

Research Synthesis Methods, 2014, 5(2): 131-41. DOI: 10.1002/jrsm.1099

Critical reflections on realist review: insights from customizing the methodology to the needs of participatory research assessment

Justin Jagosh,^a Pierre Pluye,^a Geoff Wong,^b Margaret Cargo,^c Jon Salsberg,^a Paula L. Bush,^a Carol P. Herbert,^d Lawrence W. Green,^e Trish Greenhalgh^b and Ann C. Macaulay^a

- a. Participatory Research at McGill, Department of Family Medicine, McGill University, Montréal, Canada
- b. Centre for Primary Care and Public Health, Queen Mary, University of London, London, UK
- c. School of Population Health, University of South Australia, Adelaide, Australia
- d. Centre for Studies in Family Medicine, Schulich School of Medicine & Dentistry, The University of Western Ontario, London, Canada
- e. Department of Epidemiology and Biostatistics, University of California at San Francisco, San Francisco, USA

ABSTRACT

Realist review has increased in popularity as a methodology for complex intervention assessment. Our experience suggests that the process of designing a realist review requires its customization to areas under investigation. To elaborate on this idea, we first describe the logic underpinning realist review and then present critical reflections on our application experience, organized in seven areas. These are the following: (1) the challenge of identifying middle range theory; (2) addressing heterogeneity and lack of conceptual clarity; (3) the challenge of appraising the quality of complex evidence; (4) the relevance of capturing unintended outcomes; (5) understanding the process of context, mechanism, and outcome (CMO) configuring; (6) incorporating middle-range theory in the CMO configuration process; and (7) using middle range theory to advance the conceptualization of outcomes – both visible and seemingly ‘hidden’. One conclusion from our experience is that the degree of heterogeneity of the evidence base will determine whether theory can drive the development of review protocols from the outset, or will follow only after an intense period of data immersion. We hope that presenting a critical reflection on customizing realist review will convey how the methodology can be tailored to the often complex and idiosyncratic features of health research, leading to innovative evidence syntheses.

Keywords: realist review; middle-range theory; context-mechanism-outcome configuring; participatory research; critical realism; health research

1. INTRODUCTION

Realist review methodology is increasingly being used to study health and medical interventions to support the synthesis of complex assessments. It is a theory-driven, interpretative approach to configuring contextual factors and mechanisms of change related to outcomes. Seminal texts and published reporting standards for realist review provide valuable guiding principles for the inclusion of key elements (Pawson, 2006b; Greenhalgh et al., 2011; Wong et al., 2013), yet the underlying logic of the approach makes it antithetical to standardized, predetermined or prescriptive application. We argue that in applying the logic of critical realism, the customization of protocols may be necessary to account for and incorporate the idiosyncrasies of a particular evidence base. Examples of such idiosyncrasy may include the following: (a) a dearth or an overabundance of evidence; (b) the type of available evidence (qualitative, quantitative, etc.); (c) the availability of evidence describing context and process; and (d) the characteristics of the evidence in relation to theory. Customization is part of a growing trend in systematic reviews, for cases in which a flexible and iterative syntheses process has value (particularly in qualitative research). Here, we clarify that the value of customization for a review type is dependent on the nature of the synthesis process involved (i.e., quantitative, qualitative, or mixed methods data analysis). Customization may not be relevant or productive in a review that is based on a logic of data aggregation, in which case procedural uniformity may have value. However, in the case of realist review, we suggest that suitably customizing the realist approach to the characteristics of a body of evidence can optimize review outcomes that can improve knowledge translation and facilitate decision-making.

1.1 Purpose and Aim

The purpose of this paper is to provide an in-depth consideration of the methodological challenges we faced in conducting a realist review and how this led to a customization of our protocols. As such, this is not an introductory paper on realist review; for that we recommend other sources (Pawson et al., 2005; Pawson, 2006b; Rycroft-Malone et al., 2012; www.ramesesproject.org). Rather, our aim is to contribute to the methodological literature that describes how the logic of applied critical realism can serve the diversity of research areas in the health sciences. To give some context to this reflection, we first provide a few background sections about the subject area of the example review presented here and the logic underpinning realist review.

1.2 Overview of Participatory Research

The purpose of the review we conducted was to better understand and synthesize the benefits (if any) of participatory research. The review has been published in four separate articles describing the rationale for applying realist review methodology (Macaulay et al., 2011), description of the protocol (Jagosh et al., 2011), and two papers reporting on the findings (Jagosh et al., 2012; Macaulay et al., In Press). Participatory research (PR) refers to the co-construction of research through partnerships between researchers and people affected by, and/or responsible for action on, issues under study (Green et al., 1995; Macaulay et al., 1999). PR is not a methodology, but rather a research approach, in which multistakeholder partnerships (i.e., academic + community, or other non-academic stakeholders such as professional practitioners, organizations, or policy makers) plan and implement quantitative, qualitative, or mixed methods research to generate health promotion/disease prevention interventions, products, and knowledge (Cargo and Mercer, 2008). The aim of our review was to conceptualize, assess, and provide explanatory insight into the types, extent, and quality of outcomes that arise from PR partnerships for health research.

At the time we conducted this synthesis, (2009–2011), few examples of realist reviews had been published. Most examples looked at sets of similar or related interventions working on similar program goals to better understand the causes of their successes and challenges (Kristjansson et al., 2007; Wong et al., 2010). Similar to our review, these reviews handled significant complexity and heterogeneity in their respective research areas. Our area had an added dimension of complexity in that rather than reviewing the success of an intervention goal, we set out to study the impact of the PR approach on programs of research. This characteristic of our area of inquiry presented challenges in constructing the review in a way that was similar to other published prototypes and examples.

1.3 Overview of the Research Team and Protocol

The research team comprised a core group of investigators at McGill University: A. C. Macaulay, J. Jagosh, P. Pluye, J. Salsberg, J. Henderson, P. L. Bush, and E. Sirett; an international network of researchers: G. Wong, M. Cargo, T. Greenhalgh, C. P. Herbert, and L. W. Green; and six decision-making partners from granting agencies, public health organizations, an ethics board, and a community-engaged scholarship organization. Details regarding the protocol and our partners can be found in a separate publication (Jagosh et al., 2011). Briefly, a librarian-guided search strategy yielded an initial 7167 abstracts from the literature. The literature was then filtered through tools that were inductively developed for the identification, selection, and appraisal stages. Although not an essential component of realist protocol, we conducted these stages in a stepwise (linear) fashion, to manage and filter through

the vast number of publications caught in the literature search. We encountered self-identified applications of PR that framed the partnership concept in ways that were inconsistent with community co-decision-making or equitable co-governance with academic research. Due to the presence of this heterogeneity in our dataset, we iteratively refined the selection tool and refined our conceptualization of PR (explained in further discussion). Consequently, the selection tool was revised and re-applied to all the included papers in the selection phase. At each stage, two members of the research team independently reviewed and coded articles, and brought discrepant codes to the wider team for resolution. At the selection stage, 83 full-text publications were retained and one author from each of the papers was contacted to ensure that we had a comprehensive set of documents for each study. Thirty-two (39%) of these authors responded to complete the sets for their respective partnerships, and only those sets were retained for synthesis. After passing these 32 sets publications through an appraisal process, a final sample of 23 sets of documents ($n = 276$) remained – each set pertaining to an academic–community partnership that used a participatory approach for health intervention research. The documents included peer-reviewed and non-peer-reviewed publications, websites, and program activity descriptions. The partnerships included were those describing ‘high’ levels of participation by non-academic stakeholders. This meant that we retained the literature of only those partnerships described as having all partners involved in the following: (a) identifying or setting the research questions; (b) setting the methodology, collecting data, or analyzing the data; and (c) using or disseminating the research findings. Coming to this set of criteria required considerable team discussion and was an important aspect defining the parameters of our review. Please refer to our protocol study paper for greater detail about our identification, selection, and appraisal tools (Jagosh et al., 2011).

1.4 Main Findings

Our findings revealed PR outcomes across seven domains. They were the following: that PR supports (a) the production of culturally and logistically appropriate research; (b) the capacity to recruit participants to projects and interventions; (c) the capacity building of all participating stakeholders; (d) productive conflict resolution and negotiation processes; (e) the accumulation of partnership synergy, which increases the quality of outputs and outcomes over time; (f) the capacity to sustain project goals beyond funded time frames and during gaps in external funding; and (g) the generation of systemic changes and new unanticipated projects and activity. A detailed account of these findings can be found elsewhere (Jagosh et al., 2012).

2. THE LOGIC UNDERPINNING REALIST REVIEW

The following section is a brief overview of realist logic, method, and concepts, which need description to understand why the approach supports the customization of protocols to particular areas of investigation. Realist review has developed from the philosophical traditions of critical realism (Archer et al., 1998; Sayer, 2000) which question and address the perceived or assumed interrelationship between science, knowledge, and reality (Bhaskar, 1998a; 1998b) both by the scientific community and society at large. Describing the underlying philosophical cannon of critical realism is beyond the scope of this paper. However, a central distinguishing tenet of critical realism is that there are both visible and hidden forces at play in the generation of outcomes and that research in complex areas needs to be receptive to both in order to theorize upon and unearth root causal factors. From this tenet, realist review can move beyond the internal validity or efficacy question of ‘did an intervention work or not?’ to the effectiveness or external validity questions of ‘what works, for whom and in what circumstances, how and why?’ (Pawson, 2006b). The assumption driving these questions is that complex programs and interventions offer resources (e.g., education, information, and disciplinary actions) to a target population and are built on implicit or explicit assumptions that such efforts will lead to positive change in or for that population. The process of assessment involves uncovering underlying middle-range theory (or logic) driving interventions and their multiple components, as well as illuminating the contextual factors that influence mechanisms of change to produce outcomes. Whereas the empirical testing approach of post-positivism has researchers interpreting their results after the research analysis has concluded, the realist approach involves an ongoing interpretative (abductive) process to configure context, mechanism, and outcome (CMO). It is through this process that seemingly hidden outcomes can be synthesized. Such interpretative activity may also involve bringing in hunches and ideas that are not necessarily linked to the researchers’ immediate view of the evidence. Table 1 provides a detailed definition of terms.

The remainder of this paper describes the challenges and issues encountered in following the steps for conducting a realist review as outlined by Pawson and offers suggestions for customizing review protocols for future realist reviews. We describe here the following: (a) the challenge of identifying suitable middle-range theory (MRT), (b) addressing heterogeneity and lack of conceptual clarity in the literature; (c) the challenge of appraising the quality of complex evidence; (d) the relevance of capturing unintended outcomes; (e) understanding the process of CMO configuring; (f) incorporating middle-range theory in the CMO configuration process; (g) using MRT to advance the conceptualization of outcomes – both visible and seemingly ‘hidden’

2.1 The challenge of identifying a suitable of middle-range theory

Our initial database search (Jagosh et al., 2011) yielded a large body of literature, which captured

a diversity of research designs, health topics, definitions, and concepts used in PR. Seminal texts on realist review emphasize the importance of identifying MRT at the outset in tandem with establishing the review questions. On-going attempts at MRT identification were made at the beginning, yet we found the degree of diversity captured in our search precluded the possibility of identifying suitable MRTs at the outset. Although occasionally revisiting the question of MRT identification, our attention was focused on developing an identification, selection, and appraisal protocol to sort and manage the literature, and to narrow the evidence base to a manageable set of partnerships. Not initiating and shaping the review with MRT appeared to be a departure from the suggested approach found in seminal realist review texts. Yet, this did not seem to be in contradiction with realist logic (as we argue in the sections later). Table 2 shows the goals and types of document managed at each stage of the review.

2.2 Addressing heterogeneity and lack of conceptual clarity in the evidence base

By immersing in literature from the steps outlined in table two, and discussing disputed cases, opportunities arose for conceptual and methodological barriers to surface in our team meetings. One such barrier was that the PR literature exhibited heterogeneity in the degree of participation by end-users in research (i.e., on a spectrum from lesser forms such as community consultations – to full co-governance and shared decision-making by academic and community members). This was important, because we felt that PR defined as having end users on board only to interpret research results versus being involved at every step of the research process (equitably co-governing) were two very different and incomparable PR designs. To add to this complexity, we saw variation in how authors of retained publications defined PR. There were examples in which the authors identified the research as PR, yet upon closer inspection, we observed that the partnership concept was framed in terms of placing members of a community through a training program rather than academic-community co-construction of research. Such revelations of our heterogeneous dataset helped to shape the iteratively-designed selection tool and also pushed us to refine our conceptualization of the subject area. As a result of such scrutiny, we revised the selection tool and re- applied it to all the included papers in the selection phase to allow for the retention of only those partnerships that exhibited multistakeholder co-governance across all stages of research. Such scrutiny of the diverse characteristics of participation clarified our field of assessment and generated a more conceptually homogeneous dataset which facilitated the application of realist methods.

2.3 The challenge of appraising the quality of complex evidence

In realist reviews, documents are appraised on the basis of the relevance and rigor of evidence in relation to the MRT being scrutinized (for more information on this, see Pawson 2006, pp. 87–90). It

was challenging to appraise the quality of PR partnerships according to these criteria, because of the heterogeneity of our sample and the fact that MRT did not guide the parameters of the review from the outset. In studying complete sets of documents for each partnership, we observed that the community–academic relationships were often sustained for many years, some into the decades. These partnerships often planned and implemented multiple interventions using diverse methodologies over many years. We had team debates over whether and how we should appraise the quality of methods used in the sequence of partnered research within a partnership life course and across our sets of partnerships. Key to developing our appraisal stage was deciding that the central criterion of appraisal should be that each partnership must have at least one instance of the authors describing or reflecting upon the process and impact of PR. Thus, the appraisal of partnerships was framed in terms of the degree of richness of partnership description. Two factors moved us away from appraising the quality of methods (i.e., quantitative, qualitative, and mixed methods) used in partnered research: (a) we determined that those methods, even if deemed as somehow compromised from a purely scientific perspective, could still reveal important PR outcomes; and (b) the majority of PR outcomes (e.g., capacity building and infrastructure development) fell outside the domain of intended research outcomes. Thus a methodological appraisal would not have been relevant and would have eliminated potentially valuable material from the review. This realist logic supported our customization of the appraisal process, and as a result sensitized us to the range of possible outcomes that accrue from the participatory process.

2.4 The relevance of capturing unintended outcomes

Another challenge concerned capturing the outcomes of PR, which accrued separate to the outcomes of the planned interventions. The intended outcomes of planned interventions included, for example, the number of women screened for cervical cancer from a lay health worker program. We initially thought that the review would comprise an assessment of the quantitative or qualitative intervention outcomes such as those. Yet, this proved to be unsuccessful, because we could not make a causal link between the PR approach and such ‘final’ outcome results (e.g., how could we say that the number of women screened in a cervical cancer awareness intervention was due to the participatory design?) What proved to be fruitful however, was dissecting the partnership processes across stages of activity and program implementation. At each step, we saw advantages to research processes and the building of community infrastructure, community empowerment, trust, positive systemic changes in communities, and a momentum building from one phase of partnership activity to the next. For example, we observed that PR changed or enhanced the infrastructure for community health promotion delivery over time, which was part of the process, but not often measured or directly linked to the outcomes the

interventions planned. In many instances, we observed outcomes of PR that were not noted by the study authors themselves. In these cases, we needed to apply a higher level conceptualizing to the evidence to synthesize these outcomes (explained in detail in further discussion).

2.5 Understanding the process of context, mechanism, and outcome configuring

A general description of the CMO configuring process has been presented in previous publications (Jagosh et al., 2011, 2012) and in the definition of terms provided in Table 1. Typically, pieces of evidence from primary studies are extracted and placed in one of the three categories of CMO. Our experience with CMO configuring suggests that: (a) there are instances in which an extracted piece of evidence can simultaneously be considered a contextual factor, a mechanism or an outcome; (b) along a chain of implementation steps, an outcome of one CMO can become context in a subsequent CMO; (c) differentiating context from mechanism is secondary to ensuring the assessment has captured all or most of the key elements that determine outcome success or failure (i.e., although it may be difficult to decide if a piece of evidence is context or mechanisms, it is not that critical, because there is natural overlap); (d) at times, elements of context can be seen as those which trigger mechanisms yet, at other times, the link between context and mechanism may not be causative in relation to an outcome of interest; (e) the tendency to conflate program strategy with mechanism should be identified and avoided; (f) CMO configuring typically begins with capturing an outcome from the a primary study, upon which elements of context and mechanism are added; and (g) the content of context and mechanism categories can be searched in the data or else generated through abductive reasoning.

Once we had retained our dataset, all articles were read (and reread) by at least two reviewers from the team, and passages were annotated whenever they described any impact or effect that participation made on any aspect of the research process or context. Those passages were then extracted into MS Word files in preparation for CMO configuring. Each passage was referenced in terms of the article it came from, the page and paragraph number. The initial step in CMO configuring involved an iterative process of determining what parts of the quotation fit the elements of context, mechanism, or outcome. Placing quotations in one of the CMO categories then spurred us to think about how the other categories should be completed. In the end, each CMO configuration involved a combination of an extracted quotation, along with abductive thinking reflecting our interpretation of the evidence. Typically, the quotation from the dataset was placed in the outcome category, which pushed us to think ‘abductively’ about the context and mechanism. All mechanisms involved abductive thinking. In some instances the context category also involved abductive thinking, and other times, it was populated with a quotation from an article.

We found that many quotations extracted for CMO configuring could be reasonably placed in one or more of the context, mechanism, or outcome categories. An example of this was the concept of trust. Trust (and mistrust) that was felt amongst the academic and community coalition members was understood as a mechanism that produced certain outcomes. However, deepening trust was also seen as an outcome, which became an element of the context in subsequent stages of partnership activity. Mistrust was also a common contextual factor. Seeing that this kind of evidence could be placed in any of the C-M-O categories pushed us to think about how CMOs arise in sequence, showing the building of trust (as an element of synergy) over time. The point here is that rather than belaboring the question of which category such an element would be best placed, we approached CMO configuring with the intention of making sure we did not miss any PR outcomes, regardless of where an element was placed.

In refining our understanding of mechanism, we had initially confused the program strategy with the mechanism. To overcome this, we found it useful to initially place elements of the program strategy in the mechanism box, but then tried to understand what new resource was offered to participants by the strategy and how the resource impacted participant's reasoning. This was also challenging, because we had to figure out just who the mechanism concept pertained to (academic or community coalition members? Or the community at large?) Overtime, it became clearer that the mechanism pertained to all these groups, depending on the evidence we were examining at that moment. As the synthesis progressed, we iteratively revised the mechanisms, based on clearer understandings of the concept and a more informed idea of what matters in PR assessment.

2.6 Incorporating middle-range theory in the context, mechanism, and outcome configuration process

Concurrent with data immersion during the identification, selection, and appraisal stages, numerous theories were pulled from the literature and examined for their utility as potential middle-range theories to explain the benefits of PR. Our criterion for MRT was that it had to provide overarching explanatory support that was 'adequate enough' to create the analysis of how PR works to produce intended and unintended outcomes. Due to the amount of heterogeneity encountered at the outset of our review, we were not able to identify any overarching theory to explain the benefits of PR at that time. Eventually, all theories except the theory of partnership synergy were excluded due to being too broad, too abstract, or too specific to only one dimension of PR. Table 3 outlines many of the theories we considered, and our decision and rationale for exclusion/inclusion.

Finalizing the decision for MRT occurred during synthesis as we were conducting (and re-reviewing) the CMO analysis, and it then shaped the way we assembled and understood the overarching

analysis.

Middle-range theory integration and development occurred as we were creating CMO configurations. At one point during synthesis, one of the co-investigators brought an article on the theory of Partnership Synergy (Lasker et al., 2001) to the attention of the rest of the team. The team read the article and agreed that the concept resonated with the ongoing synthesis. Partnership synergy is the notion that multistakeholder groups can pool their knowledge and resources to create something beyond what could be created by a single party working under similar conditions. This generated an overarching explanatory theory about why people choose to engage in partnered research and provided a general structure to understand how benefits accrue. For our review, although MRT did not shape the identification, selection, and appraisal protocols (as it was developed later into our review), it subsequently played a significant role in conceptualization and assessment of outcomes of data from our retained set of publications, specifically in the capture of seemingly ‘hidden’ outcomes of participatory research.

2.7 Using Middle-Range Theory to advance the conceptualization of outcomes – both visible and ‘hidden’

There are many emergent approaches to theory use in mixed studies reviews, the comparisons of which are beyond the scope here. However it should be noted that realist review is intended to be a theory-driven approach in which the identification of theory takes place early on in the review process and shapes the protocol to facilitate the capture of corresponding evidence (Pawson et al., 2005). We suggest that this works better in cases where the area of investigation has a singular question [e.g., Should smoking be banned in vehicles? (Pawson et al., 2011)], and less well in cases in which the research question is open ended (e.g., What are the benefits of PR?). For realist syntheses, MRT is expected to be an abstracted (i.e., middle-ranged) form of program theory that can be divided into theory parts and tested against corresponding to program components (Pawson, 2006b). Yet, our MRT did not resemble such middle-range program theory, because we were assessing what works in terms of the collaborative research approach rather than in relation to a set of programs per se. As such MRT contributed to the synthesis, not by dissecting the evidence into intervention components, but by facilitating the conceptualization of complex PR outcomes – complex outcomes in the sense that they could entail any evidence of the impact multistakeholder collaboration made on research process, health change, context features, infrastructure, and so on. MRT became a kind of currency that was most valuable in the unearthing of ‘hidden’ or ‘obscure’ outcomes. During the synthesis stage when we mapped the

chronology of events and activities of retained partnerships, we noticed that outcomes broadly conceived were either directly perceivable in the data or else hidden (i.e., perceivable through higher level conceptualizing). The partnership synergy theory helped to make sense of both of these types of outcomes, but especially the latter. Directly perceivable outcomes were immediately apparent and were often described clearly by the authors of the publications in the discussion sections of their papers. Outcomes through higher level conceptualizing were ones that required us to make inferences based on our broad examination of the partnership activity.

To portray how our middle-range theory supported the synthesis of these two types of outcomes, we present the following two CMO examples, which represent a common pattern observed in the literature. The first CMO configuration has to do with community leaders and academics converging to plan a community-wide needs assessment intervention:

Example of CMO Configuration Demonstrating a Directly Perceivable Outcome

Context: The needs assessment targets a refugee community in an urban American city. Prior to the project, complex historical traumas in the community created a collective psychology of fear; lack of community self-esteem and empowerment; mistrust of outsiders; it was anticipated that these factors created barriers to the participant engagement process in intervention implementation.

Mechanism: The academic-community coalition valued the “insider” knowledge of the community coalition members; academic researchers understood the importance of modifying the research design to increase its relevance and sensitivity for the community; community coalition members appreciated the willingness of academics to adapt research protocols to community needs.

Outcome: Through the collaboration process between academic and community leaders, the content and language of the needs assessment survey is modified to increase its relevance to the community; community leaders are able to elect lay health workers from the community who they know from being in the community that they are natural leaders, to implement the in-home survey; as a result, community participants feel safe in allowing door-to-door surveyors in their homes; this results in high participation in an otherwise difficult-to-reach population.

How partnership synergy theory supported this analysis: Synergy was demonstrated here in the creation of community-tailored research products and protocols and participant engagement outcomes which go beyond what could have been achievable by either party (academic or community) working in isolation toward the same goals.

The second type of outcome conceived through higher level conceptualizing is supported through

synergy theory in the following example: Here, the survey success alters an aspect of the community context for subsequent research. The outcome of one CMO configuration becomes the context for the subsequent. This is a continuation of the same example described in the previous CMO; however, the abductive–interpretive process is more heavily relied upon:

Example of CMO Configuration Demonstrating an Outcome Through Higher-Level Conceptualizing

Context: The process of implementing the tailored needs assessment survey and experiencing its success had strengthened the academic-community coalition member relationships, and created a bridge between the coalition and the community at large. Their project then had higher visibility in the community. Historical mistrust of outsiders was now offset by the communities’ experience of the culturally respectful approach by the community-academic partnership.

Mechanism: Coalition members felt enthusiastic and motivated by the success of the survey. They are encouraged by new interest from community members who want to become involved in the project.

Outcome: New synergy is created (here is the abduction). The coalition expands its membership due to new interest among community members in joining the coalition and informed by the results of the survey begins planning the next phase of the research.

How partnership synergy supported this analysis: Synergy is used as a conceptual currency, to understand that a successful outcome of partnered research created a momentum over time that moved the research forward.

Through the CMO analysis using partnership synergy theory, we were able to observe how community– academic partnerships transformed community settings, fraught with barriers, into fertile contexts for implementing health interventions.

3. DISCUSSION

In this article, we have reflected on our process of conducting a realist review for studying the complex phenomenon of participatory research. In describing our process and how we resolved our challenges, we offer a number of consideration points for the interest of experienced realist reviewers or those who are planning to conduct such reviews. The first point is in understanding the role of theory in shaping the review protocols. Theory may shape the protocols at the outset if there is prior conceptual clarity of the study area, a singular research focus and a relatively small, homogeneous evidence-base. If this is not the case, then an inductively driven, systematized screening procedure by using multiple coders may help in gaining clarity on the scope and nature of the evidence base. We found the

inclusion/exclusion process to be helpful in surfacing competing conceptions of the study area among team members, clarified conceptualizations in the diverse PR literature and helped to shape the parameters of the review.

We suggest that the conditions and characteristics of the literature under scrutiny, as well as the choice of scope will determine whether theory will guide the review or emerge as a product of data immersion. Variations in the role and purpose of MRT are described in realist review texts. Key to our discussion here is the way the interpretation of the evidence is facilitated by the candidate theories identified. We argue MRT should support an advanced understanding and conceptualization of the evidence. MRT that remains unintegrated in synthesis indicates a lack of fit between theory and evidence. In such cases, the theory needs to be refined, a more suitable theory needs to be retrieved or developed, or a different evidence- base needs to be captured.

Our experience corroborates the logic of realist appraisal as described by Pawson, who suggests that quality appraisal criteria that sorts ‘high’ versus ‘low’ quality research before the synthesis stage may prevent the unearthing of mechanisms of causation – a central feature of the realist approach (Pawson, 2006a). This point is even more relevant when MRTs are only brought in late in the review process. An important consideration for quality appraisal is ensuring that the retained dataset contains rich description of program processes and contextual elements. This type of evidence may be best located in the authors’ reflections and speculations on the research outcomes.

Although quality standards and guidelines for realist review may provide some beneficial guidance, it is still difficult to know before engaging with the area of inquiry, how methods should be customized. While the approach holds the promise of gaining deeper insights into the workings of programs and interventions, customizing the method can be time consuming and potentially fraught with conceptual challenges. We have suggested that customization can be beneficial for areas that are heterogeneous or have multiple operational definitions in the literature. It is likely that for realist reviews, much of the diversity of literature in a given area will need to be set aside and examined in separate syntheses.

In following the logic of critical realism, however, it is possible to synthesize outcomes that may not be readily apparent through a metasynthesis of final outcome measures. Realist texts suggest that mechanisms are usually ‘hidden’. Here we suggest that outcomes may also be hidden, particularly if they capture dynamic activity across time. Studying such types of outcomes requires a detailed study of the process of intervention implementation by reflections of the authors of primary studies and mapping the chronology of events and activities. Such added perspective lends external validity or generalizability to the evidence. Even when the literature does not provide author reflection or program process description,

realist logic encourages the inclusion of abductive reasoning in which reviewers can incorporate their ideas and hunches about how or why the evidence manifests in the way that it does.

There is no one way to distil/retain an evidence base, configure CMOs, and integrate theory into the process of undertaking a realist review. Each review may take a different approach. How a research team decides to customize the review protocols will depend on their collective expertise and approach, as well as the nature of the evidence base under scrutiny. It is therefore important that reviewers publish their approach to synthesis and reflect on the process. We also suggest that conducting a realist review with a multidisciplinary team of researchers and knowledge users is helpful in bringing to light debates and discussions that ultimately help to shape the protocol and push the synthesis forward. We hope the reflection presented in this paper and the insights we have gained from our experience can benefit the future use of realist review in the applied health sciences.

ACKNOWLEDGEMENTS

We would like to thank the knowledge-user (decision-maker) partners who participated in designing and implementing this review: Susan Law, David Clements, Marielle Gascon-Barré, David L. Mowat, Sylvie Stachenko, Sylvie Desjardins, Ilde Lepore, and Sarena Seifer. We also thank David Parry for his comments on the grant proposal and Robbyn Seller for her major contribution in writing the original grant application and designing the protocol. This review and postdoctoral fellow Justin Jagosh were supported by a Canadian Institutes of Health Research (CIHR) KT-Synthesis Grant (KRS-91805), funding from Participatory Research at McGill (PRAM), and the Department of Family Medicine, McGill University. Pierre Pluye holds a New Investigator Award from the CIHR. Geoff Wong and Trish Greenhalgh were funded by the National Institute for Health Services and Delivery Research (project number 10/1008/07) to provide methodological support. Margaret Cargo holds a Future Fellowship from the Australian Research Council (FT 100100312).

For G. Wong and T. Greenhalgh, the views and opinions expressed herein are those of these authors and do not necessarily reflect the opinion of the SDO programme, NIHR, NHS or the Department of Health.

REFERENCES

- Archer A, Bhaskar R, Collier A, Lawson T, Norrie A. 1998. *Critical Realism: Essential Readings*. Routledge: New York.
- Astbury B, Leeuw F. 2010. Unpacking black boxes: mechanisms and theory building in evaluation. *Am J Eval* 31: 363–381.

- Bhaskar R. 1998a. Philosophy and scientific realism. In Archer A, et al. (eds.). *Critical Realism: Essential Readings*. Routledge: New York.
- Bhaskar R. 1998b. The logic of scientific discovery. In Archer A, et al. (eds.). *Critical Realism: Essential Readings*. Routledge: New York.
- Cargo M, Mercer SL 2008. The value and challenges of participatory research: strengthening its practice. *Annu Rev Public Health* 29: 325–350.
- Freire P. 1970. *Pedagogy of the Oppressed*. Continuum: New York.
- Goodman R, Speers M, Mcleroy K, Fawcett S, Kegler M, Parker E, Rathgeb Smith S, Sterling T, Wallerstein N. 1998. Identifying and defining the dimensions of community capacity to provide a basis for measurement. *Health Educ Behav* 25: 258–278.
- Green L, George M, Daniel M, Frankish C, Herbert CP, Bowie W, O'Neill M. 1995. Review and recommendations for the development of participatory research in health promotion in Canada. The Royal Society of Canada, Ottawa.
- Greenhalgh T, Wong G, Westhorp G, Pawson R. 2011. Realist and meta-narrative evidence synthesis: evolving standards (RAMESES). *BMC Medical Research*, viewed May 9th, 2013 <http://www.biomedcentral.com/1471-2288/11/115>
- Habermas J. 1990. *Moral Consciousness and Communicative Action*. MIT Press: Cambridge.
- Hibbert P, Huxham C, Ring P. 2008. Managing collaborative inter-organizational relations. *The Oxford Handbook of Inter-Organizational Relations*, 391–416.
- Jagosh J, Pluye P, Macaulay AC, Salsberg J, Henderson J, Sirett E, Bush PL, Seller R, Wong G, Greenhalgh T, Cargo M, Herbert CP, Seifer S, Green LW. 2011. Assessing the outcomes of participatory research: protocol for identifying, selecting, appraising, and synthesizing the literature for realist review. *Implementation Sci* 6: 24.
- Jagosh J, Macaulay AC, Pluye P, Salsberg J, Bush PL, Henderson J, Sirett E, Wong G, Cargo M, Herbert C, Seifer S, Green LW, Greenhalgh T. 2012. Uncovering the benefits of participatory research: implications of a realist review for health research and practice. *Milbank Q* 90: 311–346.
- Kristjansson E, Robinson V, Petticrew M, Macdonald B, Krasevec J, Janzen L, Greenhalgh T, et al. 2007. School feeding for improving the physical and psychosocial health of disadvantaged elementary school children. *Cochrane Database Syst Rev* 1: CD004676.
- Lasker R, Weiss E, Miller R. 2001. Partnership synergy: a practical framework for studying and strengthening the collaborative advantage. *Milbank Q* 79: 179–205.
- Lawson T. 1997. *Economics and Reality*. Routledge: London.
- Macaulay AC, Commanda LE, Freeman WL, Gibson N, McCabe ML, Robbins CM, Twohig PL. 1999.

- Participatory research maximises community and lay involvement. *Br Med J* 319: 774–778.
- Macaulay AC, Jagosh J, Seller R, Henderson J, Cargo M, Greenhalgh T, Wong G, Salsberg J, Green LW, Herbert CW, Pluye P. 2011. Assessing the benefits of participatory research: a rationale for a realist review. *Glob Health Promot* 18: 45–48.
- Macaulay, AC, Jagosh, J, Pluye, P, Bush PL, Salsberg J. in Press. Quantitative methods in participatory research: being sensitive to issues of scientific validity, community safety, and the academic-community relationship. *Nouvelles pratiques sociales* 25: 2.
- Merton R. 1967. *On Theoretical Sociology. Five Essays, Old and New*. The Free Press: New York.
- Pawson R. 2006a. Digging for Nuggets: How ‘bad’ research can yield ‘good’ evidence. *Int J Soc Res Meth* 9: 127–142.
- Pawson R. 2006b. *Evidence-based Policy: A Realist Perspective*. Sage Publications: London.
- Pawson R, Greenhalgh T, Harvey G, Walshe K 2005. Realist review a new method of systematic review designed for complex policy interventions. *J Health Serv Res Policy* 10: 21–34.
- Pawson R, Wong G, Owen L. 2011. Myths, facts, and conditional truths: what is the evidence on the risks associated with smoking in cars carrying children? *Can Med Assoc J* 183: 680–684.
- Rycroft-Malone J, McCormack B, Hutchinson AH, DeCorby K, Bucknall TK, Kent B, Schultz A, Snelgrove-Clarke E, Stetler CB, Titler M, Wallin L, Wilson V. 2012. Realist synthesis: illustrating the method for implementation science. *Implementation Sci* 7: 33.
- Sayer A. 2000. *Realism and Social Science*. Sage, London.
- Wong G, Greenhalgh T, Pawson R. 2010. Internet-based medical education a realist review of what works, for whom and in what circumstances. *BMC Med Educ* 10: 12.
- Wong G, Greenhalgh T, Westhorp G, Buckingham J, Pawson R. 2013. RAMESES publication standards: realist syntheses. *BMC Medicine* 11: 21.

TABLES AND FIGURES

Table 1. Definition of realist review concepts.

Realism: Realism is a scientific epistemology (position of a community of researchers addressing the philosophical aspects of the scientific question: How do we know what we know?) that is said to be realistic about evidence gathering, given the limits of human capacity in garnering the truth of phenomena through observation, perception, and interpretation (Sayer, 2000). The philosophy of realism ponders on the fallibility of knowledge, bringing attention to the limits of both logical empiricism (e.g., as it manifests in the logic of randomized controlled trials and reviews of such) and constructivism as approaches to knowledge synthesis. Realism would suggest that one limitation of empiricism is that it sets aside important causal explanations if they are not immediately linked to observable evidence. On the other hand, constructivism, which emphasizes explanatory theories and mainly in terms of human or cultural interpretation, does not pursue an objective assessment of evidence that is deemed (by realism) to be foundational to a comprehensive understanding of causality. Realist modes of research reflect a mix of these two epistemologies by posing the kinds of questions that seek out the truth of matters, while at the same time operating from a view of the historically, socially, and linguistically contingent nature of human knowledge.

Realist review (RR): RR is an approach to synthesizing quantitative, qualitative, or mixed methods research, based on a realist logic of enquiry. It answers questions of the general format ‘what worked, for whom and in what circumstances, how and why?’ Initially designed to assess social policy, programs, and interventions, RR originated from a perceived inadequacy of usual empirical methods for evaluating social programs (Pawson, 2006b). Pawson’s critique of empiricism brings into question the applicability of empirical methods to assess causality (i.e., assessment of cause and effect) for social phenomena found in open systems. In systematic reviews of randomized controlled trials, for example, causality is investigated by counting the repeated occurrence of cause and effect across a number of cases. In contrast, realists point out the causal *mechanisms* (Astbury and Leeuw, 2010) involved in investigated phenomena. Realists argue that searching for causal mechanisms is more appropriate for investigating social phenomena in open systems (e.g., health programs in community-based settings) than for investigating material phenomena in controlled trials (e.g., pharmaceutical testing), because the former are open to an infinite array of influences that impact

outcomes, whereas the latter have been placed in studies designed to control extraneous and unknown influences and to account for effects of known, interjected influences.

Complex intervention: A complex intervention can be understood as a program (often multiple interacting component ‘interventions’) that offers resources or information, or enforces actions upon a target group (e.g., a population sample) and is built on implicit or explicit logic that such efforts will lead to positive change in or for that group. Complexity arises from the following: (a) the volitions of the target group; (b) their implementation in local (community or institutional) settings, which bring an array of uncontrollable contextual variables that exercise an impact on outcomes in addition to the intervention itself; and (c) emergent phenomena, where an outcome might change the underlying context and produce feedback that alters the prior implementation.

Middle-range theory (MRT): MRT is an implicit or explicit explanatory theory that can be used to explain specific parts of programs and interventions. ‘Middle-range’ means that it can be tested with the observable data and is not abstract to the point of addressing larger social or cultural forces (i.e., grand theories) (Merton, 1967). MRT is sought throughout the review and may initially help to shape the review protocols.

Context, mechanism, and outcome (CMO) configurations: CMO configuring is a heuristic used to generate causative explanations pertaining to outcomes in the observed data. The process draws out and reflects on the relationship of context, mechanism, and outcome of interest in a particular program. A CMO configuration may pertain either to the whole program or only to certain aspects. One CMO may be embedded in another or configured in a series (in which the outcome of one CMO becomes the context for the next in the chain of implementation steps). Configuring CMOs is a basis for generating and/or refining the theory that becomes the final product of the review. A simple example of a CMO configuration is as follows: A community experiences a high level of unemployment to which an employment training program is offered in a remote part of town (context). But the program has low enrollment and high attrition, and few people are trained (outcome). The reason is that people feel disillusioned by the lack of effort by program planners to ensure adequate public transportation to the venue (mechanism).

Context: Context often pertains to the ‘backdrop’ of programs and research. For example, in our review of participatory research, it pertains to the conditions connected to the development of research partnerships. As these conditions change over time, the context may reflect aspects of those changes while the program is implemented. Examples of context include cultural norms and history of the community in which a program is implemented, the nature and scope of existing social

networks, or built program infrastructure. They can also be trust-building processes, geographic location (e.g., rural or urban), types of funding sources, and other opportunities or constraints. Context can be broadly understood as any condition that triggers and/or modifies the mechanism.

Mechanism: A mechanism is the generative force that leads to outcomes. It often but not always denotes the reasoning (cognitive or emotional) of the various actors in relation to the work, challenges, and successes of the partnership. Mechanisms are linked to, but not synonymous with, the program's strategies (e.g., a strategy may be an intended plan of action, whereas a mechanism involves the participants' reaction or response to the intentional offer of incentives or resources). Identifying the mechanisms advances the synthesis beyond describing 'what happened' to theorizing 'why it happened, for whom, and under what circumstances'.

Outcomes: Outcomes are either intended or unintended and can be proximal, intermediate, or final. Examples of PR outcomes from our review include the following: increased self-empowerment, participant enrolment, higher education opportunities, acquisition of skills and knowledge, development of program infrastructure, and enriched research processes. Examples of intervention outcomes are improved health status, increased use of health services, and enhanced research results.

Demi-regularity: Demi-regularity means semipredictable patterns or pathways of program functioning. The term was coined by Lawson (1997), who argued that human choice or agency manifests in a semipredictable manner – 'semi' because variations in patterns of behavior can be attributed partly to contextual differences from one setting to another.

Abductive process: This concept means 'inference to the best explanation'. It involves an iterative process of examining evidence and developing hunches or ideas about the causal factors linked to that evidence.

Table 2. Goals addressed and types of documents handled in the review stages.

Stage of review	Type of document	Goal of that stage
Identification	Bibliographic citation including abstract	(a) To gain an appreciation of the heterogeneity within the field of literature; (b) To eliminate citations that made no reference to our area of investigation (PR) or did not reported any type of outcomes
Selection	Full-text paper	To clarify the assessment field and generate a conceptually homogeneous sample (in our case, to eliminate all papers that did not show equitable co-governance of PR across all stages of the research)
Appraisal	Partnership (sets of papers)	To prioritize sets of papers that demonstrated the richest data demonstrating impacts of PR and to eliminate all partnerships that did not have at least one example of evidence describing the impact of PR on research and health outcomes

Table 3. List of theories considered, excluded, or included.

Theory/Concept	Reason for Consideration	Exclusion or Inclusion	Reason for Exclusion or Inclusion
<i>Theory of Capacity Building</i> (Goodman et al., 1998)	Capacity building is a well-documented outcome of participatory research. The dialectical principle	Excluded	Too specific – did not apply to all PR outcomes
<i>Communicative Action Theory</i> (Habermas, 1990)	The dialectical principle of colonization vs. resistance as potential explanation for the tensions between researchers and community partners in community-based research.	Excluded	Too broad (metatheory)
<i>Empowerment/emancipation</i> (Freire, 1970)	Empowerment is considered a driver of participatory research.	Excluded	Not a theory
<i>Collaborative Management Theories</i> (Hibbert et al., 2008)	Possibility that theories of collaborative management from the business literature could be transferred and adapted to participatory research in health.	Excluded	Comprised theories on ‘how to’ collaborate. They did not generate explanatory power to understand PR outcomes
<i>Partnership Synergy Theory</i> (Lasker et al., 2001)	Synergy could provide explanatory power to understand both why people choose to collaborate and the extent of outcomes produced	Included	Provided explanation about how PR works to produce short and long-term outcomes