeholders and lessons for applicability elsewhere. By examining the evaluation context, activities and consequences (Mark, 2008), and considering potential theoretical tensions or complementarities, the study applies the RE notion of 'what works, for whom in which circumstances' to the RoE endeavour as well as the evaluation itself. Hence it attempts to produce 'contingency' theories (Christie, 2012) to indicate when the combined methodology may be useful.

The study is essentially a reflective case study (Cousins and Chouinard, 2012) aiming to provide practice insights for evaluators and contributions to the RoE evidence base through a transparent and rigorous approach. Whilst there is criticism of the limited generalisability of case study research in RoE (Vallin et al., 2015), the context-specific findings from such studies can potentially be 'cumulated' through realist synthesis¹ (Pawson, 2006). A process of continual reflection was employed, recorded in fieldnotes and a research diary by the researcher, complemented through regular discussion with project supervisors. This was augmented towards the end of the study by semi-structured interviews and focus groups with participants, focusing primarily on the usefulness of the research approach for each organisation. Given the existing critiques of each approach and the difficulties experienced by previous attempts to combine them, issues of practicality were emphasised, and whether the combination mitigated or exacerbated the separate methodological challenges. The key questions addressed through these reflective processes are summarised in Table 4. These questions were used as a framework for analysis of the data from interviews, focus groups, fieldnotes and reflective diary in Nvivo.

¹ Realist synthesis (RS) enables the 'cumulation' of evidence relating to particular mechanisms across different fields of study. Applying RS to RoE case studies would therefore involve a focus on the mechanisms that are operating in evaluation practice and the contexts within which they generate outcomes in the form of productive evaluations.

Table 4 – Key questions for methodological reflection/evaluation

Feasibility and impact

What are the implications of combining the two methodologies in terms of time and resources required?

To what extent does the addition of RE analysis help with difficulties of specifying ToC models?

To what extent does the addition of a ToC framework help with the challenges of specifying contexts, mechanisms and outcomes in RE analysis?

How accessible are the approaches for participating individuals and organisations?

How can the differences in collaboration and control within ToC and RE approaches be managed in practice?

In what ways and to what extent are ToC models developed in this way useful?

In what ways and to what extent are CMO-Cs generated by the RE analysis useful?

Theoretical issues and wider applicability

How complementary in practice are the causal models generated by the two methodologies?

How does the notion that ToC and RE approaches can be used for different levels of explanation within one evaluation work in practice?

Are there any additional challenges created by combining ToC and RE approaches?

Feasibility and impact

In relation to feasibility, the concerns regarding time and resource requirements for each methodology raise the possibility that combining them may create even greater practical demands. Indeed, the time-consuming nature of articulating and specifying ToCs, and of defining and refining CMO-Cs was undoubtedly challenging, particularly given the limited contact time available for each participant organisation:

“Made some progress with the logic model for [Organisation B] today, but I’m worried that it’s still a bit vague, particularly in terms of the intermediate stages. And I’m not going to be able to come back down [to fieldwork site] for at least another month or two.” (Fieldnotes)

As the research developed, however, the time required was reduced to some extent, specifically because of the ToC/RE combination. Most importantly, it proved unnecessary to specify models in elaborately measurable detail with clearly defined success criteria, previously identified as a particular challenge for a strictly rigorous ToC approach (Mackenzie and Blamey, 2005).

In terms of the broad analysis of CP policy and practice, the combination of ToC and RE approaches enabled the research to focus on key areas, rather than attempting to examine everything. The rapid realist review (Saul et al., 2013) of the existing evidence relating to

particular mechanisms within the generic ToC model of CP policy, enabled the local ToC work to make assumptions about some steps in the causal chains and focus data collection on areas where the existing evidence base was thin. For example, the review highlighted the breadth of evidence suggesting that communities can deliver their own services to meet needs which remain unaddressed by public services, but indicated considerable uncertainty about supportive or obstructive contextual factors. Thus, the analysis of the case study data could be focused on these contextual factors, without requiring the specificity within the local ToC models to be sure about causal attribution.

In terms of local practice, the participant organisations were content with loosely specified models, partly because their approaches were necessarily flexible and adaptive, and partly because they did not require causal certainty for their own formative learning or to make the data useful:

“[Organisation F] is like an onion – you peel off one layer and there’s always another layer underneath. And we have to constantly adapt what we do to keep meeting people’s needs. So it can be hard to make our case...but these numbers will help a lot with the Council.” (Staff member, Organisation F)

Thus, from an organisational perspective, the effort required to generate a fully specified model would have been wasted, since it would likely be measuring details which had already become irrelevant.

Therefore, whilst combining ToC and RE approaches in the manner of this study may still require a significant investment of time and resources, it seems reasonable to argue that it may generate additional benefits in return. Hence the combination goes some way to addressing Blamey and Mackenzie’s (2007: 451) warning that researchers combining these two approaches may find that the requirement for “measurement at multiple levels (the individual, group, organisation and community) [makes] the processes fraught with practical and conceptual difficulty”. In particular, employing a ToC approach to provide a framework for RE analysis, rather than as a rigorous evaluation methodology in itself, can ease the burdens of specification and data collection for the ToC work, whilst also helping to narrow the focus of the RE inquiry.

However, the evidence from this study also suggests additional challenges of accessibility and practical usefulness arising from the combined methodology.

It was apparent from relatively early in the research that there may be significant barriers to engaging in the evaluation process, due to limited organisational knowledge and skill in evaluation amongst the participant organisations:

“Difficult session with [Organisation E]. They really engaged with the idea of developing a theory of change when we last met, but looking at the draft model it was obvious that some of the committee members were pretty lost. Only the Chair and the Secretary really grasped the key points.” (Fieldnotes)

Moreover, whilst some participant organisations brought a range of skills and experience which enabled them to engage productively with the ToC methodology and models, it was equally clear that even these organisations found the complex causal language of RE off-putting.

Hence it was decided to undertake the ToC work collaboratively with the participant organisations, whilst the RE analysis was done separately by the researcher, combining participative ToC work and evaluator-led RE work in different phases. Thus the study was able to draw on the local knowledge and understanding of research participants to explore implementation theory and generate tentative programme theory, which was then refined into CMO-Cs by the researcher, retaining the possibility of a more critical analysis.

From one perspective, this approach was very productive. As an example, the development of the ToC models with Organisations A and D, benefited considerably from the knowledge of activists who understood the local context. However, only the more ‘distant’, critical RE analysis highlighted the different levels of impact arising from differences in legislative context and weaker community participation in Organisation D’s approach, which was obscured in the ToC by Organisation D’s internal narrative.

Whilst this approach maintained a productive balance between the more collaborative ethos of ToC and the emphasis on evaluator control in RE, it does raise a political and ethical tension by creating a mixed landscape of different levels of theory and analysis, echoing concerns about ‘ownership’ of ToCs (Sullivan and Stewart, 2006). Alongside this, there are practical barriers to completing the cyclical process by using the CMO-Cs with participant organisations to refine their ToC models and learn formative lessons. Aside from the inevitable length of the process of shifting from ToC to RE and back again, which in this instance meant that the final stage was not completed due to timescale constraints, there are also issues of the accessibility of RE methodology and terminology.

Theoretical issues and wider applicability

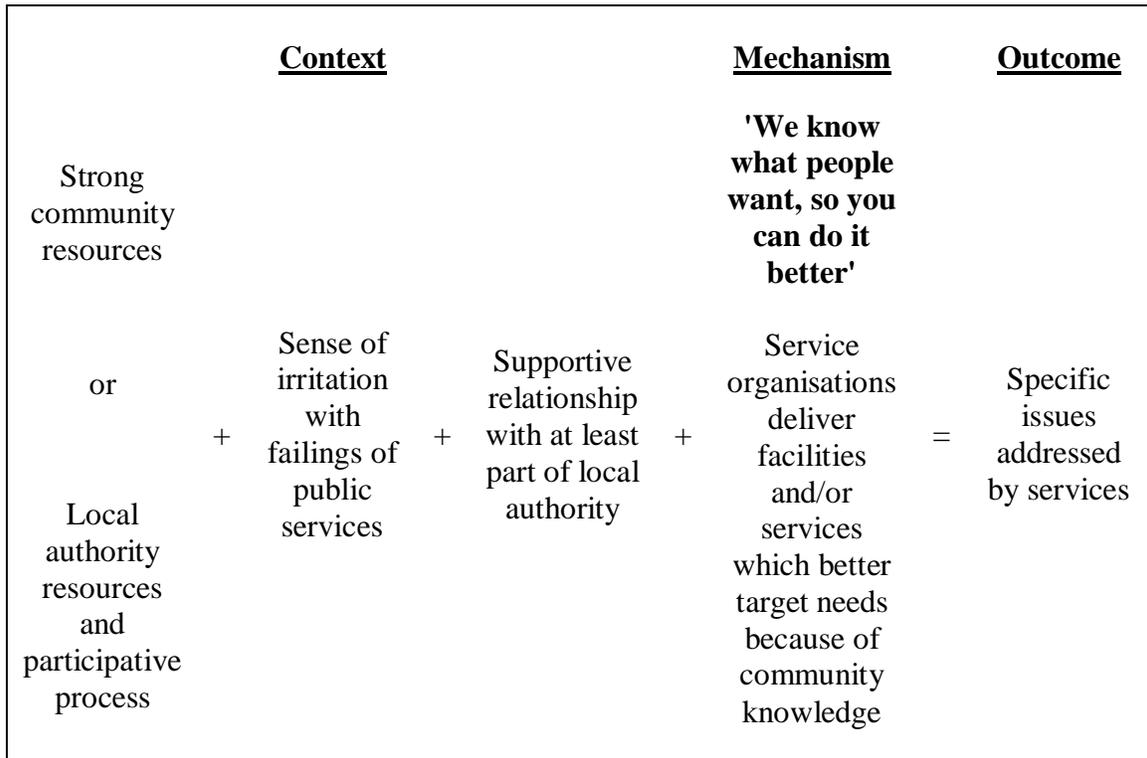
The technique of using the ToC models to provide a framework within which to identify and locate possible causal mechanisms is perhaps the most obvious advantage of combining the two approaches in the stepped fashion outlined in Table 3. As with most social policies, community participation is inherently complex, with significant non-linear elements, so an over-arching theory is necessary to identify mechanisms without becoming mired in a tangle of causal pathways (Pawson, 2013). Thus the development of a generic ToC model for CP policy provided a structure for the subsequent rapid realist review of evidence, whilst the individual ToC models similarly helped to focus data collection on the mechanisms of most interest within each case study.

Whilst the timescale of this study precluded its full realisation, this approach also has the potential of cyclical learning, whereby RE analysis of particular CMO-Cs could feed back into ToC models. As an example, Figure 3 sets out one of the key CMO-Cs

developed from the case study data. This demonstrates the contexts within which a causal mechanism identified from the earlier realist review, enables community knowledge to generate service improvement outcomes. Alongside mechanisms such as market pressures on organisations (Simmons et al., 2012), motivational effects on staff with direct connections to service users through CP (ODPM, 2005) and influencing accessibility of a service through improved community perceptions (Findlay, 2010), there is substantial evidence demonstrating the mechanism whereby organisations improve the targeting of services (Rogers and Robinson, 2004, Burton et al., 2004) or create innovative new services (ODPM, 2005) in response to improved knowledge of community needs from CP processes. The primary example of this mechanism operating in a context of strong community resources is Organisation D, whose well developed lobbying skills enabled them to influence service decisions and even to generate innovative approaches, such as the ‘Community Enforcement Initiative’ developed by the local authority in response to concerns about litter, flyposting and graffiti. Feeding this back into the ToC models for a number of the organisations could potentially help them to identify weaknesses in their approach (e.g. reliance on local authority support in a time of cuts) and consider alternatives to manage such fragilities (e.g. attempting to identify and attract additional community resources). In the case of Organisation C, these lessons could be of particular import, given that this mechanism was only triggered in a context which included facilitation by the Locality Officer.

Whilst the extant CP literature is not written in realist terms, these organisational-level service improvement mechanisms can be seen as examples of ‘interactions’ or ‘feedback processes’ in Westhorp’s (2018) typology of mechanism constructs. Hence, this evidence regarding contextual factors which influence their operation could also be useful in refining the ‘middle-range theory’ (MRT) of community participation as a means for community knowledge to improve the targeting of public services (Burton et al., 2006), offering a stronger basis for future research. Moreover, following the eclectic approach to cumulation of MRT across different types of programme proposed by Pawson (2006), such a CMO-C may provide useful evidence for evaluations of organisational change and, indeed, might be further refined itself by drawing on such literature.

Figure 3 – Example of CMO-C



However, this approach of using ToC models as a framework within which to identify key points for RE analysis is not without its challenges. In particular the number of potential mechanisms at play in causal chains which are ‘long and thickly populated’ (Pawson, 2006) can be somewhat daunting. To assist with this process of moving from broad ToC models focused on implementation theory to more specific analysis of programme mechanisms, the notion of ‘mechanism spaces’ was developed during the research. Rather than attempting to incorporate the detail of all the potential causal pathways within over-arching ToC models, the idea of a ‘mechanism space’ conceptualises the points within a policy or programme where a range of causal mechanisms may be operating. In the study, five mechanism spaces were identified within the generic ToC model of CP policy, as shown in Figure 4.

Figure 4 – Mechanism spaces within the generic ToC model of CP policy



For example, mechanism space 1 encapsulates all of the potential causal mechanisms which may be triggered by policy inputs (funding, legislation, etc.) to generate changes in community activity or community capacity by altering individual or organisational incentives and opportunities. This concept helped to focus the study, providing a structure for the realist review of existing evidence and consequently directing the RE analysis towards mechanisms within space 5, where the evidence base is weakest. Whilst a generic ToC model of this form risks over-simplification, when allied with the concept of mechanism spaces, it provides a heuristic to focus evidence reviews and empirical work.

The notion of mechanism spaces also helps to separate contexts and mechanisms, and identify which contextual factors are relevant. Although there is a danger of assuming linearity, the ToC model explicates the processes whereby outcomes of ‘earlier’ mechanisms become contexts for ‘later’ mechanisms. Thus national policy (operating in space 1) generates the contexts within which community participation processes occur (spaces 2, 3 and 4), which in turn create contexts for mechanisms which generate wider social outcomes (space 5). This perspective augments Pawson and Tilley’s (1997) suggestion that RE explanations can be understood in terms of agency and structure, reaching ‘up’ to structural levels to identify context, and ‘down’ to individual reasoning to examine mechanisms, by highlighting the value of reaching ‘back’ to previous

elements in a ToC model to understand contextual factors and the prior mechanisms through which they may have been generated.

As an example, Localism policy combined with austerity has triggered mechanisms within local authorities, generating outcomes in the shape of asset transfers to communities. This created a context for Organisation B in this study to activate a range of mechanisms within their community, such as a collective desire for autonomy and control, enabling them to take over the community centre. This in turn provided the context within which further mechanisms were triggered by new activities within the centre.

Combining ToC and RE approaches in this way also reinforces Westhorp's (2018) argument that RE theory and practice tends to focus too heavily on mechanisms operating at the level of the individual. Whilst some studies have addressed potential mechanisms operating at institutional or societal levels (e.g. Marchal et al., 2010, Prashanth et al., 2014), the emphasis placed on individual 'reasoning and resources' by Pawson and Tilley has tended to direct attention to this level of causal process. In examining community participation processes, it is apparent that some of the causal mechanisms cannot be usefully understood in terms of individual reasoning, since they are operating at the level of the organisation. Just as it is of little value to analyse complex individual behaviour by examining causal processes at the level of molecules or cells, so it may be both reductive and inefficient to examine organisational behaviour at the level of individual reasoning².

Hence, an additional benefit of undertaking RE analysis within a ToC framework is that this approach may help to identify the appropriate level at which to examine mechanisms and contexts. Drawing on Brante's (2001) notion of a 'level ontology', the suggestion here is that causality needs to be examined at different levels of sociological explanation, including the individual, institutional and structural. Considering all of the causal phases within a ToC framework enables an understanding of the different mechanisms which may be operating at different levels within complex large-scale interventions such as CP policy (Byrne, 2018). Combining this overview with the notion of mechanism spaces helps to identify when to reach 'up', 'down', 'sideways' and/or 'back' to identify relevant contextual factors and to consider which level of explanation is most useful in chains of causation.

Using the example of Organisation B once more, the generic ToC model of community participation in Figure 4 suggests that policy inputs, such as Localism and austerity, may trigger mechanisms operating at the level of the community and/or organisation (Mechanism Space 1). Thus the combination of budget cuts and support for asset transfer initiated action by Organisation B to take over and revitalise the community centre, with mechanisms operating at the level of organisational/collective decision-making:

² There is a significant parallel here with Little's notion of 'methodological localism', which sets out the value of considering meso-level causal mechanisms in response to the excesses of methodological individualism (Little, 2012)

'In order to help some of the old ones that live around here, we want to find new ways of using the centre, putting on things that will make them want to get out of the house and come here. Now we're taking over the centre, we can do things differently...if we can be proactive we can be preventative' (Organisation B Board member)

This in turn provides the context for mechanisms operating at the individual and community levels (Mechanism Space 5), as new opportunities within the centre lead to individuals engaging in new activities affecting their health, social contacts and so on. For example, the opportunity for enjoyable exercise and peer support provided by dance classes for older people may have triggered increased motivation and self-belief amongst individuals, generating health benefits.

However, whilst the ideas of mechanism spaces and level ontology may help, the essential difficulty of untangling contexts, mechanisms and outcomes within ToC models is not entirely removed, particularly given the non-linearity and complexity of community change processes (Gambone, 1998). In order to assist with this issue, contextual factors were analysed separately, to explore commonality and difference across the cases, illustrated in Table 5. Rather than presenting contexts as unified structures, as the prime movers of RE have tended to do (Pawson and Tilley, 1997; Pawson, 2013), this approach presents the context within each CMO-C as a combination of elements, for two reasons. Firstly, it assists with the comparative analysis, highlighting demi-regularities such as that shown in Table 5 (the final CMO-C from this table is illustrated in Figure 3 above). Whilst this study did not attempt to assess necessity or sufficiency amongst contextual factors, considering them in this way may facilitate such analysis, as others have suggested (Befani and Stedman-Bryce, 2017, Byrne, 2013).

Table 5 – Summary of CMO-Cs for Mechanism relating to service influence

	Community strengths	Contextual factors Relations with local authority	Motivation	Mechanism	Outcome
Organisation D	<p>Significant levels of skill, confidence and experience across the CC (and wider community)</p> <p>Substantial time commitment by (skilled, experienced) key individuals</p> <p>Strong organisational capacity, built on above</p>	<p>Support from Councillors and positive relations with some officers, alongside conflict with other officers</p>	<p>Sense of irritation with failings of public services</p>	<p>'We know what people want, so you can do it better' - Service organisations deliver facilities and/or services which better target needs because of community knowledge</p>	<p>Specific issues addressed by services</p>
Organisation C	<p>Low levels of skill, confidence and experience within NAGs (and wider community)</p> <p>Weak organisational capacity</p> <p>Substantial time commitment by Locality Officers, compensating for above</p>	<p>Support from Councillors and positive relations with Locality Officers, alongside negative views of services</p>	<p>Sense of irritation with failings of public services</p>	<p>'We know what people want, so you can do it better' - Service organisations deliver facilities and/or services which better target needs because of community knowledge</p>	<p>Specific issues addressed by services</p>

Secondly, it enables an analysis of contextual factors across different, but related mechanisms. For example, the importance of supportive relationships with at least part of the local authority (albeit often combined with conflictual elements) arose in other CMO-Cs involving different mechanisms, such as that illustrated in Figure 5.

Figure 5 – Example of CMO-C illustrating common contextual factor

<u>Context</u>			<u>Mechanism</u>	<u>Outcome</u>						
'Goldilocks zone' finance – enough money, but not too secure	+	Supportive relationship with at least part of local authority	+	Key individual(s) with professional skills to lead organisation	+	'It wouldn't happen otherwise'	+	Community organisations and communities deliver facilities and/or services that would not otherwise be delivered	=	Services or facilities delivered (plausibly leading to wider social outcomes)

In this CMO-C (which operated across three of the six case studies), supportive relationships with the local authority were part of the context within which a mechanism of community self-help was delivered, enabling community organisations to deliver their own facilities or services.

Whilst Pawson and Manzano-Santaella (2012) warn against the dangers of 'unconfigured' tables of contexts, mechanisms and outcomes, this approach of retaining contextual factors as separate elements within CMO-Cs has the potential to facilitate an important additional level of RE analysis. Alongside considering common mechanisms in the task of 'cumulating' evidence across studies (Pawson et al., 2005; Pawson, 2006), this suggests that there may also be value in 'cumulating' knowledge about the role of common contextual factors within classes of related mechanisms, potentially developing MRT in relation to contexts as well as mechanisms. This relates to the broader challenge of distinguishing between mechanisms and contexts in RE (Marchal et al., 2012, Dickinson, 2006), since it would be possible to argue that such contextual elements conceal underlying mechanisms. Indeed, it may be useful in some circumstances to analyse such contextual factors further, but in the interests of parsimonious explanation (Williams, 2018), it seems reasonable to suggest that it may be productive to cumulate knowledge about contextual influences on groups of related mechanisms.

Discussion

The evidence from this study provides significant support for Blamey and Mackenzie's (2007) hypothesis that combining ToC and RE approaches may have significant value. Thus ToC approaches can provide an analysis of programme theory within which RE approaches can be employed to examine the role of particular causal mechanisms and the contexts within which they operate to generate outcomes. Given the difficulties experienced by previous published attempts to combine the two methodologies (Barnes et al., 1999; Judge, 2000; Benzeval, 2003), this represents an important step forward in the evidence base for evaluation methodology.

Although the study focuses on relatively small projects and does not pretend to have solved all of the intractable problems of evaluating complex, non-linear processes, it has nevertheless demonstrated that it is practically possible to combine ToC and RE approaches within one evaluation. Indeed, the experience of this research indicates that the regularly expressed concerns about the time and skills required for theory-based evaluation (Blamey and Mackenzie, 2007; Brown, 1995; Sullivan, 2011) are far from insurmountable. Crucially, the combined methodology can assist with the challenge of deciding which aspects of a programme or policy to focus on within an evaluation, which can be particularly difficult in complex interventions targeting multiple outcomes at different levels, such as CP. As Gambone (1998: 150) argues, 'no research design with finite time, money, and human resources can test all the possible relationships among activities, outcomes and contexts in a community.'

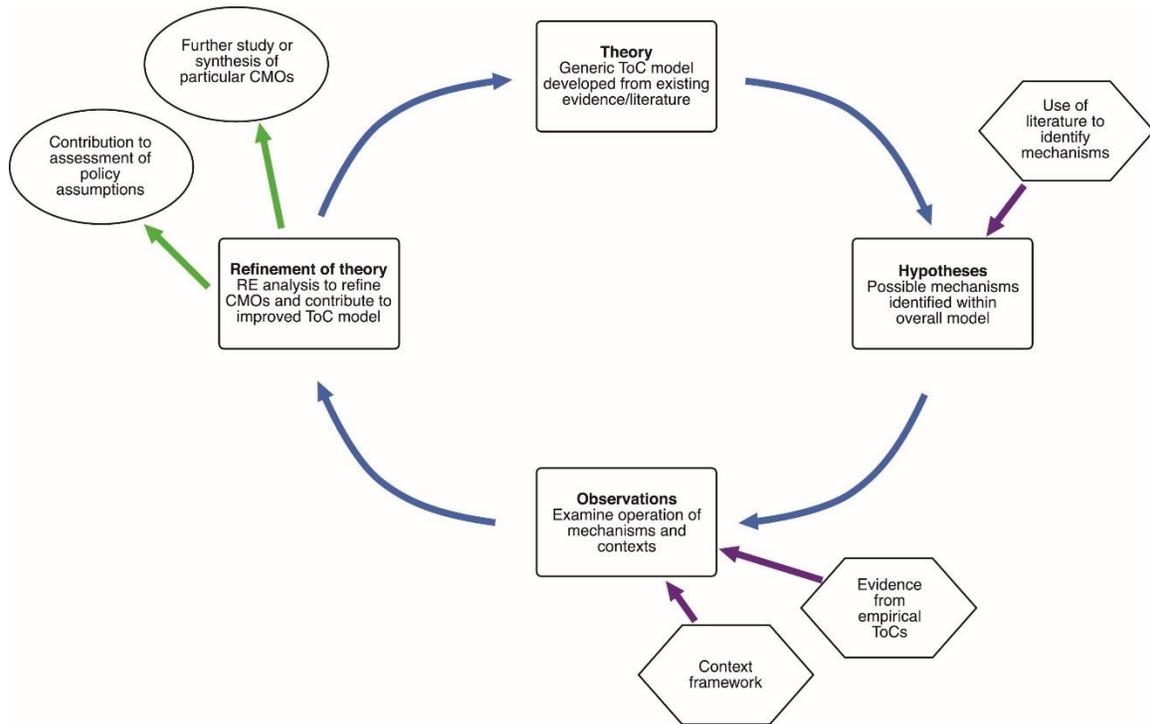
Whilst combining the two approaches may not reduce the overall time and resources required, it can help to target effort by providing a strong theoretical basis for focusing on the causal processes of most value for stakeholder organisations and the wider evidence base, potentially generating greater impact than either approach alone. Thus, applying Miller's (2010) criteria for evaluating the theory-practice relationship in evaluation, the study provides evidence that the ToC-RE combination is feasible in at least some circumstances and has the potential to create significant impact for a range of stakeholders, including participating organisations, researchers and evaluation practitioners. The ToC approach within each case study provided findings of value to participating organisations, even though the restricted form of ToC limits the possibility of generalizing from these individual models. Alongside this, the RE findings open the door to potential generalizability through cumulation, perhaps particularly in combination with the generic ToC model.

The addition of RE analysis to ToC approaches also has the potential to address at least some of the criticisms levelled at ToC methodology (see Table 1). The application of RE within the broader ToC framework clearly helps to tackle the concern that ToC approaches tend to focus on relatively descriptive models of implementation theory (Coryn et al., 2011), since the RE analysis delves deeper into the underlying causal mechanisms. And, by retaining a non-collaborative element in the RE analysis, the combined approach can offer a critical perspective which helps to manage issues of unquestioning consensus (Mason and Barnes, 2007; Breuer et al., 2016; Vogel, 2012).

Furthermore, the combined approach in this study has also developed three related innovations, in order to tackle some of the challenges of RE methodology, particularly in terms of the difficulty of identifying relevant contextual factors (Marchal et al., 2012) and of separating contexts, mechanisms and outcomes in complex, open systems (Barnes et al., 2003, Byrne, 2018). Firstly, the concept of ‘mechanism spaces’ within ToC models provides a structure for realist review and the identification of relevant MRT regarding potential mechanisms within a programme or policy. Secondly, combining this with an understanding of ‘level ontology’ enables a focus on the different levels at which mechanisms may operate and the types of contextual factors which may trigger or shape their operation. And lastly, by analyzing contextual factors as well as mechanisms across cases and studies, the door is opened to an expansion of realist synthesis methodology (Pawson, 2006), potentially developing or refining MRT around contexts as well as mechanisms. Whilst all of these innovations require further research to examine their utility across different settings, the evidence from this study suggests that they are of value in fruitfully combining ToC and RE approaches.

It seems plausible, therefore, to refine the RE cycle (Figure 2) into a ‘combined evaluation cycle’, whereby exploration and refinement of particular CMO-Cs can be used to populate and elaborate a broader ToC, leading in turn to the identification of further mechanisms and contextual factors for detailed study, as illustrated in Figure 6. As with RE methodology in general, this includes the possibility of developing MRT, as CMO-Cs are cumulated beyond the study (in the top left of the diagram) and also utilising such theory as an input to elaborate the ToC model through review of the existing literature (in the top right). Thus the limited generalisability of ToC models is countered by the opportunity for cumulation of RE findings emerging from each study, which also helps to provide a (middle-range) theoretical basis to underpin the models themselves.

Figure 6 – Combined ToC/RE evaluation cycle



Clearly further work will be necessary to test the value of this model. In particular, this study was not able to complete the cycle by taking the RE analysis back to participant organisations in order to help them refine their ToC models. And whilst Marchal et al. (2012) suggest that the inclusion of processes and contexts can enhance the attractiveness of RE findings for policy makers, there is as yet little evidence with regard their potential value for practitioners and activists. Hence, further work would be necessary to explore the practical utility of CMO-Cs identified through this approach.

Moreover, it would be of significant value if further studies combining ToC and RE approaches in different fields and in different ways were published, particularly if they incorporate at least some element of RoE in order to assess the combination’s value for evaluation methodology more broadly. Within such further studies it may also be productive to incorporate and assess some of the additional methodological approaches which aim to firm up contribution claims and understanding of necessary and sufficient contexts, such as process tracing through Bayesian analysis or Qualitative Comparative Analysis (Befani and Stedman-Bryce, 2017, Byrne, 2013). Indeed, until further studies of the ToC/RE combination are published, it is difficult to assess wider applicability (Miller, 2010) in terms of the types of social programme that may be best evaluated through this dual methodology.

Nevertheless, the innovations outlined above and the combined ToC/RE evaluation cycle do provide some degree of ‘contingency theory’ (Christie, 2012) to facilitate the

application of the combined methodology across different fields. In particular, it is possible to argue that the experience of examining both policy and practice in this study says something valuable about the evaluative purposes to which the combined methodology is most suited. The suggestion here is that the two approaches can be utilised in slightly different ways to serve different purposes. Thus, ToC approaches can provide an ‘agile heuristic’ (Funnell and Rogers, 2011: 79) to assist organisations to evidence impacts of adaptive, flexible action, which can be combined with elements of RE analysis to provide learning which can be more readily generalised across different situations. Alongside this, the combination of the two methodologies can provide a valuable lens through which to examine policy intentions and to compare these aims with evidence from practice on the ground, with potential value for policy makers.

Crucially, this suggests that the approach is best suited to evaluations which are targeted at formative learning, rather than dogmatic summation of success or failure. Thus the aim is to continually refine both the organisation-specific ToCs and the more generalisable CMO-Cs (Mason and Barnes, 2007; Funnell and Rogers, 2011: 517). The combined methodology may be less suited to evaluations where there is a requirement for a more complete causal understanding of a policy or programme, or conversely, where there is a greater interest in outcomes than in unpacking causality.

Indeed, exploring the types of evaluations and situations where the ToC-RE combination may be of most value can perhaps be facilitated by extending the RE mantra of ‘what works, for whom in what circumstances’ to the choice of evaluation methodology. In other words, selecting an evaluation approach requires consideration of the purpose, audience and context. This study does not pretend to suggest that combining ToC and RE approaches will work everywhere for everyone, but it can work productively to evaluate both policy and practice for a range of potential audiences, including community organisations, policy makers and researchers themselves. Moreover, by applying RoE principles to examine the feasibility, impact and applicability of the combined methodology, the study provides a basis for further RoE assessments of this and other combinations of evaluation methodologies to complex, multi-level policies and programmes.

Declaration of interest

The research on which this article is based was funded by the Urban Studies Foundation.

Acknowledgements

The author would like to acknowledge the support provided by Professor Mhairi Mackenzie and Professor Annette Hastings during and after the research. The author would also like to acknowledge the enormous contribution of the community activists who participated in the research, generously giving their time and expertise.

References

- Adams A, Sedalia S, McNab S and Sarker M. (2016) Lessons learned in using realist evaluation to assess maternal and newborn health programming in rural Bangladesh. *Health Policy & Planning* 31: 267-275.
- Anderson A. (2005) *The community builder's approach to Theory of Change: A practical guide to theory development*. Washington, DC: Aspen Institute.
- Barnes M, Bauld L, Benzeval M, Judge K, Killoran A, Robinson, R et al. (1999) *Health Action Zones: Learning to make a difference*. Canterbury: University of Kent.
- Barnes M, Matka E and Sullivan H. (2003) Evidence, understanding and complexity: Evaluation in non-linear systems. *Evaluation* 9: 265-284.
- Befani B and Stedman-Bryce G. (2017) Process Tracing and Bayesian Updating for impact evaluation. *Evaluation*, 23: 42-60.
- Benzeval M. (2003) *National evaluation of Health Action Zones: The final report of the tackling inequalities in health module*. London: Queen Mary, University of London.
- Blamey A and Mackenzie M. (2007) Theories of change and Realistic Evaluation: Peas in a pod or apples and oranges? *Evaluation* 13: 439-455.
- Brante T. (2001) Consequences of realism for sociological theory-building. *Journal for the Theory of Social Behaviour* 31: 167-195.
- Breuer E, Lee L, de Silva M and Lund C. (2016) Using theory of change to design and evaluate public health interventions: a systematic review. *Implementation Science* 11.
- Brown P. (1995) The role of the evaluator in comprehensive community initiatives. In: Connell J, Kubisch A, Schorr L, et al. (eds) *New approaches to evaluating community initiatives: Concepts, methods and contexts*. Washington, DC: Aspen Institute.
- Burton P, Goodlad R and Croft J. (2006) How would we know what works? Context and complexity in the evaluation of community involvement. *Evaluation*, 12: 294-312.

- Burton P, Goodlad R, Croft J, Abbott J, Hastings A, Macdonald G et al. (2004) What works in community involvement in area-based initiatives? A systematic review of the literature. London: Home Office.
- Byrne D. (2013) Evaluating complex social interventions in a complex world. *Evaluation*, 19: 217-228.
- Byrne D. (2018) Researching complex large-scale nested interventions. In: Emmel, N., Greenhalgh, J., Manzano, A., Monaghan, M. and Dalkin, S. (eds.) *Doing realist research*. London: Sage.
- Chen H-T. (1990) *Theory-driven evaluations*, Newbury Park, CA: Sage.
- Chen H-T and Rossi PH. (1980) The Multi-Goal, Theory-Driven Approach to Evaluation: A Model Linking Basic and Applied Social Science. *Social Forces* 59: 106-122.
- Chen S and Henry I. (2016) Evaluating the London 2012 Games' impact on sport participation in a non-hosting region: a practical application of realist evaluation. *Leisure Studies* 35: 685-707.
- Christie C. (2012) Advancing empirical scholarship to further develop evaluation theory and practice. *Canadian Journal of Program Evaluation*, 26: 1-18.
- Clapham K, Manning C, Williams K, O'Brien G and Sutherland M. (2017) Using a logic model to evaluate the Kids Together early education inclusion program for children with disabilities and additional needs. *Evaluation and Program Planning* 61: 96-105.
- Connell J and Kubisch A. (1998) Applying a theory of change approach to the evaluation of comprehensive community initiatives: Progress, prospects and problems. In: Fulbright-Anderson K, Kubisch A and Connell J (eds) *New approaches to evaluating community initiatives: Volume 2 - Theory, measurement, and analysis*. Washington, DC: Aspen Institute, 15-44.
- Connell J, Kubisch A, Schorr L and Weiss C. (1995) *New approaches to evaluating community initiatives*. Washington, DC: The Aspen Institute.

- Coryn C, Noakes L, Westine C and Schroter D. (2011) A systematic review of theory-driven evaluation practice from 1990 to 2009. *American Journal of Evaluation* 32: 199-226.
- Coryn C, Wilson L, Westine C, Hobson K, Ozeki S, Fiekowsky E. et al (2017) A Decade of Research on Evaluation: A Systematic Review of Research on Evaluation Published Between 2005 and 2014. *American Journal of Evaluation*, 38: 329-347.
- Cousins J and Chouinard J. (2012) *Participatory evaluation up close: An integration of research-based knowledge*, Charlotte, North Carolina, Information Age Publishing.
- Dalkin S, Greenhalgh J, Jones D, Cunningham B and Lhussier M. (2015) What's in a mechanism? Development of a key concept in realist evaluation. *Implementation Science* 10: 49.
- Davies H, Nutley S and Tilley N. (2000) Debates on the role of experimentation. In: Davies H, Nutley S and Smith P (eds) *What works? Evidence-based policy and practice in public services*. Bristol: Policy Press.
- Daykin N, Evans D, Petsoulas C and Sayers A. (2007) Evaluating the impact of patient and public involvement initiatives on UK health services: a systematic review. *Evidence & Policy: A Journal of Research, Debate & Practice* 3: 47-65.
- Dickinson H. (2006) The evaluation of health and social care partnerships: an analysis of approaches and synthesis for the future. *Health & Social Care in the Community* 14: 375-383.
- Findlay E. (2010) Review of the impact of community engagement within regeneration. Edinburgh: Scottish Government.
- Fulbright-Anderson K, Kubisch A and Connell J. (1998) New approaches to evaluating community initiatives: Volume 2 - Theory, measurement, and analysis. Washington, DC: Aspen Institute.
- Funnell S and Rogers P. (2011) *Purposeful program theory: Effective use of theories of change and logic models*, San Francisco: Jossey Bass.

- Gambone M. (1998) Challenges of measurement in community change initiatives. In: Fulbright-Anderson K, Kubisch A and Connell J (eds) *New approaches to evaluating community initiatives: Volume 2 - Theory, measurement and analysis*. Washington, DC: Aspen Institute, 149-164.
- Granger R. (1998) Establishing causality in evaluations of comprehensive community initiatives. In: Fulbright-Anderson K, Kubisch A and Connell J (eds) *New approaches to evaluating community initiatives: Volume 2 - Theory, measurement and analysis*. Washington, DC: Aspen Institute, 221-246.
- Judge K. (2000) Testing Evaluation to the Limits: The Case of English Health Action Zones. *Journal of Health Services Research & Policy* 5: 3-5.
- Kazi M, Pagkos B and Milch H. (2011) Realist Evaluation in Wraparound: A New Approach in Social Work Evidence-Based Practice. *Research on Social Work Practice* 21: 57-64.
- Kubisch A, Auspos P, Brown P, et al. (2002) Voices from the field II - Reflections on comprehensive community change. Washington, DC: Aspen Institute.
- Kubisch A, Auspos P, Brown P, Chaskin R, Fulbright-Anderson K and Hamilton R. (2010) Voices from the field III - Lesson and challenges from two decades of community change efforts. Washington, DC: Aspen Institute.
- Lacouture A, Breton E, Guichard A and Ridde V. (2015) The concept of mechanism from a realist approach: a scoping review to facilitate its operationalization in public health program evaluation. *Implementation Science : IS* 10: 153.
- Little, D. (2012) Analytical sociology and the rest of sociology. *Sociologica* 2012: 1
- Macfarlane F, Greenhalgh T, Humphrey C, Hughes J, Butler C and Pawson R. (2011) A new workforce in the making?: A case study of strategic human resource management in a whole-system change effort in healthcare. *Journal of Health, Organisation and Management* 25: 55-72.
- Mackenzie M and Blamey A. (2005) The Practice and the Theory: Lessons from the Application of a Theories of Change Approach. *Evaluation* 11: 151-168.

- Mackenzie M, O'Donnell C, Halliday E, Sridharan S and Platt S. (2010) Do health improvement programmes fit with MRC guidance on evaluating complex interventions? *BMJ* 340.
- Marchal B, Dedzo M and 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.
- Marchal B, van Belle S, van Olmen J, Hoeree T and 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.
- Mark M. (2008) Building a better evidence base for evaluation theory: Beyond general calls to a framework of types of research on evaluation. In: Smith, N. and Brandon, P. (eds.) *Fundamental issues in evaluation*. New York: Guilford.
- Martinson R. (1974) What works? Questions and answers about prison reform. *Public Interest* 35: 22-45.
- Mason P and Barnes M. (2007) Constructing theories of change: Methods and sources. *Evaluation* 13: 151-170.
- McDermott I, Checkland K, Coleman A, Osipovi D, Petsoulas C and Perkins N. (2017) Engaging GPs in commissioning: realist evaluation of the early experiences of Clinical Commissioning Groups in the English NHS. *Journal of Health Services Research & Policy* 22: 4-11.
- Miller R. (2010) Developing Standards for Empirical Examinations of Evaluation Theory. *American Journal of Evaluation*, 31: 390-399.
- ODPM. (2005) Improving delivery of mainstream services in deprived areas - the role of community involvement. London: Office of the Deputy Prime Minister.
- Pawson R. (2003) Nothing as Practical as a Good Theory. *Evaluation* 9: 471-490.
- Pawson R. (2006) *Evidence-based policy: A realist perspective*, London: Sage.
- Pawson R. (2013) *The science of evaluation: A realist manifesto*, London: Sage.

- Pawson R, Greenhalgh T, Harvey G and Walshe K. (2005) Realist review - a new method of systematic review for complex policy interventions. *Journal of Health Service Research Policy* 10: 21-34.
- Pawson R and Manzano-Santaella A. (2012) A realist diagnostic workshop. *Evaluation* 18: 176-191.
- Pawson R and Tilley N. (1997) *Realistic evaluation*, London: Sage.
- Porter S. (2015a) Realist evaluation: an immanent critique. *Nursing Philosophy* 16: 239-251.
- Porter S. (2015b) The uncritical realism of realist evaluation. *Evaluation* 21: 65-82.
- Prashanth N, Marchal B, Devadasan N, Kegels G and 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.
- Rogers B and Robinson E. (2004) The benefits of community engagement: A review of the evidence. London: Home Office/IPPR.
- Saul J, Willis C, Bitz J and Best A. (2013) A time-responsive tool for informing policy making: rapid realist review. *Implementation Science* 8: 1-15.
- Secker J, Bowers H, Webb D and Llanes M. (2005) Theories of change: What works in improving health in mid-life? *Health Education Research* 20: 392-401.
- Silvaggi F and Miraglia M. (2017) Mental Health at Work: A Review of Interventions in Organizations. *E-Journal of International & Comparative Labour Studies (EJICLS)* 6: 1-27.
- Simmons R, Birchall J and Prout A. (2012) User involvement in public services: “choice about voice.” *Public Policy and Administration*, 27: 3-29.
- Sullivan H, and Stewart M. (2006) Who Owns the Theory of Change? *Evaluation*, 12: 179-199.
- Sullivan H. (2011) 'Truth' junkies: using evaluation in UK public policy. *Policy & Politics* 39: 499-512.

- Vallin L, Philippoff J, Pierce S and Brandon P. (2015) Research-on-Evaluation Articles Published in the American Journal of Evaluation, 1998–2014. *New Directions for Evaluation*, 2015: 7-15.
- Vogel I. (2012) Review of the use of ‘Theory of Change’ in international development. London: Department for International Development.
- Weiss C. (1972) *Evaluation Research: Methods of Assessing Program Effectiveness*, New Jersey: Prentice Hall.
- Weiss C. (1998) *Evaluation*, New Jersey: Prentice Hall.
- Weiss C. (2007) Theory-based evaluation: Past, present and future. *New Directions for Evaluation* 2007: 68-81.
- Westthorp G. (2018) Understanding mechanisms in realist evaluation and research. *In*: Emmel, N., Greenhalgh, J., Manzano, A., Monaghan, M. and Dalkin, S. (eds.) *Doing realist research*. London: Sage.
- Williams M. (2018) Making up mechanisms in realist research. *In*: Emmel, N., Greenhalgh, J., Manzano, A., Monaghan, M. and Dalkin, S. (eds.) *Doing realist research*. London: Sage.
- Wong E, Yeoh E-k, Chau P, Yam C, Cheung A and Fung H. (2015) How shall we examine and learn about public-private partnerships (PPPs) in the health sector? Realist evaluation of PPPs in Hong Kong. *Social Science & Medicine* 147: 261-269.