



**Thank you for downloading this document from the RMIT Research Repository.**

The RMIT Research Repository is an open access database showcasing the research outputs of RMIT University researchers.

RMIT Research Repository: <http://researchbank.rmit.edu.au/>

**Citation:**

Rogers, P, Westhorp, G and Walker, B 2015, 'Dealing with complexity in a realist synthesis: community accountability and empowerment initiatives' in Michael Bamberger, Jos Vaessen and Estelle Raimondo (ed.) *Dealing With Complexity in Development Evaluation: A Practical Approach*, Sage Publications, United States, pp. 385-404.

**See this record in the RMIT Research Repository at:**

<https://researchbank.rmit.edu.au/view/rmit:36079>

**Version:** Accepted Manuscript

**Copyright Statement:**

© 2016 by SAGE Publications, Inc.

**Link to Published Version:**

<https://us.sagepub.com/en-us/nam/dealing-with-complexity-in-development-evaluation/book242113>

**PLEASE DO NOT REMOVE THIS PAGE**

## Chapter 20

# ***Dealing with complexity in a realist synthesis: community accountability and empowerment initiatives***

Gill Westhorp, Bill Walker, Patricia Rogers

*Community accountability and empowerment interventions – like many interventions in international development – are complex in nature. They are inserted into diverse contexts; they attempt to achieve different goals; they work in different ways; they are affected by a wide variety of factors at national, sub-national and local levels; and effective interventions are responsive and adaptive. This chapter presents the methodology and findings of a review study on this topic using a realist synthesis approach, which is particularly suited to address complexity issues. The chapter begins with an overview of the review in terms of its scope, processes and findings. It then analyses the review in terms of Pawson’s VICTORE framework to demonstrate the complexity of the review topic, how complexity was reflected in the findings, and how the methodology of realist synthesis helped us to manage and deal with complexity. We then discuss how the findings from a realist synthesis can assist in dealing with the complexities of policy and program management in the real world.*

### **1. Introduction**

Realist evaluation (Pawson and Tilley, 1997) and realist synthesis (Pawson, 2006) were developed to assist policy makers and program staff to find pathways through complexity – specifically, to develop and test usable theory about complex and varied interventions applied across multiple contexts. One of the central ideas is that understanding how and why something works, or does not, in particular contexts can assist with practical decision-making: whether to use a particular kind of program in a particular situation; how to adapt a program to a particular context; and so on. But if complication and complexity are the norm in social programs (Rogers, 2008, Westhorp, 2012, Westhorp, 2013), how exactly does realist synthesis address this? Can it in fact deal with the complication and complexity of real programs and real decisions? And how might practitioners and policy makers actually use the results? This chapter explores these questions through the lens of a recently completed realist review<sup>1</sup> of community accountability and empowerment initiatives.

### **2. Overview of the review**

#### **Review Methodology**

Realist review (or realist synthesis) is a type of systematic review of existing evidence, grounded in a realist philosophy of science (Pawson, 2006). It is theory-based and seeks to infuse every step of the review process with theory. The RAMESES standards for realist reviews (Wong et al., 2013) make it clear, for example, that a realist review:

- examines not a program but an aspect of program theory;
- seeks to answer ‘realist questions’- not whether a program works overall, but for whom, to what extent, in which respects, in what contexts, and through what mechanisms;
- uses iterative search strategies that relate to program theory, rather than a single front-end search strategy based on search terms;

---

<sup>1</sup> ‘Realist review’ is another label for a realist synthesis.

- uses inclusion and exclusion criteria of relevance (i.e. relevance to program theory), and rigour (i.e. the aspects of the primary text that will be used in the review are of sufficient quality to support the conclusions which will be drawn from them in the review). Realist reviews do not include or exclude texts on the basis of research design (e.g. RCT or quasi-experimental design);
- involves a synthesis which refines elements of theory, rather than aggregating outcomes data;
- produces an improved program theory.

For reasons that are explained below, this was more a ‘theory-building’ review than a ‘theory-testing’ review.<sup>2</sup>

### **Background to the review**

Since the 1990s, community accountability and empowerment interventions (CAEIs) have been advocated as a way to improve educational outcomes by improving the quality of educational services and participation by students and families. CAEIs seek to increase the ability of communities to hold governments, funders, bureaucracies and service providers accountable to them for the provision of services and opportunities that meet basic rights. Community accountability can also involve questioning the standards to which public organizations are held, and the extent to which these are responsive to community needs. ‘Voice’ is therefore important – processes for the community to express their preferences, opinions and views. Changing accountability structures involves changing power relationships, and so power is also critical (Walker, 2009). Key elements of accountability include transparency of decision-making, answerability, enforceability and the ability to sanction (Rocha Menocal and Sharma, 2008, pp. 5-6).

The review identified four different types of initiatives related to community accountability in education. *Specific accountability interventions*, including community scorecards, citizen report cards, text-book monitoring, and monitoring of teacher attendance, have been designed to address specific problems and operate, at least in part, at the local level. *Decentralisation* may be relevant to community accountability and empowerment because it empowers local communities directly, or because it establishes a context in which it is easier for local communities to hold (closer) levels of government to account. *School-based management* is a particular form of decentralisation in which various decision-making powers and forms of budgetary control are devolved to school level. Types of school-based management in which parents hold control, or share control with school staff, may strengthen accountability of staff to communities. *Community schools* are a relatively common response to shortages of education provision, often involve significant control by community members, and are sometimes integrated into government strategies for expanding education access and improving accountability.

While there have been examples of positive results from some of these initiatives, overall research and evaluation of community accountability and empowerment initiatives has found mixed results (Ringold et al., 2012; Joshi, 2013). This review therefore sought to understand more about how and why various strategies work differently in different contexts. It focused on the primary education sector and on accountability at the local level.

It had previously been argued that community accountability and empowerment interventions improve educational outcomes by improving the quality of educational services and the participation of students and families in education. However, there had been no agreement about

---

<sup>2</sup> See Chapter 9 for a discussion on different review and synthesis approaches.

what is meant by ‘community accountability’ or ‘community empowerment’ in relation to education. The range of interventions which might affect accountability and empowerment was broad and evidence of impacts was mixed. This set of circumstances – contested understandings and mixed evidence - is by no means uncommon in international development. It provided part of the rationale for undertaking a realist review.

### **Negotiating the boundaries of the review**

The review was contracted by DFID as part of a joint call with the Australian aid agency (AusAID, now part of the Department of Foreign Affairs and Trade), and the International Initiative for Impact Evaluation (3IE). The initial call for proposals framed the research question in terms of a dichotomous question: ‘Do community accountability and empowerment initiatives improve education outcomes, especially for the poor, in low and middle income countries?’. Given the state of knowledge in the area, we proposed a revised research question and a realist review methodology to answer it: ‘Under what circumstances does enhancing community accountability and empowerment improve education outcomes, particularly for the poor?’

Following consultation with the funding body, the foci for the review were agreed to be: low and middle-income countries (LMIC); primary school education, because one of the Millennium Development Goals is to achieve universal primary education; a focus on girls and on marginalised populations, because they are frequently disadvantaged in relation to education; public (that is, government-provided) education, because that is most directly within the capacity of governments to affect; interventions that have, as their primary intention, to improve accountability of governments and education-service providers to communities, because these were the primary mechanisms of interest in the review question; and interventions that entail local-level participation or implementation, because the focus of the question was on community-level accountability and empowerment.

These agreements, along with the key terms ‘community accountability’ and ‘empowerment’, set the conceptual boundaries for the review. Setting conceptual boundaries is important, not just from a pragmatic point of view but also from a complexity theory perspective. Complexity theory is a member of the family of open systems theories, and open systems theories recognise that in reality, systems do not have neat edges. It is, therefore, necessary for the analyst to draw conceptual boundaries around ‘the system under consideration’ and to determine what is considered to be ‘in’ and what is ‘out’ in each case (Midgley, 2000; Cabrera et al., 2008). This same principle is reflected in the RAMESES standards for realist review:

*Because a realist synthesis may generate a large number of avenues that might be explored and explained, and because resources and timescale are invariably finite, the expectation is that the review must be ‘contained’ by progressively focusing both its breadth (how wide an area?) and depth (how much detail?). (Wong et al., 2013, p. 6)*

Even with these parameters established, the scope remained broad and the potential for complexity remained high. Just as one indicator: there were over 100 low and middle income countries in scope for the review, each with different histories, cultures, sources of disadvantage, education systems, education reforms, existing accountability relationships and so on.

Program theory – at least in principle – provides a way through such complexities. It provides a focus for what questions to ask, what data to extract, how to analyse and interpret the data, and how to explain the different patterns of outcomes. Realist program theory, with its particular focus on causal processes and the features of context necessary for their operation, provides a particular lens for doing so.

## Logistics of the review

An initial literature review had informed the development of the initial proposal. After acceptance of the research proposal, a protocol for the review was developed, describing its aims and rationale, definitional and conceptual issues, and the policy and practice background, and outlining the proposed methods to be used to search for and review studies, and to extract and synthesise evidence. The protocol was revised in response to peer review and feedback from the funding agency and then formally published (Westhorp et al., 2012).

Search strategies involved a combination of keyword searches in numerous databases; document searches of relevant websites; keyword and targeted searches using Google Scholar; snowballing of references of included documents and consultation with End User group members<sup>3</sup>. Checking reference lists identified many additional sources that had not been captured through the database and website searches. In fact, of the 268 documents included after initially screening, only 46 (17.2%) were identified through the original Boolean search strategy; of those, only 28 were eventually included in the review. This is consistent with previous studies, which found traditional searches to be a relatively poor basis for theory-based reviews (Greenhalgh and Peacock, 2005). The search process continued throughout the review.

21,000 documents were initially identified. These were reduced to 140 documents using inclusion and exclusion criteria. Inclusion criteria related to the agreed foci for the review: education sector interventions that sought to improve accountability of governments and education-service providers to communities; in low and middle-income countries; including a focus on girls and on marginalised populations; and, entailing local-level participation or implementation. Studies could be included for addressing any of these elements<sup>4</sup>. Texts in English and published after 1995 were included. Exclusion criteria were being outside the agreed foci for the review or insufficient quality. Quality assessment did not refer to a hierarchy of designs and all types of empirical studies were eligible for inclusion. Coded studies included formal research and evaluations, various types of case studies, RCTs, comparative analyses, and quasi-experimental studies.

Studies outside the agreed foci for the review or of insufficient quality were excluded. In keeping with the RAMESES standards for realist review, the quality assessment was done at the level of specific claims rather than entire studies. Material was reviewed by two team members, who conferred as required to make judgements about the trustworthiness of data within reports, referring documents to one of the lead investigators for discussion where they had doubts.

Included documents fell into two groups. The first comprised 30 'core' studies, which met inclusion criteria and provided interim or education outcomes data. Of these, sixteen studies provided evidence of impacts on student learning itself, in India, Indonesia, Uganda, Kenya, El Salvador, Guatemala, and Nicaragua. Many of these studies also reported intermediate outcomes. Another 14 studies provided evidence of intermediate education outcomes such as enrolment, attendance, or reduced corruption. The second group of around 110 documents provided evidence in relation to particular mechanisms or features of context.

---

<sup>3</sup> End User groups comprise people who may use the end product of a funded review – for example, policy and program personnel, not necessarily from the funding agency.

<sup>4</sup> Realist reviews do not have to be constricted to studies within particular program areas but there was more material within the education domain than could be managed in the available resources. It was not feasible to include material from other domains.

Originally, we had intended to use NVivo to code materials and we had developed a detailed coding framework to provide consistency across the review team. In practice, given the diversity of material and the theory-building nature of the task, this proved simply too cumbersome to use. Instead we developed a set of data extraction templates tailored to the particular review.

The final version of the report, after revisions in response to peer review and feedback from the funding agency, summarised the evidence for intermediate outcomes and student-learning outcomes from the included studies and presented significant theory development in terms of the causal mechanisms by which these interventions worked, the contexts in which they worked, and the relationships between accountability and empowerment (Westthorp et al., 2014).

### **Theory development**

It became apparent very early in the review process that programs in the area of community accountability and empowerment were generally under-theorised. Many programs referred to accountability but most were vague about exactly how or why they expected to generate it, who would become accountable to whom, or how it would improve education outcomes if they did. Some reports referred to a model of accountability as requiring transparency of decision making, answerability (the requirement to justify decisions), enforceability and the ability to sanction (Rocha Menocal and Sharma, 2008, pp. 5-6). However, few were clear exactly which of these aspects would be affected by their interventions. Many more referred to empowerment, but were unclear whether this was individual or collective, psychological or political, or what the relationships between these might be. In some cases, it appeared that programs assumed that the provision of information was sufficient to both motivate and empower communities to hold schools, teachers and education departments accountable. None of the reports or research reviewed clearly theorised the relationship between accountability and empowerment.

This was problematic because realist reviews are by their nature theory-based. Our response was to accept that the review would be more theory-building than theory-testing, and as befits a broad question in an under-theorised area, to aim for breadth more than depth of understanding. The review moved through six iterative components of theory development:

1. Development of the initial programme theory, later revised to a hierarchy of outcomes
2. Operationalisation of key terms 'community accountability' and 'empowerment'
3. Development of a typology of CAI interventions
4. Identification of 11 causal mechanisms involved in producing outcomes of interest
5. Identification of 13 contexts affecting mechanisms and 30 specific propositions
6. Theoretical model for the relationship between empowerment and accountability

The first piece of theory development was undertaken during development of the research protocol. A workshop involving the research team, practitioners in accountability interventions and academics was conducted. Results from the workshop were used to develop an initial programme theory for the overall class of community accountability and empowerment interventions, using an expanded hierarchy-of-outcomes format. This framework informed the development of search terms and strategies. It was then revised during the review process. The revised hierarchy of outcomes is shown in Figure 20.1, below.

Figure 1: Hierarchy of outcomes for community accountability and empowerment initiatives

ACTIVITIES	IMMEDIATE	SHORT TERM	INTERMEDIATE	INTERMEDIATE EDUCATION	FINAL EDUCATION
Engagement strategy	Involvement of necessary stakeholders	Application of sanctions/ incentives	Reduced corruption	Increased student enrolments	Improved student-learning outcomes
Information provision	Increased awareness relevant to project	Stronger relationships between stakeholders	Improved teacher attendance	Increased student attendance	
Capacity building	Increased knowledge/ skills for project roles	Community structures established or strengthened	Improved pedagogy	Increased student engagement	
Deliberation and planning processes	Plans tailored to local contexts	Stronger voice in advocacy or decision making	Improved teaching and learning resources	Improved student retention	
Implementation processes	Plans implemented	Participation in planned activities	Improved buildings and facilities	Decreased grade repetition	
			Stronger parent support for education		

Source: Westhorp et al., 2014

The second component of theory development involved operationalising the two key terms of ‘community accountability’ and ‘empowerment’. During development of the research protocol, we defined these terms. The term ‘community accountability’ was understood to refer to the ability of communities (primarily local communities) to hold governments, funders, bureaucracies and service providers accountable to them for the provision of services and opportunities that meet basic rights. The term ‘empowerment’ has many definitions and is surrounded by some confusion. Underlying these multiple meanings are long-running theoretical controversies over the nature of power. To reflect at least some of these meanings, ‘empowerment’ was initially operationalised by slightly adapting Friedmann’s model of empowerment (Friedmann, 1992). This resulted in a model of eight bases of social power: spaces, surplus time over subsistence requirements, appropriate information, knowledge and skills, financial resources, productive assets, social networks and social organisations. These various kinds of resources are either required for communities to be able to hold authorities and service providers to account (in realist terms, are necessary features of context), or may be developed as a result of community accountability and empowerment interventions (in realist terms, are outcomes of interventions which then create a new context within which accountability interventions may operate).

During the early phases of the review, the third component of theory development was undertaken – the development of the typology of four types of interventions relevant to community accountability and empowerment in education that was outlined earlier.

A basic program theory for each of these types of programs was developed. This later helped in understanding the kinds of mechanisms that might be expected to fire in particular kinds of initiatives. There was of course a dilemma in trying to relate programme mechanisms to the four categories of intervention. On the one hand, programmes' purposes and activities do affect the kinds of mechanism that might be triggered: mechanisms involve an interaction between what the programme provides and how targets respond. On the other hand, some mechanisms fire in multiple kinds of interventions. We addressed this dilemma in two ways: firstly, by listing the mechanisms which we believed, on the basis of the review, were most likely to be triggered by interventions in different categories, and secondly by identifying the features of interventions we believed were most likely to be necessary for specific mechanisms to fire.

Our project then moved on to the meat of a realist review, identifying program mechanisms and the features of context which affect whether and how they work. After months of detailed reading, extracting, analysing and discussing, the fourth component of theory development was complete - eleven proposed mechanisms, described in some detail, and with examples of each provided from the literature. These mechanisms are summarised (in much briefer format) in Table 1 below.

**Table 1: Mechanisms identified in the review**

Label	Description
1. Eyes and ears	Community members act as local data collectors for monitoring purposes, forwarding information to another party, which has the authority to act. The outcome of this mechanism is the action taken by the party that receives the information.
2. Carrots and sticks	Actors respond to actual application of rewards or sanctions.
3. Big brother is watching	Actors respond in <i>anticipation</i> of the application of rewards or sanctions.
4. The power to hire and fire	A direct, employment-based accountability relationship is established between a school management committee and school staff.
5. Increasing community capacity	Provision of training and 'learning by doing' support communities to develop knowledge, skills, and self and collective efficacy required for other actions.
6. Elder/Council authority	Strengthened relationships between school committees and other local authorities lend credibility and authority to the school committees to take specific actions to support education.
7. Increasing the capacity of local politicians	Local representatives develop an understanding of local issues and needs and increased confidence and skill to advocate for them.
8. Mutual accountability	All parties to an agreed action plan monitor the performance of all others, building mutual accountability.
9. Mind the gap	Discrepancies between rights or entitlements and actual provision surprise or concern local citizens, who demand change in response.
10. Our children's future	Increased understanding of and support for education motivates individual or collective action by parents to support children and schools.
11. It's working!	Seeing positive outcomes from an action operates as a positive feedback loop motivating further action.

Source: Westhorp et al., 2014

The ninth mechanism, 'Mind the Gap', is possibly the most common intended mechanism in information for accountability interventions. However we found extremely limited evidence that



this mechanism fires as anticipated. Rather, we theorise that information is a necessary but not sufficient condition for changed behaviours.

In realist terms, mechanisms are causal forces or processes that generate a particular outcome. The outcome generated by a particular mechanism can lie at any stage along an implementation chain or at any level on a hierarchy of outcomes. It is worth noting that none of these mechanisms (when they work) generate education outcomes per se. Rather, they generate outcomes at intermediate levels, which then create contexts in which education outcomes are more likely to be achieved or within which other actions can be taken. In order for the final intended outcome (in this case, improved student learning outcomes) to be achieved, sequences of activity - necessarily involving participation by different stakeholders – and sequences of mechanisms, each generating their own outcome, may be required.

The next stage of theory development was therefore the fifth component of theory development - to organise a map of some 'causal pathways' (sequences of activity and mechanisms) which appeared to generate particular intermediate outcomes. Complexity theory focuses attention on interactions and relationships between elements of systems, and on the 'local rules' that guide those interactions. Pawson and Tilley's (1997) construct of a mechanism ('reasoning and resources') has interaction between the program and the decision-makers built into its very core. Unsurprisingly, every mechanism identified in our review involves interactions between different elements of systems. Equally unsurprisingly, local culture constructs the 'local rules' that guide those interactions. There were numerous examples of parents being unwilling to monitor or supervise teachers because they perceived teachers to be more knowledgeable or because teachers were more powerful than themselves. As Fitriah noted (2010, pp 87-88), in Indonesia, when primary schooling was made free, parent participation in schooling and complaints about schooling decreased. No longer making a financial contribution, they perceived that they had lost the right and authority to do so.

Features of context that appeared to affect the operations and outcomes of interventions were identified through close reading of texts and sorted into 13 categories, including the broader political and social environment, the education system, information systems, de jure (legal) and de facto (actual) powers of local management committees, attitudes and roles of school management committees and school staff, the capacities of local communities, school facilities, gender, sustainability, engaging communities and enabling voice, engaging service providers and officials, and the roles of facilitators. 30 specific propositions about the circumstances in which community accountability and empowerment interventions are more likely to generate improved education outcomes were developed. For example, one of the propositions focused on social norms, drawing on evidence including the quote above about parents' perceptions: the proposition was that "Community-accountability and empowerment interventions are most likely to engage parents where they take into account social norms, parent resources and parents' intrinsic motivations." (Westthorp et al, 2014, p 123).

Another example focused on education assessment systems: the proposition was that "Community accountability and empowerment interventions are more likely to generate improved learning outcomes when there is a high-quality national system for assessment of student learning and when assessment systems are constructed to support collective action." (ibid p 113)

Each of the propositions was explained and supported with evidence, usually from multiple programs and countries.

These 30 propositions might be considered the ‘direct’ answer to the overarching question for the review (“*In what circumstances do community accountability and empowerment improve education outcomes...*”). However, taken in isolation, they do not constitute a realist understanding, because a realist view seeks to understand how context affects mechanisms (i.e. affects whether and which mechanisms fire). While there was evidence to support the influence of each of these propositions, at least in some contexts, there was not clear evidence to support linkages with particular mechanisms. Consequently, the next stage of theory-building was the construction of a provisional CMOC (Context-Mechanism-Outcome Configurations) table (Westthorp et al., 2014, Table 4). This was developed by aligning significant features of context against mechanisms, on the basis of evidence where it was available or on the basis of logic where it was not.

The sixth and final component of theory development was a theoretical model for the relationship between empowerment and accountability, proposed on the basis of the findings. Each element of the empowerment model was re-described in relation to the specific sections of the evidence that had been reviewed. The ways in which those elements interact with elements of the accountability model was also described. The ability to use formal theories and program theory to inform each other is part of the rationale for realist reviews.

The CMOC table and the empowerment and accountability model constitute the revised theory that is the intended product of a realist review. That theory remains to be tested and further refined through future research and evaluation.

### **3. Applying the VICTORE framework to an analysis of the review**

Pawson’s VICTORE framework (2013) is a useful way of understanding the many ways in which programs are complex (or, some might argue, ways in which they are complicated). The VICTORE model refers to Volitions, Implementation, Contexts, Time, Outcomes, Rivalry and Emergence (see Chapter 5).

#### **Volitions**

The term volitions refers to participants’ motivations and reasoning. It relates closely to Pawson and Tilley’s construct of program mechanisms as constituting some interweaving between the resources and opportunities provided by the program and the ‘reasoning’ of participants in response.

A number of the mechanisms identified in our review operated at this level, although not all referred to program participants per se. One example is the ‘Our children’s future’ mechanism. As parents participate in program activities (the resource), they develop greater understanding of the value of education for their children and of their own potential roles in contributing to education outcomes (intermediate outcome and mechanism). Because (most) parents want the best for their children, this increased understanding generates action to support education. Actions might be private, such as monitoring their children’s attendance, providing school lunch, or supervising homework; or they might be public, such as contributing to building facilities or monitoring teacher attendance. The ‘Carrots and sticks’ and ‘Big Brother is watching’ mechanisms both describe staff reasoning (i.e. volitions) in response to incentives, either positive or negative (i.e. resources).

The implication for program evaluation and review methodology is clear, and in fact lies at the heart of realist methodologies. Understanding the reasoning of stakeholders and participants enables evaluators and researchers, and the stakeholders on whose behalf we work, to understand how outcomes are or are not generated and why they vary across contexts. We noted – as have many before us – that traditional RCTs were generally speaking poor sources of data about stakeholder

reasoning, that qualitative data can be a rich source of understanding, and that the few theory-based evaluations we found were particularly strong.

For example, Lieberman et al. (2012) reported on the first phase of the Uwezo initiative in Kenya, a two-phase project to provide information to communities about how much their children are learning and how parents might support learning. A second phase of the project will involve broad dissemination of student assessment results and 'stimulation of a multi-faceted national discussion about children's learning' (p. 1). The programme theory is that:

*[T]hese measures will empower citizens to hold their governments accountable for improving the quality of their children's education, and also equip them with the knowledge necessary to contribute themselves to improving their children's learning (Lieberman et al., 2012, p. 8).*

The research systematically assessed the assumptions implicit in the theory of change for the first phase in order to identify 'the ways in which the treatment may have influenced outcomes—the mechanisms' (p. 16). They found no evidence of impact on any of the outcomes of interest from the first phase (p. 28) but note that outcomes are not expected until the second phase. They were able to discount implementation failure as the cause of the early lack of outcomes, but discovered through careful testing that almost all the assumptions underpinning the program design were incorrect.

#### *Implications for community empowerment and social accountability practitioners*

How parents, children, teachers and other key actors are or can be motivated to participate and to exact accountability and how they reason and make choices – and who makes those choices – are important for program outcomes. Practitioners might consider whether the mechanisms identified in this review are likely to fire in their programs. Alternatively they may consider their own program designs and ask practical questions: who does this program intend to do what? What would motivate them to do so? What does the program provide that is intended to motivate them to do so? These questions can be used as a guide to theorise mechanisms.

## **Implementation**

Most programs have long implementation chains and involve many processes, which implies that there are very many decision-making points. This was certainly true for most of the programs included in our study. Constructing our hierarchy of outcomes highlighted the chain – even though the diagram starts at the local implementation level (because of the focus of our review) and ignores earlier stages. Three of the categories of context we identified related specifically to implementation strategies: approaches to engaging communities and enabling voice, strategies for engaging service providers and officials, and the roles of facilitators.

Processes of implementation were generally well-described in formal research studies but relatively poorly described in most evaluation studies. Differences in implementation in different sites were rarely addressed in any depth. From a realist perspective, differences in implementation can affect the resources available to stakeholders and thus their capacity to enact change. Similarly, they can affect the 'reasoning' of stakeholders and thus their motivations to participate, or to sustain changed behaviours. Without good data about implementation, it becomes significantly harder to ascertain 'what matters' about implementation, and thus to use this to inform future policy and programs.

#### *Implications for community empowerment and social accountability practitioners*

The details of implementation – who does what, at what stage along an implementation chain, with whom – are likely to make a significant difference to the nature of the outcomes achieved. This is not because the activities themselves actually 'cause' the outcomes, but because the activities

provide resources and opportunities to which decision makers along the implementation chain respond. Policy makers and practitioners need a clear sense of both the 'what' and the 'why' of implementation strategies, in order to monitor whether implementation processes are generating intended outcomes. Further, recent evidence in relation to social accountability indicates that multiple accountability and empowerment strategies are much more successful than those involving single tactics (Fox, 2014). This increases the 'complicatedness' of implementation, which in turn may increase complexity (as more parts may interact, or different elements are affected by different aspects of context, this may generate more differentiated patterns of outcomes). Long implementation chains can also mean that more attention needs to be paid to sequencing of program elements. The challenge therefore is to remain clear and focused in planning, without expecting outcomes to follow in a linear manner.

## Contexts

Contexts are a linchpin of realist analysis. The term refers to the fact that programs are by definition implemented in existing systems and communities. More specifically, a realist approach acknowledges that programs are implemented in multiple, open, interacting systems, and draws attention to the fact that particular aspects of the context will have particular implications for whether and how programs work. Pawson draws attention to some of these: program stakeholders, relationships, institutional settings, norms, values, culture, and history amongst them. In formal complexity theory terms, context is reflected in the ideas of sensitivity to initial conditions and 'controlling parameters' – features that influence whether and how systems move between different ways of operating.

As discussed above, investigation of context was a major feature of our review. Effects of context were rarely investigated directly in the original studies. However, the process of realist synthesis enabled us to draw 'nuggets of evidence' from multiple studies to examine how and why particular features mattered. One clear example was that of adult literacy, which appeared to have multiple impacts on community accountability and empowerment interventions.

*Where members of SMCs are not literate, it is more difficult for them to administer schools effectively or exert influence on schools (Fuller and Rivarola, 1998, p. 39). Low parent literacy levels also appear to impede the ability of parents to assess their children's progress at school and, therefore, to judge whether or not the school is operating effectively (Banerjee et al., 2008, re: India; Blimpo and Evans, 2011, re: Gambia). Parents who can make those assessments may be more likely to make complaints (Blimpo and Evans, 2011, p. 27) or to intervene in school management (Gunnarsson et al. 2004, re: ten Latin American countries). (Westthorp et al, 2014, p. 122)*

### *Implications for community empowerment and social accountability practitioners*

The diversity of contextual factors that can affect whether and how programs work has two implications for practice. Firstly, practitioners need to think carefully about the assumptions built into their program designs and consider whether those assumptions are in fact met in the local context.

Secondly, practitioners may need to undertake an analysis of the context, identifying the factors that are most likely to affect the ways in which their programs will work, and add or tailor strategies to suit the local conditions. Some contextual factors are 'given' (relatively fixed) while others are more amenable to change, especially at community level. For example, existing community social capital (one of the contextual factors we identified) is not a fixture but may be built before an intervention is implemented. Some programs deliberately use a developmental phase to help build contextual features that will support the program or limit those which may undermine it.

## Time

Pawson notes that the term 'time' has many implications for programs: the idea of historicity (the idea that actual history creates the specific situation in which the program operates and the program then contributes to the history of the place or organisation), as already familiar through complexity theory, but also the idea of program duration and the various effects of sequences of programs. Historicity was evident in some features of context in our review, including the broader historical context within particular countries and the evolving nature of their education systems. Sequences of interventions were also apparent in some countries. For example in India, the Janshala (Community Schools) program was succeeded by the Sarva Shiksha Abhiyan (SSA—the Education for All Movement), the Indian Government's primary policy for universal primary education. Other projects and programs took place in the context of these broader policy initiatives, both past and present.

The issue of program duration was also significant, particularly in relation to sustainability. Previous research had demonstrated that it can take a number of years for programmes to generate changes of the kind that improve student learning outcomes (Patrinos et al., 2007). Successful interventions were often those that had been sustained over a significant period, including the oft-cited Philippines Textbook Watch programme. There was also evidence that programmes that are not sustained are not effective. Turnover of personnel means that the impacts of training are lost; unless new personnel are also trained, momentum and enthusiasm are lost (Evans et al., 2012, p. xii). As our review noted, different kinds of sustainability had different contextual implications:

*For some initiatives, being sustained means that policies and funding systems are sustained (political and economic context). For others, sustainability means that capacities are developed at the local level (local community context). This may require support and training by external organisations over a period of time, until sufficient members of the community are trained (implementation context). (Westhorp et al, 2014, p. 124).*

### *Implications for community empowerment and social accountability practitioners*

Pawson (2013) suggests that policy makers and practitioners should 'map' the major programs or initiatives that have preceded a new initiative, and give careful consideration to the effects those earlier programs may have on responses to the new offering. Policy makers and funders may need to allow time (that is, provide funds over sufficient time) for development and 'bedding down' of programs. In some sectors, including education, it may be many years before impacts on final outcomes become evident and sustained effort may be required to achieve intended goals. However, longer durations for programs also means that there is an increased chance that other contextual factors will change, so careful monitoring and revisions of planning to adapt to changes in context will be required.

## Outcomes

In his discussion of outcomes, Pawson notes that they may be short, medium and long term; that there may be multiple indicators for any indicator; and that outcomes do not speak for themselves but must be interpreted.

Because it does not require aggregation of outcomes or an average effect, the selection and analysis of evidence about outcomes is one of the areas where realist synthesis is strikingly different from other forms of systematic review. Our use of a hierarchy of outcomes enabled us to include, arrange and make logical sense of evidence in relation to a wide range of outcomes as successive stages in long processes leading to improved learning outcomes.

For realists, outcomes are generated by mechanisms operating in specific contexts. A single intervention could involve multiple mechanisms, each affected by specific aspects of context, to produce outcomes at different levels. We demonstrated this point by developing a detailed Context-Mechanism-Outcome Configuration (CMOC) for a single study - the Vidya Chaitanyam Project (VCP) in Andhra Pradesh (AP), India. The project promoted collective action by a large network of mothers already in self-help groups (SHGs), using a simple traffic-light score-card process to track several dimensions of school quality (Galab et al., 2013). The table demonstrates many things: how earlier stages of work generate intermediate outcomes that create new contexts in which later stages can work; how the same mechanisms can fire for different stakeholders; and how different features of context relate to different mechanisms.

Note that CMO tables are read ‘across and as a sentence’. Using the first line of the table below as an example: Where education quality is known to be poor, existing self-help groups have high levels of social capital and there is appropriate local infrastructure, participation by local officials triggers the ‘authority’ mechanism (described elsewhere in the report), resulting in self-help groups agreeing to address education. While each CMO configuration is shown in a separate row, they can be linked – for example, the outcome from one CMO configuration can produce a changed context for another CMO configuration.

**Table 2: Context-Mechanism-Outcome Configuration for a single intervention (see Galab et al., 2013)**

Context	Mechanism	Outcome
Education quality in AP is known to be poor (p. 14) Existing self-help groups (SHGs) have built up social capital over many years (pp. 16, 17) Existing infrastructure of SMCs, Village Organisations (VOs), <i>mandal</i> officials	Authority (A variant of Council/Elder authority) Clear state-sanctioned authority for role of SHGs (p. 17); legitimation of state authority by officials regularly attending higher-level meetings (pp. 16, 18)	SHGs agree to add education to their agenda (implicit in subsequent actions)  State-sanctioned authority structure exists (p. 17)
VC project is designed for low-literacy self-help-groups, using score cards developed in conjunction with district-level officials, (p. 18) Information campaigns re: rights <sup>5</sup>	Capacity building SHGs trained to use score cards (pp. 15,18) Increased awareness of rights (pp. 15,22,25)	Parents understand school quality and how to use score cards (p18, 24)  Parents develop capacity to question both SMC and teachers
Parents understand school quality Parents use data to ask questions of SMC and teachers (pp. 25, 35) SMCs meet regularly	Eyes and ears Parents monitor and report data to SMC (pp. 23,25,26)	SMCs discuss issues raised by parents using score cards showing gaps in school quality (pp. 22, 23, 24)
Score cards reveal gaps in school quality (p28–30) Parents aware of their rights (pp. 15, 25)	Mind the gap Parents/SMCs concerned about gaps shown by score cards (pp. 28–31)	Parents report the gaps to SHGs, SMCs and District/Mandal officials and ask questions after SMC meetings (pp. 18, 25)
SMCs discuss issues monitored	It’s working!	SMCs respond regularly to parents

<sup>5</sup> Community accountability interventions usually provide communities with information about their rights and entitlements in the issue area (here ‘education’), under law and under international treaties.

Context	Mechanism	Outcome
and raised by parents (pp. 23,24)	Positive feedback loops between parents and both SMCs and teachers	and SHGs (pp. 21–4) Parents gain increased capacity and confidence to monitor, measure and report performance (pp. 24–6, 35)
Parents have increased capacity and confidence to monitor measure and report school performance (pp. 24–6, 35)	Capacity building Parents learning by doing in asking questions of teachers and SMCs (pp. 25, 35)	Parents question school performance — directly to teachers, and via SMC (pp. 25, 26)
Parents question teacher performance Parents report on teacher performance to district and <i>mandal</i> Parents report on school quality regularly to SHGs, SMCs and district/ <i>mandal</i> authorities (pp. 18, 23, 24, 27)	Big Brother is watching Teachers are aware that authority structures of SHGs, SMCs and of districts and <i>mandals</i> are ‘watching’	Improved teacher effort: Teachers attend more regularly/on time (pp. 25, 26, 28,29,31, 32)
Parent belief in value of schooling (pp. 22, 24) Increased parent capacity and confidence (pp. 24–6, 35)	Capacity building Parents trained with regard to quality of pedagogy and testing of students (p. 18)	Parents and students observe improved teacher pedagogy (pp. 32, 33)
Parents and students observe improved teacher pedagogy (pp. 32, 33)	Big brother is watching Parents backed by authority system and students backed by parents (pp. 32, 33)	Teachers adopt more inclusive/engaging teaching methods, techniques and materials (pp. 32, 33)
Teachers adopt more inclusive/engaging teaching methods, techniques and materials (pp. 32, 33)	It’s working! Teachers, parents and students see this is working (pp. 32, 33)	Teachers sustain and increase use of inclusive/engaging teaching methods, techniques and materials (pp. 32, 33) Improved student-learning outcomes (‘A’ grades) (pp. 29, 35) Increased learning reported by students and parents (pp. 22, 32, 33)
Educational authorities report SMCs meeting regularly, score cards enhance parental questioning, teacher effort in response (p. 26)	It’s working! Education authorities see positive results	Authorities support and sanction on-going use of approaches (p. 26)

Note: page numbers refer to Galab et al. (2013)

Source: Adapted from Westhorp et al., 2014, pp 59-60

#### *Implications for community empowerment and social accountability practitioners*

Practitioners in the education domain may be able to use the hierarchy of outcomes from this review to position the outcomes they expect to generate, to consider the relationships between earlier and later outcomes, and potentially to refine plans or consider timelines for their projects. Practitioners in other domains could consider constructing a similar hierarchy, using domain-specific content, to make sense of outcomes that are likely to be causally related over time.

Similarly, the CMO tables in the review may assist practitioners to identify which aspects of context are likely to affect which mechanisms do and do not fire, and therefore which outcomes are or are not generated. This information can be used to refine program design, to design evaluations, and/or to explain patterns of findings in an evaluation.

## **Rivalry**

The term rivalry in the VICTORE model refers to the fact that multiple policies and programs are likely to operate concurrently in any policy area, some of which will be mutually reinforcing but others of which may be competing or undermining each other. While this is undoubtedly true, we found no evaluation studies that overtly took account of the effects of other programs in their analysis. Few even described the wider policy context within which they operated.

### *Implications for community empowerment and social accountability practitioners*

Program outcomes are likely to be strengthened if policy and program personnel take account of 'rivalry' when designing, implementing and evaluating programs. Pawson (2013) suggests mapping surrounding policies and programs in order to understand the context for the program under consideration (the 'evaluand'). In evaluation terms, this helps to deal with problems of attribution - either to avoid falsely attributing positive outcomes to the evaluand or to understand why programs do not achieve their intended effects. We suggest that such mapping be undertaken during program planning. If it is done then, planning can seek to build synergies between programs and/or take account of the potential negative consequences of other policies for the evaluand.

## **Emergence**

Pawson's construct of emergence is that program elements interact with each other, and with existing elements of systems, to create new elements of systems; or (in more familiar terms for those familiar with realist evaluation) that successful programs change the contexts in which they worked - and therefore should not necessarily be expected to continue to work. In systems theory terms, programs might create a positive feedback loop (in which change breeds more of the same change: success breeds success) or a negative feedback loop (in which change damps down the intended change). We saw many examples of programs making some change to systems at the local level, and a handful of long term program evaluations, but no rigorous studies of long term systemic change. This might in part be a product of the kinds of studies that are funded.

The Galab et al. (2013) study referred to earlier provided a good example of feedback loops generating changes to local systems, which we summarised as follows:

*... parental action in monitoring schools appears to have been sustained by a series of positive feedback loops in which parents saw that their collective actions were yielding increasingly effective outcomes as measured by their own score cards. They saw that their increased efforts as parents in encouraging their children to attend school and learn resulted in increased school attendance and learning; their attempts to engage SMCs and teachers garnered increasing responsiveness and, over time, generated the school-quality reforms they sought; and their own observations, recorded in score cards verified that measures of school quality were indeed improving over time. Further, the study claims and provides some evidence that, over the 18 months, this project was studied, a sense of joint ownership of school issues developed, which embraced collective problem solving and action by parents and the school —that is, both parents, SMCs and schools believed that their efforts were effective over the period studied (pp. 28-31, 36). (Westthorp et al, 2014, p.48-49)*

### *Implications for community empowerment and social accountability practitioners*



Some current development practices create rigid expectations about the rates at which programs are expected to progress and tie funding disbursements to those expectations. However, the emergent nature of these programs and the operations of positive and negative feedback loops within them suggests that more adaptive and responsive planning and funding models might be more effective. This in turn implies that monitoring practices should be adapted to capture emerging outcomes – expected and unexpected, positive and negative, and including any restructuring of systems and relationships in response to change.

So far, this article has considered how features of complexity, as described in Pawson’s VICTORE model, were manifest in our review. In the final sections of the chapter we identify the contributions to knowledge from the review, and then turn to consider the ways in which a realist review helps policy makers and practitioners to deal with complexity.

## **Summary**

Despite its challenges and limitations, we believe that the review contributes to existing knowledge in a number of important ways. It identifies the categories of intervention within which community accountability and empowerment interventions fit. It collates the evidence for intermediate outcomes and student-learning outcomes from the included studies. It proposes and provides examples of 11 mechanisms through which community accountability and empowerment interventions may work and identifies 13 categories of contextual features (representing a total of 30 elements of context, or ‘circumstances’) that affect whether and where community-accountability and empowerment interventions work. It proposes relationships between mechanisms and the elements of context most likely to affect them. Finally, it presents a new conceptual model for the relationship between accountability and empowerment.

Community accountability and empowerment interventions – like many interventions in international development - are complex in nature. They are inserted into diverse contexts; they attempt to achieve different goals; they work in different ways; they are affected by a wide variety of factors at national, sub-national and local levels; and effective interventions are responsive and adaptive. Theories of change, methods of evaluation and methods of systematic review all need to be able to take these various kinds of complexity into account. By focussing on the ‘how and why’ in addition to the ‘what’, realist methods are amongst those that can.

## **Practical applications**

Realist syntheses, like realist evaluations, are intended to be useful for policy and practice. Our review provided both brief ‘main messages’ for policy makers and much more detailed advice about using the various products that the review produced. Here we both summarise the specific guidance from this review and extract some main messages about the value of realist synthesis for policy and practice.

- There is credible evidence that community accountability and empowerment interventions can contribute to improved education outcomes - in some circumstances. The programs that had best demonstrated effectiveness seemed to have fired multiple mechanisms, sometimes concurrently and sometimes in sequence. This often required multiple program strategies.
  - Reviews which identify multiple mechanisms and the program strategies to which they relate –as well as the other contextual factors necessary for the mechanisms to fire – are likely to be useful for both policy and practice.

- Planning of community accountability and empowerment interventions (CAEIs) should begin with assessment of the broader social and political context to determine whether CAEIs are *feasible*; followed by an assessment of the education system to see whether CAEIs are the most *appropriate* response to local issues; and finally of community capacities to assess the most appropriate *type* or design features for the intervention.
  - This broad sequence may be applicable to many other sorts of interventions. ‘Portable findings’ such as these are a feature of realist reviews.
  - Realist reviews can thus frame the sorts of questions that need to be answered in preparation for new policy or program initiatives.
- The planning process should articulate the program’s theory of change, and identify the different elements that are required for the multiple mechanisms that are needed for outcomes to be achieved and sustained. Programs should be designed in such a way that they can be adapted to local contexts. Focusing on ‘what it takes for the main mechanisms to operate here’ shifts planning from a step by step guide to action, in which ‘fidelity’ of implementation is valued, to a more principles-based and complexity-consistent approach, in which the ‘fit’ of the intervention to the local context is valued.
  - This is a strength of realist methodology.
- The theory of change needs to be specific about the different types of accountability involved: who is to be held accountable, to whom and for what. Accountability arrangements necessarily involve power, and careful attention must therefore be paid to the nature of existing power relationships, the perceived problems in the power relationship and how those problems contribute to poor education outcomes before design can be attempted.

We also provided more detailed advice about the way the various ‘theory products’ of the review might be used. For example, we proposed particular questions about mechanisms that could be used for program planning, formative evaluation and summative evaluation. Mechanisms, however, do not stand alone. In realist analysis, context, mechanism and outcome are inextricably linked. CMO statements are intended to be read as a sentence: ‘In this context, that mechanism generates these outcomes for those groups.’ CMO statements provide meaningful interpretation of otherwise inexplicable patterns of data. Our table of CMOs might be used as a starting point for planning particular interventions; as a basis for development of more detailed or specific CMOs for particular interventions in particular contexts; and as a basis for evaluation design.

- The new theoretical model for empowerment and accountability could be used to analyse relevant features of the context and to design community accountability and empowerment interventions.

There were of course challenges in undertaking a synthesis of this size. Not surprisingly, we found that many studies failed to collect or report data about important aspects of the interventions we wished to examine (see also Chapter 16). We therefore recommended that future studies should: take account of and test the theory of change for the intervention, including variations in responses and outcomes across contexts; identify the different mechanisms expected to operate in and explicitly gather and make available data to better understand them, and the contexts in which they work; seek to identify and understand barriers to engagement in CAEIs and how these might be overcome; and ensure that detailed information from studies is available to later researchers, including access to detailed reports and data sets, to enable secondary analysis. These

recommendations are likely to be relevant in almost any field of study. The better the quality of the evaluations, the more useful the syntheses based on them are likely to be.

## References

Banerjee, A.V., R. Banerji, E. Duflo, R. Glennerster and S. Khemani (2008). Pitfalls of Participatory Programs: Evidence from a randomized evaluation in education in India, World Bank Policy Research Paper 4584, World Bank, Washington D.C.

Blimpo, M. P., & Evans, D. K. (2011). School-Based Management and Educational Outcomes: Lessons from a Randomized Field Experiment: Stanford University.

Cabrera, D., Colosi, L., & Lobdell, C. (2008). Systems thinking. *Evaluation and Program Planning*, 31(3), 299-310.

Fitriah, A. (2010). Community participation in education: does decentralisation matter? An Indonesian case study of parental participation in school management: a thesis presented in partial fulfilment of the requirements for the degree of Master of Philosophy in Development Studies Massey University, Palmerston North, New Zealand.

Fox, J. (2014). Social Accountability: What does the Evidence really Show? Washington, DC: World Bank/GPSA.

Friedmann, J. (1992) Empowerment – The politics of alternative development. Blackwell, Cambridge MA and Oxford, UK

Fuller, B. and Rivarola, M. (1998) Nicaragua's Experiment to Decentralize Schools: Views of Parents, Teachers, and Directors, Development Economics Research Group, Working Paper Series on Impact Evaluation of Education Reforms, (5).

Galab, S., Jones, C., Latham, M., & Churches, R. (2013). Community-based accountability for school improvement: A case study from rural India: CfBT Education Trust.

Greenhalgh, T. and Peacock, R. (2005) 'Effectiveness and Efficiency of Search Methods in Systematic Reviews of Complex Evidence: Audit of Primary Sources,' *BMJ (Clinical research ed.)*, 331 (7524): 1,064-1, 065.

Gunnarsson, L., Orazem, P., Sanchez, M., and Verdisco, A. (2004) Does School Decentralization Raise Student Outcomes? Theories and Evidence on the Roles of School Autonomy and Community Participation. Department of Economics Working Paper #04005, Iowa State University.

Joshi, A. (2013). Do They Work? Assessing the Impact of Transparency and Accountability Initiatives in Service Delivery. *Development Policy Review*, 31, s29-s48.

Lieberman, E., Posner, D. and Tsai, L. (2012) Does Information Lead to More Active Citizenship? An Evaluation of the Impact of the Uwezo Initiative in Kenya. Draft Paper, Dar es Salaam: Twaweza.

Midgley, G. (2000). Systemic Intervention: Philosophy, Methodology and Practice. Kluwer Academic/Plenum Publishers, New York, Boston, Dordrecht, London, Moscow.

Patrinós, H., Fasih, T., Barrera, F., García-Moreno, V., Bentaouet-Kattan, R., Baksh, S. and Wickramesekera, I. (2007) *What Do We Know About School-Based Management?* World Bank, Washington, D.C.

Pawson, R. (2006). *Evidence-based policy : a realist perspective*. London: SAGE.

Pawson, R. (2013). *The science of evaluation : a realist manifesto*. Thousand Oaks, CA: Sage Publications.

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

Ringold, D., Holla, A., Koziol, M., & Srinivasan, S. (2012). *Citizens and service delivery: assessing the use of social accountability approaches in the human development sectors*. Washington, DC: World Bank.

Rocha Menocal, A., & Sharma, B. (2008). *Joint evaluation of citizens' voice and accountability: synthesis report (No. EV692)*. London: DFID.

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

Walker, D. W. (2009). Citizen-driven reform of local-level basic services: Community-Based Performance Monitoring. *Development in Practice*, 19(8), 1035-1051.

Westhorp, G. (2012). Using complexity-consistent theory for evaluating complex systems. *Evaluation*, 18(4), 405-420.

Westhorp, G. (2013). Developing complexity-consistent theory in a realist investigation. *Evaluation*, 19(4), 364-382.

Westhorp, G., Walker, B., Rogers, P. (2012) *Under what circumstances does enhancing community accountability and empowerment improve education outcomes, particularly for the poor? A Realist Synthesis*. Protocol. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Westhorp, G., Walker, D. W., Rogers, P., Overbeeke, N., Ball, D., and Brice, G. (2014). *Enhancing community accountability, empowerment and education outcomes in low and middle-income countries: A realist review (No. 2207)*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

<http://r4d.dfid.gov.uk/pdf/outputs/SystematicReviews/Community-accountability-2014-Westhorp-report.pdf>

Wong, G., Greenhalgh, T., Westhorp, G., Buckingham, J., & Pawson, R. (2013) RAMESES publication standards: realist syntheses. *BMC Medicine*, 11(1), 21.