REALIST SYNTHESIS

RAMESES TRAINING MATERIALS

July 2013

Geoff Wong, Gill Westhorp, Ray Pawson and Trish Greenhalgh

Authors' affiliations

Geoff Wong[1], Gill Westhrop[2], Ray Pawson[3] and Trish Greenhalgh[1]

[1] Centre for Primary Care and Public Health, Queen Mary University of London, 58 Turner Street, London E1 2AB, UK

[2] Community Matters, P.O. Box 443, Mount Torrens, SA 5244, Australia

[3] Department of Social Research Methodology, University of Leeds, Leeds LS2 9JT, UK

Corresponding author

Dr Geoff Wong
Senior Lecturer in Primary Care
Blizard Institute
Barts and The London School of Medicine and Dentistry
Queen Mary, University of London
58 Turner Street
London E1 2AB
(Tel):   +44 (0)20 7882 2483
(Fax):   +44 (0)20 7882 2200
(Email):   grckwong@gmail.com

Competing Interests

None declared.

Contents

Section 1. A guide to the training materials

This training package has been developed to provide practical guidance to reviewers who want to undertake a realist synthesis (or realist review – the terms are synonymous). There has been, over recent years, a growing demand for training but so far, no 'how to' manuals exist. Development of the training materials was funded as part of the RAMESES project (http://www.ramesesproject.org).

The package comprises eight sections.   Section 1 is this introduction to the training materials.

Section 2 provides a brief overview of realism, a glossary of realist terminology, and some information about further references and resources.

Sections 3 to 7 focus on five areas that have been a frequent source of difficulties and misconceptions for realist reviewers. The topics are:

- Section 3: Focussing reviews

- Section 4: Program theory

- Section 5: Developing a search strategy

- Section 6: Selection and appraisal of documents

- Section 7: Applying realist principles in analysis

- Section 8: Further reading and resources

The quality standards for realist synthesis developed through the RAMESES project can be found online at: http://www.ramesesproject.org/index.php?pr=Project_outputs#method.   Note that the quality standards are not the same as the publication standards: they are an additional resource. These are designed to be used as a self-assessment tool by reviewers (and in particular those new to realist review) – a way to check whether the work done is consistent with the publication standards, and to work towards excellence.   They may also assist peer reviewers of protocols and realist syntheses to provide structured and consistent feedback to review teams.

The five focus areas (Sections 3 to 7 of this training package) are not the only areas that realist reviewers find challenging. The RAMESES research team selected these as priorities based on our past experiences in practice and training, an on-line discussion list operated as part of the RAMESES project, the literature, and our work in preparing the publication standards for realist syntheses (http://www.biomedcentral.com/1741-7015/11/21).

Realist reviewers all have different needs. Some are about to embark on a review but others are already in the process of doing one; some have years of experience in realist research and others are novices. Catering to such diverse needs is a big task. Also, how a realist synthesis is undertaken depends on issues such as the research question, the resource available, the nature and amount of relevant literature, funder's expectations, and end users' needs. Consequently, it is impossible to be prescriptive about 'what must be done'. The training materials serve more as guidance than as 'must-do's'. In true realist fashion, we expect that these resources will operate differently for different groups: as a structured introduction for newcomers and an aide-memoire for old hands. We also anticipate that they will evolve as others question, challenge and seek to improve them.

Our approach is to set out the main principles for each of these challenging areas and provide a series of resources for each one. Each section provides:

- Learning objectives for the topic

- An explanation about why the topic matters – why it is important to get right

- 'Quality standards' for the topic.

- Examples drawn from published reviews.
  In each case, there is at least one example from the published literature of how the topic area has been tackled successfully, and at least one worked example (also drawn from the published literature) of how the topic in a review might be improved. Our case examples are necessarily brief and so learners may need to read in full the document each example is drawn from to more readily appreciate the comments we have made.

- A learning activity that provides an opportunity to practice a key skill or work with a key idea.

- A set of focused questions to help reviewers to reflect on their own review project and how to achieve the standards in their particular project

Further reading for all the sections may be found at the end of these training and learning materials in section 8.

The examples we have selected are not featured to 'name and shame', or 'name and fame', particular authors. Rather, our purpose is to clarify and explain how the challenges can be tackled by giving 'real life' examples and by suggesting potential solutions. Our goal is to improve the overall standard of realist syntheses.

If you would like to help us further improve these training materials, please contact us either by:
Email   Geoff wong – grckwong@gmail.com
Or
Via the RAMESES JISCM@il email list
www.jiscmail.ac.uk/RAMESES

Section 2. Background materials

2.1 What is realism?

> "Realism is a methodological orientation, or a broad logic of inquiry that is grounded in the philosophy of science and social science." (1)

'Realism' refers to a philosophy of science. It sits, broadly speaking, between positivism ('there is a real world which we can see and understand directly through observation') and constructivism ('given that all we can know has been interpreted through human senses and the human brain, we cannot know for sure what the nature of reality is'). Realism agrees that there is a real world and that our knowledge of it is processed through human senses, brains, language and culture. However, realism also argues that we can improve our understandings of reality because the 'real world' constrains the interpretations we can reasonably make of it. While our knowledge will always be partial and imperfect, it can accrue over time. Below, we introduce key ideas in realist philosophy, how they apply to social programs and what they imply for the role of researchers and reviewers.

Mechanism

Realism can help us understand the social world. Used in this way, it acknowledges the existence of an external social reality and the influence of that reality on human behaviour. To understand the relationship between context and outcome, realism uses the concept of 'mechanism'.

There are many definitions and conceptualisations of mechanisms (even within realism) (see Section 2.3, Glossary). In realist philosophy, mechanisms are causal forces or powers. They cause things to happen, something realist have termed – generative causation (see Figure 1). Mechanisms in social science are comparable but not identical to mechanisms in natural science (e.g. the mechanism of gravity accounts for why a dropped object falls to the ground). Social mechanisms may usefully be defined as '… underlying entities, processes, or [social] structures which operate in particular contexts to generate outcomes of interest.' Here 'entities' may refer to things such as norms or belief systems, 'processes' are sequences where later events depend on earlier ones, and social structures may refer to things such as gender, class, or cultural patterns of relationships. Like the mechanisms in natural sciences, they possess a number of features: they are not 'visible', but must be inferred from the observable data; they are context sensitive, and they generate outcomes (2).
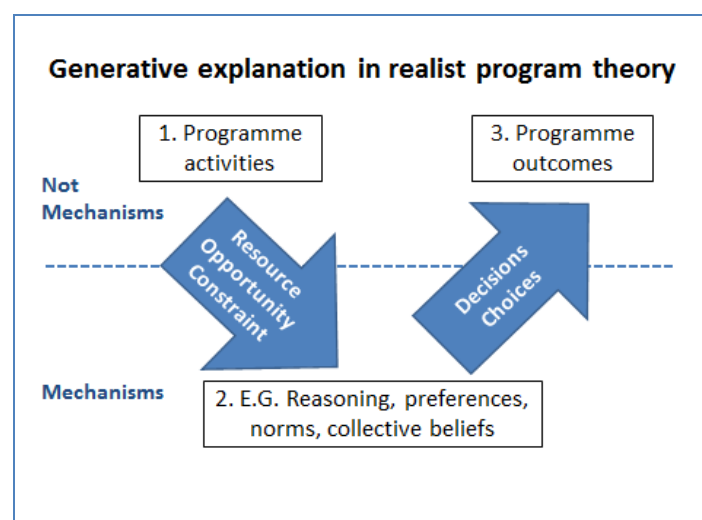

Figure 1: A pictorial representation of mechanisms

Social programs or interventions work by changing the decision-making of subjects. (We use the term 'subjects' here as shorthand for all those who may be directly affected by an intervention and whose decision-making does or could affect outcomes. In Pawson and Tilley's Close Circuit TV-in-car-parks example, 'subjects' included potential offenders, car owners, car park security staff, police, and passers-by. In human services programs, it usually refers to participants. However, mechanisms at earlier stages of implementation might involve funding providers, agency managers, service providers and so on.) The program or intervention changes the resources or opportunities available to subjects and, in that sense, changes the context for those subjects. The new context then triggers new mechanisms.

Using this logic, potential program mechanisms can be identified by asking what it is about a program that generates change. An intervention itself does not directly cause outcomes; it is the participants' reaction to the opportunities provided by the program that triggers the change. A realist approach therefore looks for interactions among the opportunities or resources provided by the intervention and the reasoning or responses of the participants.
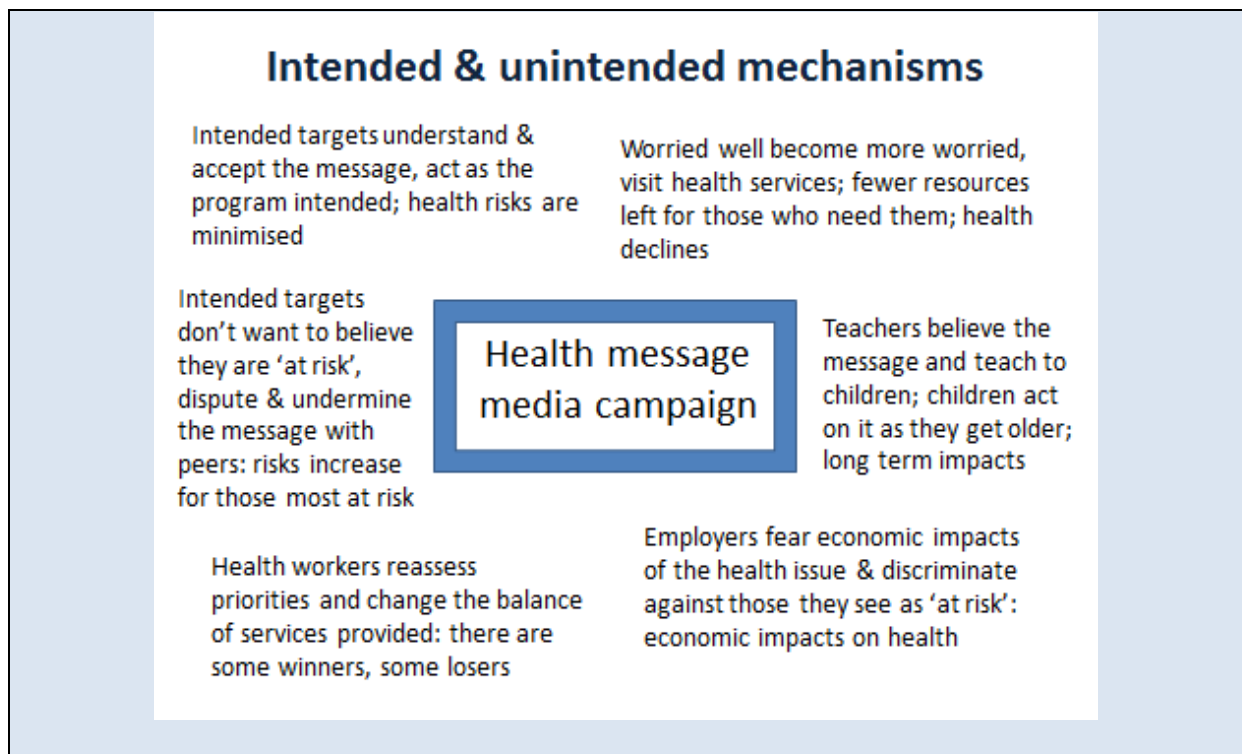
One route to identifying program mechanism is to reconstruct, in imagination, the reasoning of participants or stakeholders. When asked how the intervention was influential, a subject might reply, 'It made me ponder A, see alternative B, grasp opportunity C'. Mechanisms may also generate negative effects. So other participants may say, 'I've tried D previously, I'm bored with E, I prefer to do F'. Starting in this way generally reveals to researchers that they are dealing with a potentially a large number of mechanisms. The role of the realist researcher therefore necessarily involves identifying 'main mechanisms' – those that are common and significant enough to contribute to the pattern of outcomes of the intervention.

An important principle of realism is that the 'causes' of outcomes are not simple, linear or deterministic. Programs often work through multiple mechanisms. Some mechanisms are obvious and correspond to those intended by the program's designers, some are less obvious, and some are not anticipated by the designers. A mechanism is not inherent to the intervention, but is a function of the participants and the context. Consequently, the same intervention can trigger different mechanisms for different participants, even within one location. Programs run across very different social contexts are quite likely to generate different patterns of outcomes in those different contexts (see Box 1).

Box 1: An example of the intended and unintended mechanisms

Intended and Unintended Mechanisms

Consider the example of a health promotion media campaign (safe sex, safe needle usage, more fruit and vegetables, more exercise …). It is likely that the campaign intends to reinforce healthy behaviour by the majority of the population, but change the behaviours of those 'at risk'. However, the campaign is likely to trigger multiple mechanisms concurrently, for a number of different groups. Some of those mechanisms support the intended outcomes of the campaign, some undermine it, and some generate unintended outcomes. Some will operate over the short-term, some over the long term. See the results of a quick thought experiment in the diagram. Note that in each case we have identified (in very brief form) the 'reasoning' of the subgroup, the action that they take as a result, and the outcomes those actions may generate.

Intended & unintended mechanisms

Intended targets understand & accept the message, act as the program intended; health risks are minimised

Worried well become more worried, visit health services; fewer resources left for those who need them; health declines

Intended targets don't want to believe they are 'at risk', dispute & undermine the message with peers: risks increase for those most at risk

Health message media campaign

Teachers believe the message and teach to children; children act on it as they get older; long term impacts

Health workers reassess priorities and change the balance of services provided: there are some winners, some losers

Employers fear economic impacts of the health issue & discriminate against those they see as 'at risk': economic impacts on health

It is also important to note that different kinds of mechanisms operate at different levels of reality and through different kinds of programs (see Figure 2). Consider, for example, the different kinds of interventions that could be tried in an attempt to reduce domestic violence. Family therapy and Cognitive Behaviour Therapy (CBT) are different kinds of therapeutic approaches and work through different mechanisms. Family therapy (at least according to its program theory) works by changing family dynamics, which changes family members' experiences in the family, which changes their responses to experiences and thus their behaviour. CBT works, so the theory goes, by retraining the cortical and limbic systems in the brain, and does so regardless of whether the program focus is reducing family violence or overcoming addiction. Drug therapy works by changing chemical processes within the brain, while education and training programs work by changing knowledge, skills and/or attitudes. Legislation may work through deterrence or – where prison terms result – by temporary disablement (i.e. the offender cannot offend against the family members while imprisoned –although their violence may simply be directed to others (a mechanism known as displacement). Community development may work by changing community norms about acceptable and unacceptable behaviour or by changing bridging social capital, such that victims can access support to escape from violence. Note in each case that we have identified a program type or strategy, and that the mechanisms through which they work operate at a different level of reality than the actions of the programs themselves.
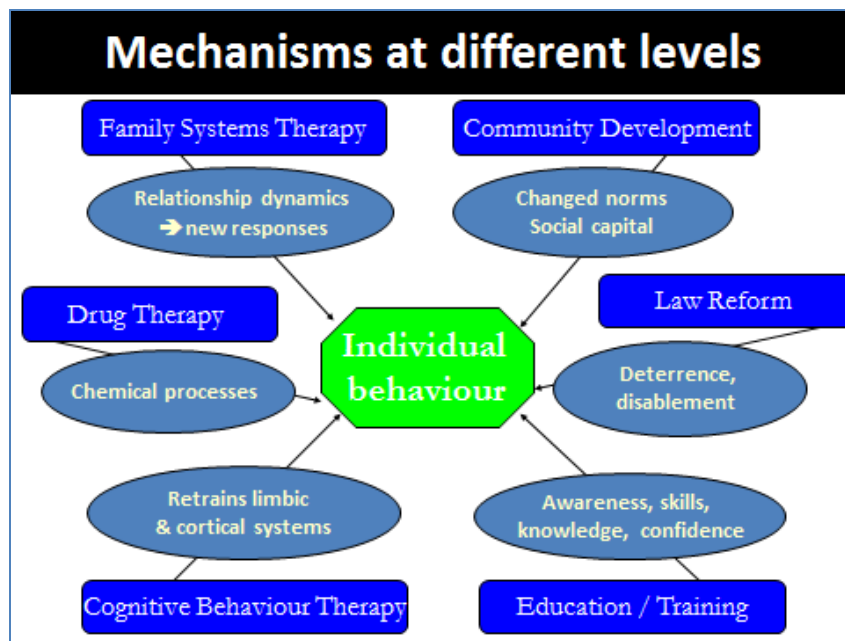
Figure 2: Mechanisms operate at different 'levels'

In summary, generative mechanisms:

- cannot be seen or measured directly (because they happen in people's heads, or at different levels of reality than the one being observed);

- are context-sensitive;

- are multiple (hence, when researched, they need to be unpicked, defined and prioritised);

- are best expressed at a somewhat abstracted level, so that they are not tied unnecessarily to particular people, places or things.

Context and Mechanism

Researchers will be familiar with the observation that varying outcomes occur in different contexts. The realist explanation for this variability revolves around mechanisms and their interactions with other mechanisms and context. Although the endless permutations and combinations of interactions might be expected to produce no observable patterns, the fact is that patterns do occur. Realism suggests that this is because similar mechanisms are being triggered in some contexts, producing the similar bits in the pattern; and different mechanisms are being triggered in other contexts, producing a different part of the pattern.

The term 'context' may refer to broad social or geographical features (for example the country in which an intervention operates and its cultures); to features affecting the implementation of programs (for example whether the program occurs in a jail, a hospital or health service, whether there is adequate funding, the qualifications of staff). It could also relate to the make-up of the participants on a program or the different population profiles of locations in receipt of an intervention. It could also relate to the conditions in which subjects seek to enact their choices (graduates of a vocational training program will find it easier to get work in a context of high employment; recipients of a housing subsidy will find it harder to use that subsidy in a context of housing shortages.) 'Context' in short can take on a multitude of forms. The realist use of context,

however, is not just a matter of listing the infinite potential 'surrounds' to an intervention. What matters is developing an understanding of how a *particular* context acts on a *specific* program mechanism to produce outcomes – how it modifies the effectiveness of an intervention. We now modify our earlier diagram (Figure 1) to represent this interplay and the patterns of outcomes that can be expected to result (see Figure 3).
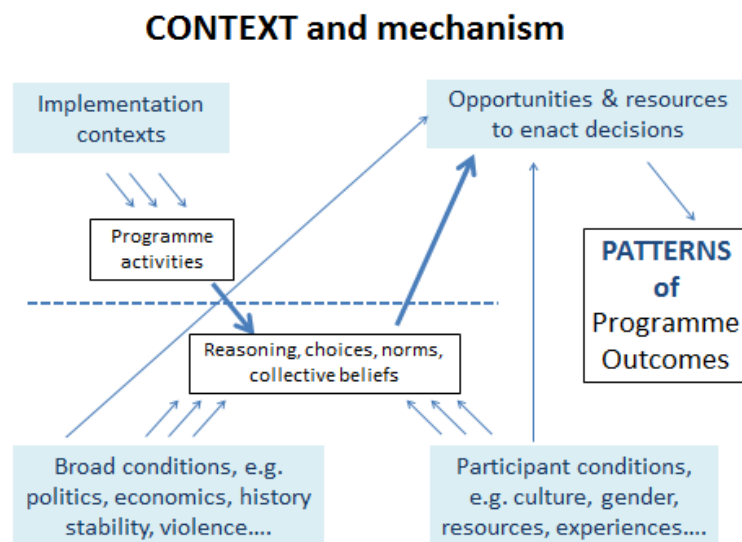

Figure 3: The interplay between context and mechanism

In summary, realism holds that mechanisms matter because they generate outcomes, and that context matters because it changes (sometimes very dramatically) the processes by which an intervention produces an outcome. Both context and mechanism must therefore be systematically researched along with intervention and outcome. By implication, research or evaluation designs that strip away or 'control for' context with a view to exposing the 'pure' effect of the intervention limit our ability to understand how, when and for whom the intervention will be effective.

2.2 Theory and Realist Synthesis

Realist review was developed as a theory based approach to synthesising existing evidence. It is, therefore, part of the school of 'theory-led' or 'theory-driven' research. 'Theories' rather than 'programs' are thus the basic unit of analysis. This move represents a considerable intellectual leap that is often misunderstood. Research begins with program theory and ends, if it has been successful, with a revised, more nuanced and more powerful program theory. Being disembodied from any specific application in this manner, research findings have the potential to be applied on any subsequent occasion that such a theory comes into application. This borrowing of policy ideas happens all the time. Some realists are fond of saying that there is nothing new under the sun when it comes to developing the ideas for interventions.

There are a number of reasons why theory is important. As we have already noted, theory is implicit in all programs. Using an abstracted program theory for the review helps reviewers to move beyond the minutiae of particular programs to focus on the major ideas within them. Once developed, reviewers can also compare the program theories (explicit or implicit) of the various interventions they examining to more general substantive theory, which can help in understanding the differences

between interventions and perhaps, therefore, the differences in how they work and the outcomes they generate.

Other reasons for using a theory-based approach include:

- There are always more questions that you could ask in a review than you will have the capacity to answer. Using theory helps to focus the review and to decide which questions to ask.

- There is always more literature that you could examine than you will have the resources to do. Theory helps to determine which literature is most relevant.

- The information gathered in a review always has to be interpreted and theory provides a guide for interpretation.

- The information gathered in a review is usually complex and messy: theory provides a basis for abstraction and for understanding 'the patterns in the data'

- Given that evaluations of the 'same' program (or kinds of programs) almost always show different results in different situations (or for different groups), theory provides a basis for explanation of the patterning of outcomes.

- Attributing outcomes to programs is complex. In primary evaluation, theory provides a basis for causal attribution. In reviews, theory provides a framework with which to assess the plausibility of attributions made by original authors.

However, the word 'theory' has many meanings. As Pawson and Tilley noted in their book Realistic Evaluation (3), is can refer to everything from grand over-arching theories such as Marxism to specific hypotheses that are tested in a laboratory experiment. In theory-based research and evaluation, there are four kinds of theory that matter (see Figure 4).

One is the underlying philosophy (realist philosophy). Realism takes particular positions both in relation to the nature of reality (that is, ontology) and the nature of knowledge (epistemology). This is realist philosophy (also sometimes called **realist theory**).

The second is **methodological (i.e. research and evaluation) theory**, or the implications for research and evaluation methodology that realist philosophy implies. One most often hears of 'realist evaluation' and 'realist synthesis' (or review), but other forms of realist research are also possible (see Section 8. Further reading and resources).

The third is **program theory**. This is the theory about what a program or intervention is expected to do and in some cases, the theory about how it is expected to work. Realist program theory goes a little further and includes descriptions of contexts, mechanisms and outcomes.

The fourth is '**formal theory**' or '**substantive theory**'. This is theory within particular domains – sociology, economics, psychology, education, health and so on. Examples include game theory in economics, constructivist learning theory in education, attachment theory in human development and so on. Sometimes substantive theories are used to design programs. They may also be used to inform program theory that is developed 'after the event' – for example, when evaluators develop program theories for programs that are already underway, or when reviewers develop the 'initial rough theory' for their review. Substantive theories are often used to help make sense of CMO patterns – to contribute to the 'synthesis' stage of a realist synthesis.

It is not necessary to have a strong command of all these types of theory before beginning a realist review. However it is worth remembering that the philosophical underpinnings of the various studies in a review may be different. It is still possible to undertake a realist review, regardless of the philosophical assumptions built into the program or the research methodologies used within the documents that are reviewed.
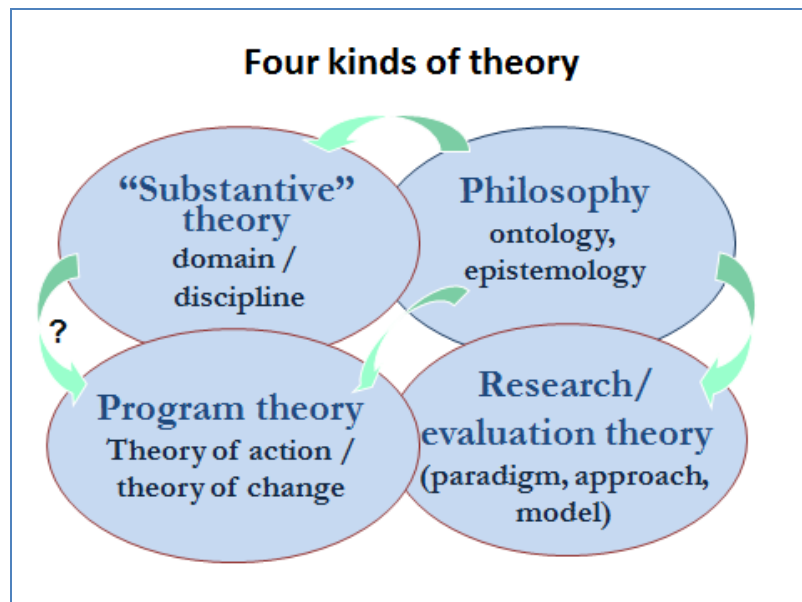

Figure 4: Main types of theory relevant to realist reviews

There are three other uses of the term 'theory' that are important in realist synthesis. These are 'initial rough theory', 'refined theory' and middle range theory.

The term **'initial rough theory'** refers to the initial sketch of a theory that is used to guide a realist synthesis. This is often a program theory. However, some questions for realist synthesis do not refer to a particular program. Here, the 'initial rough theory' sketches 'whatever it is that the question is investigating' (what would be called 'the evaluand' in evaluation) and how it is expected to work.

Initial rough theories may or may not be constructed in realist terms. Program theories provided by commissioners of reviews, for example, are not usually framed that way. To support the process of undertaking a review, it is at least useful to construct an initial rough theory of action (What is supposed to happen?) and an initial rough theory of change (Why is that supposed to work?). If it is possible to construct the theory of change in realist terms, so much the better.

A '**refined theory**' is the product of a realist review. In the process of conducting a review, some aspects of the initial rough theory may have been proved wrong. Others may have been supported with strong evidence. Many (perhaps most) will have been refined to some extent. Refining a theory might mean becoming clearer about the contexts in which, or population groups for which, an intervention will or will not work. It might mean developing a more sophisticated understanding of how particular mechanisms work. It might mean refining understandings of outcomes ('to what extent', 'in what respects', 'over what timeframe'). The product may be a set of CMOs – statements about the contexts in which particular mechanisms generate particular outcomes. The final stage of theory refinement is to develop an explanation of why the CMO patterns look the way they do.

Perhaps a useful analogy to use for the theory building and testing process of realist reviews is that of being a police detective. Both work 'backwards' and 'outwards' (by this we mean casting the net wide) from the outcome of interest. For the detective the outcome is the crime – say murder most dreadful. The detective starts to form ideas about how and why it might have happened. The former may seem obvious – a stab wound is seen, but that may not be the cause of death – and so the detective can speculate from the clues available at the crime scene, but cannot be sure until there is a post mortem (and perhaps not even then). The analogy here is clear – realist reviewers can speculate based on clues from (for example) the articles they read, but need to continue to seek data to confirm their speculations. The detective will search for more clues, interview suspects, obtain forensic evidence (such as looking for fingerprints on the murder weapon), all in a bid to confirm, refute or refine any initial speculations as to 'who dunnit'. This is analogous to the realist reviewer seeking more data – be it from formal searching or though lay or content experts. Both may cast the net wide in search of data to test speculations. Neither knows for sure what they are looking for, but follow leads or clues that show promise. Finally, neither knows for sure the 'truth' but instead builds up an evidence base to support their theories – a case that is judged on coherence and plausibility.

The term '**middle range' theory** refers to the level of abstraction at which useful theory for realist work is written: detailed enough and 'close enough to the data' that testable hypotheses can be derived from it, but abstracted enough to apply to other situations as well. This is a theory that lies;

> *"…between the minor but necessary working hypotheses that evolve in abundance during day-to-day research and the all-inclusive systematic efforts to develop a unified theory that will explain all the observed uniformities of social behavior, social organization and social change…*
> *It is intermediate to general theories of social systems which are too remote from particular classes of social behavior, organization and change to account for what is observed and to those detailed orderly descriptions of particulars that are not generalized at all. Middle-range theory involves abstraction, of course, but they are close enough to observed data to be incorporated in propositions that permit empirical testing.[our emphases]"* (4)

Middle-range theory can be thought of as a 'knowledge repository'. It holds promise for evaluation and systematic review methodology because it offers an approach to linking findings from program to program and from policy to policy. Interventions are normally targeted at specific outcomes in specific populations but the problems they deal with often have a common genesis. For instance, very many programs designed to change behaviour wrestle with the problem of persuading inveterate 'outsiders' (be they drug users, educational underachievers, excessive eaters, etc). to become reformed 'insiders'. We know that prizing people out of such out-groups is difficult and know that targeting 'marginal members' may be a more realistic aim. Accordingly, the opportunity for learning in realist synthesis is thus the middle-range task of trying to figure out what constitutes marginal membership of such groups. There is a healthy sociological and evaluation literature to draw on in accomplishing such a task and its accomplishment offers significant potential learning for the *next* behaviour change program.

Realist reviews may draw on existing middle range theories to develop the 'initial rough theory' that they test. The outcomes of a realist review are also ideally framed as middle range theory – that is, theory that can usefully be applied to a family of interventions, or to a problem that manifests in a number of domains. Program mechanisms are also usually described at a middle range level.

There is more detail about program theory in realist review in Section 4 below.

2.3 A Glossary of Terms

There are multiple definitions and descriptions of almost all the terms used in realist research and evaluation. Those that are provided here draw from a particular school that draws on the work of authors such as Popper, Campbell, Bhaskar, Sayer, and Archer (5). Our interpretations for the purpose of realist evaluation and realist synthesis draw heavily on publications by Pawson and Tilley (3), Pawson (1) and Pawson (5).

Context

"Context often pertains to the "backdrop" of programs and research. … 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 effects, funding sources, opportunities, or constraints. Context can thus be broadly understood as any condition that triggers and/or modifies the behavior of a mechanism." (6)

Context-mechanism-outcome (CMO) configurations

A CMO configuration is a statement, diagram or drawing that spells out **the relationship between** particular features of context, particular mechanisms and particular outcomes. In a sentence, they take the form of "In 'X' context, 'Y' mechanism generates 'Z' outcome." To more fully appreciate CMO configurations readers may wish to familiarise themselves with the concept of mechanisms (as defined below in this section).

"CMO configuring is a heuristic used to generate causative explanations pertaining to the 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 to either the whole program or only 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." (6)

More details about CMO configurations may be found on pages 21 to 27 in 'The Science of Evaluation: A Realist Manifesto' (5).

Demi-regularity

"Demi-regularity means semi-predictable patterns … The term was coined by Lawson (1997), who argued that human choice or agency manifests in a semi-predictable manner - "semi" because variations in patterns of behavior can be attributed partly to contextual differences from one setting to another." (6)

Lawson's interest was with long-term social and economic change. He was concerned to distance realism from the idea that there are 'laws' of social evolution *and* from the alternative that there is nothing but patternless fluctuation. This perspective sits comfortably with the realist notion of program effectiveness. What should not be anticipated is the discovery of intervention panaceas, nor will outcomes be complexly haphazard. There will be some patterning. We should be able to discern broad lessons on for whom, in what circumstances and in what respects an intervention is more likely to succeed and these 'demi-regs' are the basis for decision support in the policy making process.

Hypothesis

"A hypothesis is a logical supposition, a reasonable guess, an educated conjecture. It provides a tentative explanation for a phenomenon under investigation." (7).

Hypotheses can be developed and used at many levels in realist research – for example, hypotheses about the main ideas in program theory, about mechanisms, about the aspects of context that will influence whether and how mechanisms work.

Mechanism

There are many definitions of mechanism. What they all have in common is that mechanisms generate outcomes. Examples include:

"Mechanisms are the agents of change. They describe how the resources embedded in a program influence the reasoning and ultimately the behaviour of program subjects." (p13)(5)

"…mechanisms are underlying entities, processes, or structures which operate in particular contexts to generate outcomes of interest. There are three essential clues located in a ''realist'' reading of mechanisms.
These are that:
1. Mechanisms are usually hidden;
2. Mechanisms are sensitive to variations in context; and
3. Mechanisms generate outcomes." (2)

In their book 'Realistic Evaluation' Pawson and Tilley (3) provide explanations for a number of aspects of 'mechanism':

- a generative process which creates or constitutes a regularity (p 67);

- which is located at a different 'layer' of social reality than the regularity it explains (p 67);

- an "underlying" process, which cannot usually be directly observed (p 65);

- operating at both micro (individual) and macro (social/structural) levels (p 65);

- involving both people's choices (agency) and "the capacities they derive from group membership" (structure) (p 66);

- demonstrating "how program outcomes follow from the stakeholder's choices (reasoning) and their capacity (resources) to put these into practice" (p 66);

- "propositions about what it is within the program which triggers a reaction from its subjects" (p 66)

- a hypothesis about how programs work which always works as "a 'weaving process' which binds resources and reasoning together" (p 66);

- "not variables or correlates which associate with one another; rather we are trying to explain how the association *itself* comes about" (p 67): "the mechanism is responsible for the relationship itself" (p 68).

- "not expressible as properties of the individual" (p 68)

- "A mechanism is thus not a variable but an *account* of the make-up, behaviour, and inter-relationships of those processes which are responsible for the regularity. A mechanism is thus a theory… which spells out the potential of human resources and reasoning." (p 68).

Middle-range theory

A theory that is specific enough to generate hypotheses (for example in the form of propositions) to be tested in a particular case, or to help explain findings in a particular case, but general enough to apply across a number of cases or a number of domains.

Stratified reality

Realists believe that the world we live in is layered or stratified. All around us real social structures and systems exist and operate independently of our conception of them. Events can be seen, but the mechanisms that cause them are not readily observable because they exist at a different layer or strata. Understanding them requires theory and abstraction. This way of thinking about the world and how it operates implies that reality is stratified.

"The stratification of reality in the philosophical ontology of CR [critical realism] has two dimensions. The first is the … central distinction between the events that we can experience and describe, and the hidden, but nonetheless real, mechanisms behind them. The second dimension is that reality is assumed to consist of hierarchically ordered levels where a lower level creates the conditions for, but does not determine, the higher level. The distinction between the levels lies not in the entities, but in the generative mechanisms that operate at each level. It is not possible to reduce the causes of what occurs to one level to those of another level (whether lower or higher), because at each level something qualitatively new emerges … . These levels and their causes form an open, interactive world of things and contingent tendencies, which, according to CR, constitutes the proper object of scientific investigation." (8)

The most basic distinction within the notion of a stratified reality is that between 'structure' and 'agency'. Programs are attempts to induce social change and it is important to understand how these different strata play a part in producing social transformation. A grand attempt to understand the anatomy of societal change can be found in 'realist' methodology, most especially in the works of Archer (9). Her theory of 'morphogenesis' attempts to answer the age old sociological chicken-or-egg question about what comes first in propelling social change – is it 'structure' or is it 'agency'? Put simply, her answer is 'chickenegg'. People's immediate actions are shaped within social structures in which they sit – communities, organisations, legal systems, power relationships, etc. However, in a longer time frame, these structures themselves change as a result of the activities and choices of the historically situated individuals who make them up. Social change, in short, occurs through a never ending cycle: 'structural conditioning' shapes 'social interaction' which in turn shapes 'structural elaboration', which then provides 'structural conditioning' and so on, and so on. Realists thus suppose that change is something that no one steers. It happens perpetually and of its own accord whenever people and groups reflect on their own position.

What happens, as with policies and programs, when someone or some institution tries to steer change? The same dynamic persists. The intervention sits alongside other structures (organisations, communities) that clamour for the agents' attention. Some agents will choose to respond to the program and in so doing will subtly change the nature of the intervention and its place in social structure. Realist evaluation and realist synthesis attempt to retain this idea of the stratified and temporal unfolding of change. Programs and their effects cannot be considered in isolation from the rest of society. They are shaped by and shape history.

Substantive theory

'Substantive theories' are existing theories within particular disciplines. They may be used to help understand interventions. For example, in the social sciences theories may deal with topics such as

'cognitive development', 'deviance control', 'incentivisation' or any of the wider ambitions of interventions.

Theory

There are multiple definitions for the word 'theory'. One simple definition is that, "A theory is an attempt to organize the facts – some 'proven', some more conjectural – within a domain of inquiry into a structurally coherent system." (10) For a discussion of the different types and roles of theory in realist synthesis, see section 2.2 above.

Section 3. Focussing reviews

> **Learning objectives for this section**
>
> - Explain the importance of focussing a review
> - Know what constitutes good practice when focussing reviews
> - Describe the steps that may help to focus your review

3.1 Why focus reviews?

A realist research question contains some or all of the elements of 'What works, how, why, for whom, to what extent and in what circumstances, in what respect and over what duration?'. Realist analysis is then used to answer the question. 'Realist analysis' requires the application of a realist philosophical 'lens' to data; it seeks to analyze data using realist concepts. Specifically, realism adheres to a generative explanation for causation - that is, an outcome (O) of interest was generated by relevant mechanism(s) (M) being triggered in context (C). It is this process that above all else distinguishes a realist review from all other review types.

As with any other review method, realism has to wrestle with complexity. The problems that interventions hope to deal with are complex and multifaceted. Programs themselves are complex and adaptive. The situations in which programs are inserted are complex and changing. It is impossible for any review method, realist or otherwise, to be comprehensive in covering all the contributory processes and contingencies. This brute fact sometimes runs against the grain for reviewers who bring to the exercise the idea that they are providing an 'overview' or that they operate with a census of 'all' relevant primary studies. Even if one is undertaking a commissioned review and that commission points the reviewer to a well circumscribed family of interventions of type 'X', the questions that could be asked are still infinite. It is always necessary to 'focus'. It follows that no one should be unduly defensive about the fact that a realist review will provide partial knowledge. The key is to prioritise and make the chosen lines of investigation absolutely clear.

Because a realist synthesis will generate a large number of avenues that might be explored and explained, and because resources and timescale are invariably finite, it may be necessary to 'contain' a review. Many different aspects of a realist review might need to be focussed. Examples of *how* a review might be narrowed include:
- the question(s) to be answered (refining from broader to narrower)
- the aspect of program theory to be investigated
    - a sub-set of programs within a program family (e.g. routine screening rather than all health screening)
- scale of review (e.g. focus on particular countries in international development reviews, or cultures, or timeframes),
- the extent to which the review aims to be comprehensive
    - rapid review – using realist analytic processes within a more limited literature set
    - systematic review – aiming to include all evidence on the topic

Focussing may also take place at different time points in the review process. Not all eventualities can be anticipated at the start of a review and the theory 'axe' may be wielded too hastily. There is a normal process in any inquiry whereby new, unforeseen aspects are uncovered whilst one 'reads into' a project. The proper rhythm of inquiry is thus for its potential scope to widen before a well-

informed choice can be made on how and when to narrow it. Examples of time points *when* focussing may be needed include:

- o when negotiating the research project or funding contract;
- o while writing and negotiating the research protocol (where required for funding projects)
- o when an advisory group is established
- o when content experts are consulted
- o when it becomes clear how much evidence is available for particular aspects of the question;
- o when evidence suggests new pathways that could be explored

It is entirely legitimate for the synthesis' objectives, question and/or the breadth and depth of the review to evolve or be refined as the review progresses.

3.2 Quality standards for focussing reviews
When focussing reviews, we recommend the standards in Table 1 should be used.

Table 1: Quality standards for focussing the review

	Inadequate	Adequate	Good	Excellent
The review question is sufficiently and appropriately focussed.	The review question is too broad to be answerable within the time and resources allocated. There is no evidence that progressive focussing occurred as the review was undertaken.	Attempts are made by the review team to progressively focus the review topic in a way that takes account of the priorities of the review and the realities of time and resource constraints. Attempts are documented so that they can be described in publications as appropriate.	Adequate plus: The focussing process is iterative. Commissioners of the review are involved in decision-making about focussing. Decisions made about which avenues are pursued and which are left open for further inquiry are recorded and made available to users of the review.	Good plus: The review team draws on external stakeholder expertise to drive the focussing process in order to achieve maximal end-user relevance.

3.3 Examples from the literature

Case study 1

An example of (initially inadequate) focusing can be found in an early review undertaken by one of us: 'Internet-based medical education: a realist review of what works, for whom and in what circumstances' by Wong et al. (11). In this realist review, the stated focus of the review was:

> "…to [a] explain what sort of Internet-based medical education 'works', for whom and in what circumstances, …"

Whilst this review objective encompasses the expected elements of a realist question, it is rather broad – and potentially too broad to be answerable, even in the context of this review, which was as a doctoral thesis. A lesson here is that even in the relatively long time frames allowed for a thesis (in comparison, for example, to the tighter timescales of commissioned reviews), the temptation can be to be too ambitious in the scope.

Focussing in this review only occurred, much later on during data analysis and was driven more by chance than planning:

"As the review progressed we became aware of … the emergence of two prominent demi-regularities [that] prompted us to narrow our review focus to the two candidate theories discussed below."

Whilst such a strategy does mean that the review is focussed on what was present within the included studies, there are potential problems that might impede progress. For example:
- it is theoretically possible that there are no identifiable demi-regularities in the included studies
- there may be many demi-regularities and it is not clear which are more prominent and should be pursued (in preference to others)
- the focus chosen by concentrating on "prominent demi-regularities" may not be important to the potential users of a review's results.

With the example above, if we now 'grade' it against the quality standards set out in Table 1, the focussing strategy used can at best be described as 'Adequate'. Adequate because progressive focussing has occurred and it is reported. One aspect of this review that at least tries to take into account the needs of other stakeholders can be found thus:

"Several previous systematic reviews and two meta-analyses have compared the efficacy and utility of Internet-based education with conventional teaching methods or no teaching [References x2]. Two main questions face researchers in this field: efficacy (can Internet-based medical education work, and if so what is the 'effect size' compared to conventional teaching?) and effectiveness (under what real-world circumstances does it actually work, and how might its impact and cost-effectiveness be maximised?)."

Use is made of what is already known about the review's topic area to provide a very rough direction of focus. More iterative focussing, (for example) by consulting other educators on what they perceive their 'burning' unanswered questions to be might have helped improve the review further.

Case study 2

Other reviewers have approached the issue of focussing in different ways. In 'Implementing successful intimate partner violence screening programs in health care settings: Evidence generated from a realist-informed systematic review', O'Campo et al. have set out to uncover, "…why and how universal IPV [intimate partner violence] screening programs in health care settings are effective." (12). Two strategies are used to focus the review:
- the authors construct a program theory for IPV (Figure 1, page 856). This is then used to explain their initial focus – only on the initial steps of the IPV program theory, namely "…screening and risk assessment and identification of IPV Victims."
- the literature is consulted to check where most interest lies in IPV screening programs and this is identified as being in routine screening.

Case study 3

In Jagosh et al.'s review 'Assessing the outcomes of participatory research: protocol for identifying, selecting, appraising and synthesizing the literature for realist review' (13), the review team involves knowledge-user partners very early on to assist in formulating their review question and focus:

"Through the initial funding application process, the research questions were developed by the core group and sent to the [knowledge-user] partners to further define the aim of the proposed review according to their experiences and the priorities of their organizations"

In summary, reviewers should try to iteratively refine the focus of their reviews. This process should be planned from the outset and negotiated with funders, so that they are aware of, and able to participate in, key decisions that will structure the outcomes of the review.

3.4 Learning activity
This activity is designed to provide practice in refining and focussing a realist review question.

Imagine that you have been asked to undertake a realist review to inform a new government policy to promote healthy eating in adults. Your colleagues propose that the review question should be: "What interventions promote healthy eating in adults?"

As a learning activity, you might like to try the following tasks:
- firstly, rewrite this question in realist terms;
- secondly, propose a first *refinement* to *focus* this question more narrowly; and
- thirdly, make notes about how you might go about further refining the question at particular later stages of the review.

3.5 Reflection activity
How a review is focussed will depend on a range of issues (for example, your research question, resource allocation, funder's expectations, end users' needs and so on). It is therefore impossible to be prescriptive and restrictive on what must be done. However, this does not mean there are not strategies that may be employed to start this process. To assist on this front, we have developed a list of questions a reviewer / review team might like to ask themselves: These questions are based on the quality standards in Table 1 and are listed in Box 2. We suggest that a reviewer might like to go through the questions in Box 2 to work out if the questions are relevant to their review and then how each question might be addressed.

Box 2: Questions to assist the focussing process in realist syntheses

- Can you complete your review within the time and resources allocated?
- Have you discussed the need to focus your review with (where relevant):
  - your supervisor?
  - within your review team?
  - your funding body / commissioners of the review?
  - potential users of your review?
- What processes will you develop and put in place to focus your review? For example:
  - 'What' will you focus?
  - 'When' will you do your focussing?

Section 4. Program theory

Learning objectives for this section

- Explain the importance of program theory
- Describe how program theories are developed
- Set out the steps needed to develop and use program theory
- Distinguish what constitutes good practice in using program theory

4.1 Developing and refining realist program theory

"*Intervention are theories*" (1)

Realist synthesis has most often been used to make sense of complex interventions. These interventions or programs often have multiple components (which interact in non-linear ways), outcomes (some intended and some not) and long pathways to the desired outcome(s). The term 'program theory' refers to an abstracted description and/or diagram that lays out what a program (or family of programs or intervention) comprises and how it is expected to work (also see section 2.2 and 2.3).

Program theory serves two main functions in a realist synthesis. The first is to 'sketch the terrain' that will be investigated, and in the process to assist in refining the elements and scope for the review. The second is to provide a structure for review findings.

There are many forms of program theory (see Funnell and Rogers (14)), but the general idea is to identify and map out:
- o the key components (functions, strategies or activities) of the program;
- o the outcomes is the program intended to generate;
- o the components that contribute to particular outcomes. In some programs (but not all) it's useful to develop a rough sequence in which things need to happen, or a rough hierarchy of outcomes, in order to develop a sense of how the program is expected to work.

This can be done in a workshop with stakeholders, by reviewing program documentation (policy documents, funding applications, program descriptions and so on) or by reviewing a small selection of literature about the program type.

In multi-faceted (or multi-component) programs, where each element may trigger several mechanisms and various mechanisms will be affected by different contextual factors, the potential extent and complexity of the program theory can make the task seem overwhelming – or at least, it can sometimes be unclear where to start.

This is one of the stages where the previous topic (Section 3: Focussing reviews) is important. There are a number of ways to go about focussing the theory development stage.

One is to concentrate on specific outcomes of interest and then work backwards and 'outwards' from the outcome of interest to construct an initial rough theory.

Another is to spend a limited amount of time (perhaps a few hours) developing a very rough sketch of the program and then ask: Is there a particular component of this program on which everything else hinges? One aspect of the program that, if it fails, might bring the whole program down?

For example, programs which work by engaging a smaller group to pass on information or expertise to a wider group (peer education programs, training teachers to work differently with students, training managers to work differently in their staff teams) all share a common potential 'breaking point'. If the first group does not successfully engage the second group, the program will not 'work'. If a 'breaking point' exists, that might make an appropriate focus for thorough development of program theory.

A third way is to identify the areas of a very rough program sketch are already well understood, and to pick an area of the program theory that is not yet well understood, and focus attention there.

Once the focus is selected, the rough theory needs to be developed. This is not an easy task and alas, mostly goes unreported in publications. The realist reviewer above all else needs to be curious and critical. A useful heuristic is keep asking 'Why?' and 'How?' questions and to only stop when a sufficiently coherent and plausible set of theories has emerged.

For instance, when investigating a policy document, one will often be confronted with a rather glossy set of expectations that if 'X' is put in place 'Y' will follow. The document may be read in several ways:
   1) Uncritically, much in the way that the author intends you to read it, where the relationship between X and Y seems obvious, automatic and sensible.
   2) Cynically, where you delve immediately for the political ideology that gave birth to the program and the benefits to power-holders that may follow from it.
   3) Critically, where you interrogate what the policy architect has said (or failed to say) about how X will relate to Y. You should always put some basic questions to the absent author – 'How does the program work? Why is it that you suppose X should bring about Y?'

Option 3 above describes the realist mind-set. You may be lucky. You may immediately find the odd explanation or two. The document might go on to suggest what it is about the proposed intervention that will make a difference to individual behaviour or community life. The author might well have made a credible link between the resources on offer in the intervention and the reasoning that follows in the minds of program subjects. Rarely will the notion of 'mechanism' be employed but, nevertheless, some nascent program theories may be evident.

However, if you keep on reading, a wider set of justifications for the program will unfold. They may well involve alternative and competing explanations for mounting the program. If you keep pursuing the 'why?' question a broader set of explanations will build. If you read other materials (those produced a little further down the policy chain), these program theories will become more detailed. Program practitioners are often deeply sensitive to the key realist question – what exactly is it about their intervention that makes a difference to participants in their program. Dig deeply here. After a while your reading will more than likely take you to oppositional accounts. For every policy protagonist there is an antagonist. Again, avoid purely ideological critiques (option 2) in favour of relentless chasing of the 'why?' question (option 3). The literature will often suggest reasons why the program resources might be misinterpreted, reinterpreted or simply ignored by the some members of target community. All programs have unintended consequences and here are the program theories that might begin to account for them.

A critical reading should continue along these lines. With increasing familiarity with the literature you (the reviewer) will be able to see similarities and differences in the emerging program theories. Realists are fond of saying that there is nothing new under the sun in the world of intervention theory and you may well be able to embed your emerging set of program theories into previous policy analysis. Academics are great classifiers and typologists and you may be able use their ideas to bring some order to the chatter of program theories you have uncovered. So begins the journey of program theory development.

The second main function of program theory is to provide a structure for review findings. Realist reviews bring together diverse sorts of evidence from diverse sorts of research. The 'nuggets' of evidence need to be aligned with the particular elements of the program theory to which they are relevant. The evidence will support some parts of the theory, refine other parts, and refute some parts all together.

Once the evidence is aligned, it is necessary to then 'step back' and synthesise it – to produce a refined theory that provides the portable lessons for translation to other circumstances.

The ideal theory resulting from a completed realist synthesis would consist of:

1) An outline of the contexts in which, populations for which, and main mechanisms by which, particular outcomes are achieved (that is, a description of the CMO configurations identified through the review);

2) One or more middle-range theoretical explanation of how and why particular mechanisms generate certain outcomes within certain contexts; and

3) A middle-range theoretical explanation of the *pattern* of outcomes found – why the pattern of CMOs looks the way it does. This usually draws on formal theory in the domain in which the review is being carried out.

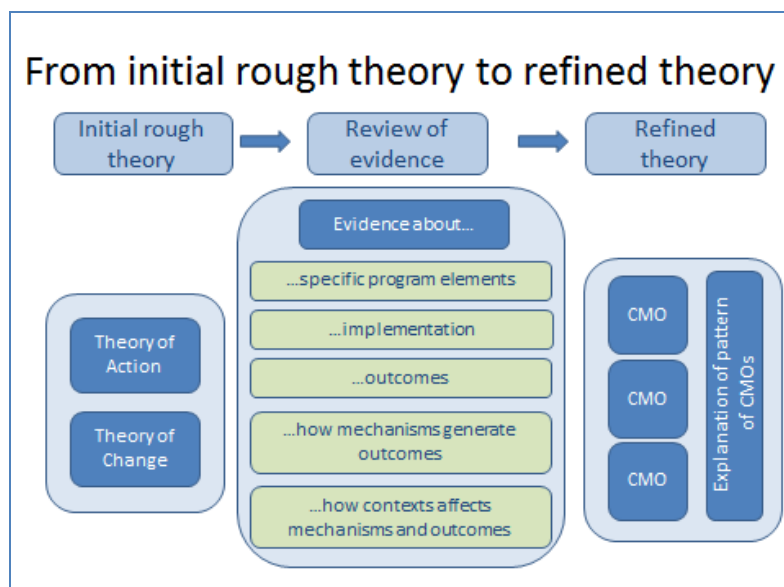The whole process of moving from initial rough theory to refine theory is summarised in Figure 5.

Figure 5: The process and sources of evidence used to refine theory

4.2 Quality standards for program theories

A program theory has an important role in a realist review and efforts to construct one should begin early in a review. Invariably any initial attempts at such a task will be tentative and will need to be progressively refined as the review progresses and understanding about the topic under study grows. For this topic area, we suggest that quality is defined as set out in Table 2.

Table 2: Quality standards for constructing and refining a realist program theory

	Inadequate	Adequate	Good	Excellent
An initial realist program theory is identified and developed.	A realist program theory is not offered or A program theory is offered but is not converted to a realist program theory at any stage of the review.	An initial program theory is identified and described in realist terms (that is, in terms of the relationship between contexts, mechanisms and outcomes). The refined theory is consistent with the evidence provided.	Adequate plus: An initial realist program theory is identified and described at the outset. The theory is refined iteratively as the review progresses.	Good plus: The relationship between the program theory and relevant substantive theory is identified. Implications of the final theory for practice, and for refinements to substantive theory where appropriate, are described. The final realist program theory comprises multiple context-mechanism-outcome configurations (describing the ways different mechanisms fire in different contexts to generate different outcomes) and an explanation of the pattern of CMOs.

4.3 Examples from the literature

Case study 1

The implications of not developing a program theory as a 'road map' are illustrated in the review by Daykin et al. (15). In this review, the reviewers are clearly dealing with a heterogeneous complex intervention – Patient Advice and Liaison Services (PALS) in England, United Kingdom – which has multiple components, long implementation chains and multiple intended outcomes. The question used to guide their review was:

> "*What context and mechanism factors have been identified as leading to favourable/unfavourable outcomes in patient and public involvement in NHS services?*"

Perhaps a first difficulty, in relation to program theory, lies with the structure of this question itself. It asks for context and mechanism factors to be identified, but it does not explicitly ask for an explanation of how contextual factors affect the operation of mechanisms[1].

---

[1] There may also be a clue in the term 'mechanism factors'. In realist theory, mechanisms are not 'factors' in the traditional sense of 'mediating and moderating factors', but causal processes.

Daykin et al. searched the literature on PALS and related interventions comprehensively and systematically and were able to identify a number of 'themes' in relation to each intervention type. For example, for PALS specific literature, they identified four themes: Resources for PALS; Accessibility of PALS; Priorities for PALS users; and Organisational issues (pages 61 to 62). Here perhaps lies a second issue in relation to program theory. It is not clear whether the 'themes' are intended to be contextual factors or causal processes, and so it is not clear what role they play in a realist explanation of outcomes.

There are instances where relationships between these elements are clear. For example, 'Accessibility of PALS' is reported to contribute to "take up" (an intermediate outcome) and one element of accessibility is 'physical location' (a contextual factor). However, 'take-up' may also be affected by a range of other factors, and those relationships are not clarified. Our suggestion is that developing a program theory could have assisted the team to explain the relationships and influences of between the various components of the program and the 'themes' and between those themes and the pattern of program outcomes in different contexts. It may also have assisted greater integration of the review's findings across PALS and other related programs. This is one of the goals of using program theory, rather than focusing on programs per se: it allows synthesis of learning across similar programs.

Case study 2

McMahon and Ward's realist review, 'HIV among immigrants living in high-income countries: a realist review of evidence to guide targeted approaches to behavioural HIV prevention', sought to develop a program theory from the outset (16). Their review's initial goal was to examine and make sense of the adaptations that are made to interventions trying to change behaviour of immigrants living in high-income countries so as to reduce HIV transmission. We have used this example to illustrate:
- use of the literature to inform the initial stages of theory building
- theory refinement – making it more specific and testable
- testing theory

This review also demonstrates the idea raised in 'focusing the review' above – that any single review will only be able to examine part of an entire program theory in any detail.

Scoping literature and initial theory development

When scoping the literature, the authors found that the predominant thinking behind adapting interventions related to making the interventions "culturally appropriate". Their initial attempts to develop a program theory therefore sought to explain and make sense of adaptations to enhance cultural appropriateness and the responses of individuals to those adaptations.

This provides a clear example of the process of focussing reviews, undertaken during literature scoping. As a result, other kinds of adaptations not related cultural appropriateness were deliberately excluded.

The authors used a "known set" of articles (which "… were comprised of studies known to the lead author and studies found in preliminary searches of databases in the initial stages of the review") to construct their initial, generic theory. The initial theory (see Figure 6) outlined the relationship between what was done (Intervention Adaptation Activity), the Theorised Mechanism of Adaptation Activity and the Anticipated Response to Adaptation Activity (outcome), and included 'counter-mechanisms' (what the authors have called Potential Resistance to Adaptation Activity). Whilst this initial program theory is portrayed in a linear fashion, they acknowledge that:

"It is important to note that in reality these 'chains' can operate in non-linear and unpredictable ways depending on the context [Reference x1]. Here, for simplicity, the implementation 'chain' is presented in a linear 'path' with the participant response and participant resistance represented as outcomes that point in different directions."
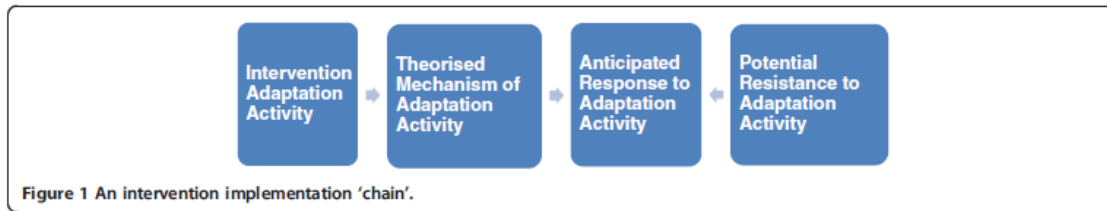

Figure 1 An intervention implementation 'chain'.

Figure 6: An intervention implementation 'chain' sourced from McMahon and Ward (16)

Theory refinement

The first step in theory refinement was to identify what sort of adaptations were made to interventions to make them culturally appropriate. McMahon and Ward grouped the adaptations they identified under the following headings: 'staffing', 'language', 'content', 'ethnic diversity', 'settings', 'community consultation' and 'priority setting'. For each adaptation group, they then inferred possible mechanisms – their initial list of inferred mechanisms being – 'authenticity', 'understanding', 'consonance', 'specificity', 'embeddedness', 'endorsement' and 'framing'. In effect, this process made the theory of cultural appropriateness more specific and testable (hence more middle-ranged). The authors provide a detailed account of how they went about constructing their initial generic program theory in their main article as well as in an Additional File (http://www.systematicreviewsjournal.com/content/supplementary/2046-4053-1-56-s3.pdf).

Testing theory

McMahon and Ward's review explicitly set out to test and refine their theory. A program theory was specified for each inferred mechanism, using the same structure as the initial generic theory. For example, the adaptation activity of "integrating cultural values and elements into the intervention context" works through the inferred mechanism of 'consonance'. The outcome of interest is that "… immigrants understand intervention in a symbolic sense – 'shared values'" and the important influencing context for 'consonance' is titled "balance 'old country' and 'new country values' (see Figure 7). Within each of their included studies, data (confirming and disconfirming) were sought to identify whether the inferred mechanism was in operation and to understand its relationships to context and outcome. The authors provide a detailed discussion of how they have tested and refined their program theory for four (out of the seven) inferred mechanisms in their paper.
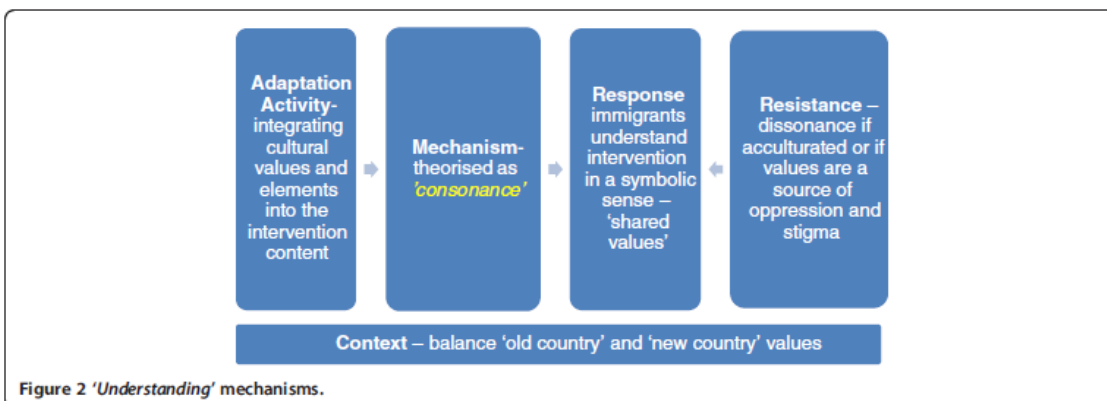

Figure 2 'Understanding' mechanisms.

Figure 7: 'Understanding' mechanisms sourced from McMahon and Ward (16)

26

4.4 Learning activity

This activity is designed to provide practice in developing initial program theory. Note that we have constructed the activity to operate 'backwards from intended outcomes'. It is equally possible to construct a theory 'forwards from intended activities'.

Choose a program or intervention in an area of interest to you. The program should be large enough and well-established enough to have generated both policy and research documentation.

Collect three documents about that program.
1. A formal policy or program description produced by the central agency for the program.
2. Implementation: either instructions or guidelines about how to implement the program at the local level or a piece written by a practitioner about their experiences in implementation of the program.
3. Research or evaluation report about the program.

Begin by reading the policy or program document. Note each of the outcomes that the program is **intended** to achieve. For this activity, constructing a diagram might be helpful.

(For an additional challenge, consider whether there is an 'order' to these outcomes. Do some need to be achieved earlier for later ones to be achieved? Are some to be achieved for individuals and others for communities? Using short names for each of the outcomes, organise the outcomes into a format that makes 'logical sense' to you. If you are constructing a diagram, put these at the 'end' of the diagram).

Next, select **one** of the outcomes that the program is expected to achieve. Ask yourself: how or why is this outcome expected to be achieved? Check sections or sentences that describe the rationale for the choice of strategy. Note too that the choice of strategy itself contains a clue (strategies or activities usually reflect the perceived 'solution' to the perceived nature of the 'problem'). If intended causal processes are not immediately apparent, additional questions might include: Who is expected to do what differently, in order for this outcome to be achieved? What different choices or decisions would they need to make, in order to do that? What will the program do or provide to assist them to do that?

Make notes or give titles to each of the main ideas about how or why the outcome is expected to be achieved (there's often more than one). If you are constructing a diagram, add each one separately. Draw arrows between the elements to show how they relate to each other and how they contribute to the outcome you are considering.

Pause and consider the diagram so far. There may be a number of elements comprising one chain that contributes to the outcome. See if you can give that chain overall a name – a noun or short phrase that describes the overall causal process at work (e.g. 'deterrence' in crime prevention; 'changing norms' in community development; 'the Hawthorne effect' in research). At one level of abstraction this chain constitutes a mechanism.

In order to describe the mechanism in more depth, ask yourself: how or why does each link in the chain work? (This has previously been described as working out 'what lies beneath the arrow'). These causal processes 'below' each link are also mechanisms, but viewed at a greater level of detail. (One of the characteristics of realist explanation is that one can 'zoom in' and 'zoom out' to different levels of abstraction, using the same basic explanatory framework). Continue asking 'how' and 'why' questions until a sufficiently coherent and plausible explanation has been developed.

Now read the other two documents you collected. Do they mention the outcome that you were considering? What do they tell you about it? More importantly, do they suggest changes to how do they change the description of how the outcome is achieved? Amend your sketch or notes to reflect these changes.

As you read the two additional documents, you might also start to collect clues about the contexts in which the intended processes do and do not work as intended. These too can be named and added to the sketch of program theory. This is the beginning of the process of theory refinement. In a review, evidence about each aspect of the initial theory is collected and the theory is gradually refined in the light of that evidence.

4.5 Reflection activity

To assist program theory development, we have put together a list of questions that review teams might like to ask themselves. These questions are based on the quality standards in Table 2 and are listed in Box 3. We suggest that reviewers might like to go through the questions to work out whether the questions are relevant to their review and then how each question might be addressed.

Box 3: Questions to assist constructing and refining a realist program theory

- Do you need to construct a realist program theory for each outcome of interest? If not, why not?
- What sources and resources (e.g. other researchers, experts, service users) will you draw on to help you develop your realist program theory?
- What processes will you develop and put in place to:
  - develop *and*
  - iteratively refine your program theory / theories?
- Are there existing substantive theories that are will help to inform your program theory / theories?
- What assumptions are built into the program theory?:
  - What assumptions are we (the reviewers) making?
  - What assumptions are there in the data?
  - Which ones do we need to challenge and why?
- What data, from where, might help to test and refine the theory?

Section 5. Developing a search strategy

Learning objectives for this section

- Explain the importance of developing a search strategy that meets your review questions' needs
- Define what constitutes good practice for developing searches for realist syntheses
- Develop and use a search strategy for realist syntheses

5.1 Search strategies suitable for realist syntheses

What constitutes 'the right evidence' is different in a realist synthesis than it is in other form of review. Data that may usefully contribute to a realist synthesis are:

- not decided by research type (e.g. randomised controlled trial (RCT)) but by relevance to the review question;

- not restricted to research into or evaluations of programs per se, but related to the program theory that underpins the program;

- not necessarily about the whole research question, but relevant to a sub-section of it;

- not necessarily drawn from a whole text/document, but from a sub-section of it relevant to a particular aspect of the review question;

- able to shed light on any aspect of C, M or O for any element of the theory;

- different for theory building (which does not need to be as rigorous) as opposed to theory testing (which needs to be sufficiently rigorous to support the conclusion being drawn on for the review).

There are two related processes when trying to find the 'right' evidence for realist reviews. One process is searching (this section) and the other is about making judgments on whether or not a data should be included (Section 6).

While initial searches are undertaken to develop theory and need not be as systematic, the searching processes used to test theory should be more systematic and transparent.

Searching for theory testing should be guided by the objectives and focus of the synthesis and revised iteratively in the light of emerging data. Data relevant to a realist synthesis may lie in a broad range of sources that may cross traditional disciplinary, program, and sector boundaries. The search phase is thus likely to involve searching for different sorts of data, or studies from different domains, with which to test different aspects of the provisional theory.

Search methods using forward and backward citation tracking may be particularly valuable in finding the documents necessary to refine and test provisional theories. Realist syntheses do not exclude sources solely on the basis of their study design; hence, 'methodological filters' (e.g. to identify randomised controlled trials) add little to the search and run the risk of excluding relevant papers (see Section 6 for an explanation about inclusion process and how this impacts on searching).

Searching is likely to need to be iterative because, as the synthesis progresses, new or refined elements of theory may be required to explain particular findings, or to examine specific aspects of particular processes. As new elements of theory are included, searches for evidence to support, refute, or refine those elements may be required. Imagine, for example, that the review of interventions to improve healthy eating by adults described in learning activity in section 3.4 above revealed that different cultural groups respond in different ways to a particular intervention. It may be necessary to undertake additional searches to gather more information on aspects of those cultures in order to theorise and explain this pattern of responses. A single pre-defined search is unlikely to be sufficient and may suggest insufficient reflection on emerging findings. The important judgement centres on whether or not the searches carried out in a review were likely to have located the sources needed for further theory development and/or testing.

A final point about search strategies for realist reviews addresses the balance between a search process being comprehensive versus theoretical saturation. Comprehensive searches set out to find (as much as is practically possible) each and every document on the topic of interest. For example, searches in Cochrane systematic reviews are designed to be comprehensive. In a realist synthesis, the ultimate product is explanatory theory. Providing the necessary evidence to demonstrate that a theory is coherent and plausible does not necessarily require the unearthing of every document about that theory. Thus in a realist synthesis, searching can be stopped if saturation is reached – that is a judgement can be made to stop searching if sufficient evidence is found such that it is reasonable to claim that the theory is coherent and plausible.

The extent to which theoretical saturation and comprehensiveness are goals in a particular review will be influenced by the specific question the review is addressing. A review that seeks to understand and build theory about the effects of context in relation to a specific program will need to be more rigorous about ensuring that research and evaluation documents about that specific program are included, so that important contexts are not overlooked. This is an example of a particular type of 'comprehensiveness'. A review that seeks to understand the operations of mechanisms at a broader level (such as Wong et al.'s review of smoking in cars, below in section 5.3) will prioritise theoretical saturation over comprehensiveness.

Many review teams have significant searching expertise. The key is to ensure that such expertise is adapted to the requirements for a realist review.

5.2 Quality standards for search strategy

For this topic, we suggest that quality is defined as set out in Table 3.

Table 3: Quality standards for developing a search strategy

	Inadequate	Adequate	Good	Excellent
The search process is such that it would identify data to enable the review team to develop, refine and test program theory or theories	The search is incapable of supporting a rigorous realist review. Common errors include: • The search is driven by a methodological hierarchy of evidence (e.g. privileging RCTs) rather than the need to identify data to develop, refine or test program theory/ies • The search process is not informed by the objectives and focus of the review • The database(s) selected are narrow in the subject matter that they contain (e.g. limited to specific topics rather than extending to social science, psychology etc.) • Searching is undertaken once only at the outset of the review and there is no iterative component	Searches are driven by the objectives and focus of the review. The search strategy is piloted and refined to check that it is fit for purpose. Documents are sought from a wide range of sources which are likely to contain relevant data for theory development, refinement and testing. There is no restriction on the study or documentation type that is searched for	Adequate plus: further searches are undertaken in light of greater understanding of the topic area. These searches are designed to find additional data that would enable further theory development, refinement or testing.	Good plus: the searching deliberately seeks out data from situations outside the program under study where it can be reasonably inferred that the same mechanisms(s) might be in operation.

5.3 Examples from the literature

Case study 1

Connelly et al.'s synthesis set out to "…present practice-relevant guidance on interventions to reduce at least one measure of adiposity in child populations that do or do not contain overweight or obese children."(17). They undertook a comprehensive search, drawing on the literature to guide

the development of their searching. They also searched more than one database and the reference lists of included trials.

> *"We used the search strategy and inclusion criteria described by Summerbell et al. [Reference x1] applied to the following electronic databases: Medline, Embase, Cinhal, PsycINFO up to 30 April 2006. This identified randomized, controlled trials or controlled trials of interventions to prevent overweight or obesity in populations that included non-overweight children with or without overweight or obese children. Trials had to include an outcome that measured an index of adiposity … We also searched the reference lists of included trials and published reviews for potentially relevant studies."*

Only one search was undertaken, only randomised controlled trials were included and the authors did not perform any preliminary searching to develop a program theory. The databases the authors searched were likely to contain additional relevant data, as the topic areas for their review are in health and psychology. Improvements the authors could have made to their search strategy to make it more likely to find data that would inform a realist synthesis include:

- Not using a methodological filter to decide which trials to include (i.e. consider including trials other then randomised controlled trials);

- Developing a program theory for interventions to reduce adiposity in children and basing their search strategy on finding the data needed to test and refine their program theory;

- Searching more than once, guided by the need to seek more data to test and refine that theory.

Case study 2

Morgan's review set out to, "…identify some of the underpinning factors that promote the development of evidence based health policy."(18). Only one health related database, PubMed, was searched. Data sources that contain data on health policy are likely to be found outside of the 'health' databases. A potential improvement to this search strategy would be to search in more than one database that is likely to contain relevant data.

Case study 3

In Jackson et al.'s review of the program 'Moving To Opportunity' (MTO), their iterative search strategy included three databases; four search engines; the reference lists of included articles. They also consulted external expertise. All of this was driven by the needs of the review:

> *"A ''snowball'' approach was used in which one reference led to others. Other evaluations were revealed through correspondence with Dr. Jeffrey Kling, one of the principal MTO researchers. … Additional literature was also accessed to help understand key concepts and issues raised through the review including housing theories, studies of poverty, housing and health, and social determinants of health."*(19)

Case study 4

In common with other review methods, piloting of a search is an important process and should be undertaken where possible. The purpose of any piloting is to check that the search has the ability to find documents that have already been identified. The process of using a 'known set' of documents to develop and refine a search is clearly described in McMahon and Ward's review (see the 'Systematic searching for primary studies' section of their paper) and can be a valuable way of testing if a search strategy is able to find relevant documents (16).

Case study 5

Searches can sometimes reveal a lack of evaluation studies on a topic area. In the review undertaken by Wong et al., the goal was to develop a, "…framework of threats to the program theory of public health legislation…" and then to test this framework using the case example of banning smoking in private vehicles carrying children. A threat they wanted to understand better was that of enforcement of such legislation. However, they report that they were:

> *"… unable to find any formal studies that evaluated the enforcement of smoking bans in vehicles carrying children, we deliberately chose to seek out studies which examined the closely related topics of the enforcement of cellular phone use and child restraints in vehicles. Our logic for searching in these areas were that they involved enforcement of 'in vehicle' behaviours that were potentially equally hard to enforce and (in the case of child restraints) involved the safeguarding of children."*(20)

The rationale behind such a search was based on the realist principle of causal mechanisms. The assumption was that it may be reasonable to extrapolate from studies of cellular phone use and child restraints in vehicles because similar mechanisms *may* be in operation. The Wong et al. paper provides only a brief explanation of whether a search directed by mechanism can yield helpful transferrable lessons. A more detailed analysis is provided in Pawson et al.'s paper, starting at 'Is the Law Enforceable?' on pages 536 to 541 (21).

Learning and reflection activities

The learning and reflection activities for this section have been combined with those of section 6 (see sections 6.4 and 6.5).

Section 6. Selection and appraisal of documents

Learning objectives for this section

- Explain how documents are selected and appraised for realist syntheses
- Define what constitutes good practice when selecting and appraising documents for realist syntheses
- Select and appraise documents appropriately for use in realist syntheses

6.1 Selecting and appraising documents for use in realist syntheses

Realist synthesis requires a series of judgements about the relevance and robustness of particular data items for the purposes of answering a specific question.

A wide range of documents may contain data that contribute to a realist synthesis. For example, outcome and impact studies, qualitative interviews, ethnography, questionnaire surveys, mixed-method case studies, and close reading of policies, business plans, websites, project initiation documents and 'grey literature' write-ups may all contribute in different ways to identifying and elucidating program theories. Within any document, different data may be relevant to different aspects of a review. Hence, rejecting a document on a global assessment of its methodological quality is illogical. Instead, inclusion and exclusion decisions are based on two criteria:

- *Relevance* – whether it can contribute to theory building and/or testing; and
- *Rigour* – whether the methods used to generate the relevant data are credible and trustworthy.

It can be difficult to 'whittle down' the number of documents eligible for inclusion in a review. However, it is necessary to make the process of selection and appraisal manageable and this includes developing inclusion and exclusion criteria for the initial literature search. The inclusion and exclusion criteria are guided by the focus of the review. As with other review methods, the titles, abstracts and or keywords of any documents included in from this initial search may then be screened for inclusion. Further stages of screening and discussion may be needed to produce a provisional 'short-list' of documents that then need to have the criteria of relevance and rigour applied to them. It will only be at this stage that data sources are analysed in detail. Thus, in practice, the selection and appraisal stage may need to run in parallel with the analysis stage.

The process and nature of selection and appraisal of documents for realist syntheses has two implications for researchers. Firstly, the 'short list' of documents that need careful scrutiny for inclusion (using the criteria of relevance and rigour) may be quite long – this may thus be a time consuming process and time and resources need to be put aside for this. Secondly, because potentially any part of any document might be relevant for theory testing and refinement, good data management processes are needed. For example, as a review progresses a review team may find they need to find data to help them test a modification they have made to their existing explanatory theory. Such data may exist in documents that had been excluded in earlier stages. So rather than re-run a search or develop a new one, reviewers may find it helpful to have set up a process whereby they can easily access the documents (or the citation of documents) that had initially been thought to be irrelevant. In other words, a good filing and tagging system helps.

Once the documents have been 'screened in', assessment of the data within the document begins. In any document, there may be several pieces of data that serve different purposes, such as helping to build one theory, refining another theory and so on. An appraisal of the contribution of any section of data within a document should be made on the same two criteria of *relevance* and *rigour*.

6.2 Quality standards for selecting and appraising documents

For this topic area, we would expect quality to be defined as set out in Table 4.

Table 4: Quality standards for selection and appraisal of documents.

	Inadequate	Adequate	Good	Excellent
The selection and appraisal process ensures that sources relevant to the review containing material of sufficient rigour to be included are identified. In particular, the sources identified allow the reviewers to make sense of the topic area; to develop, refine and test theories; and to support inferences about mechanisms	The selection and appraisal process does not support a rigorous and complete realist review. For example: • Selection is overly driven by methodological hierarchies (e.g. the restriction of the sources to RCTs to the exclusion of other forms of evidence) • Sources are appraised using a technical checklist for a particular method (e.g. assessment of quality for an RCT) rather than by making a defensible judgement on the relevance and rigour of the source • Selection and appraisal processes are overly restrictive and exclude materials that may be useful for a realist analysis • Selection and appraisal processes are not sensitive enough to exclude irrelevant materials	Selection of a document for inclusion into the review is based on what it can contribute to the process of theory development, refinement and/or testing (i.e. relevance). Appraisals of rigour judge the plausibility and coherence of the method used to generate data.	Adequate plus: During the appraisal process limitations of the method used to generate data are identified and taken into consideration during analysis and synthesis.	Good plus: Selection and appraisal demonstrate sophisticated judgements of relevance and rigour within the domain.

6.3 Examples from the literature

Case study 1

The use of checklists to judge the quality of an entire document can mean that documents may be excluded before they are even analysed for relevance and rigour. An example of this practice can be seen in Carr et al.'s report (22). Their health technology assessment review set out "... to identify, describe, classify and analyse the range of models developed to date for delivering health-related lifestyle advice (HRLA), or training, for effectiveness, mechanism of effect, cost-effectiveness, equity and acceptability in improving the health and wellbeing of individuals and communities, with particular reference to the reduction of inequalities in the UK." They undertook an exemplary and comprehensive search that identified a large number of potential documents for inclusion on this complex topic.

As can be seen from the figure 8 below, from 269 potentially eligible studies, 243 were excluded after a range of quality assessment checklists were used (these were for quantitative studies, the *Quality Assessment Tool for Quantitative Studies* and the Critical Appraisal Skills Program (CASP) checklist for qualitative research).
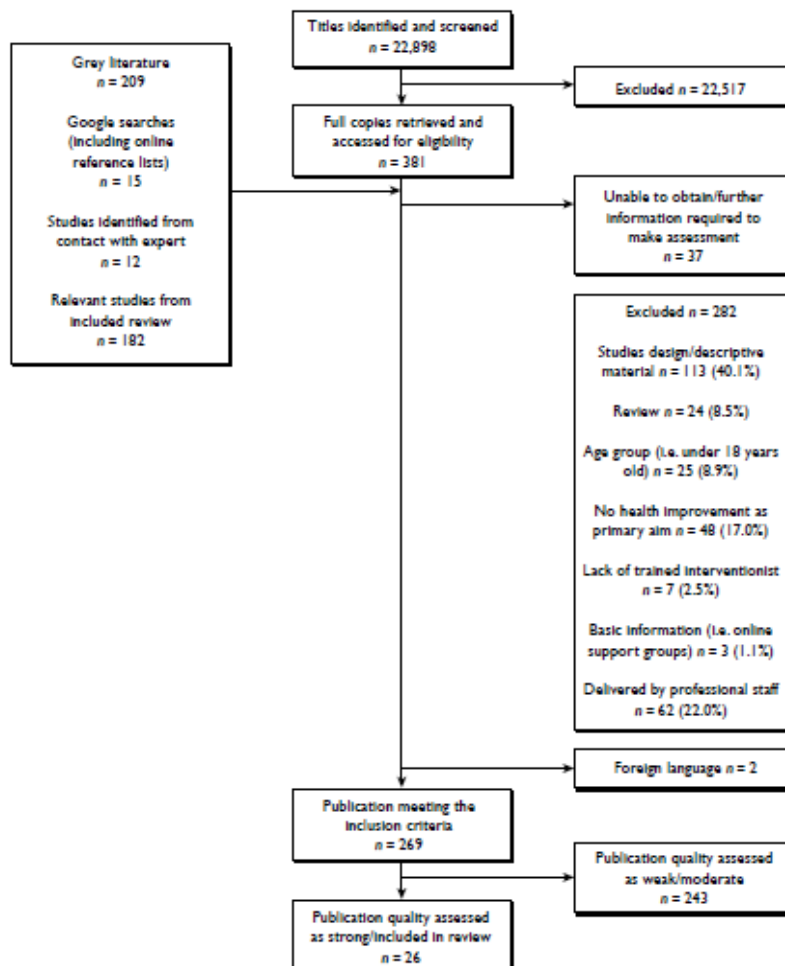

Figure 8: Study selection process figure sourced from Carr et al.

The quality checklists meant that of the 26 included studies, 23 were RCTs and there was one ethnographic study, one process evaluation and one controlled before and after study. The conclusions the authors were able to reach were mixed and inconclusive. The complex nature of the family of interventions the authors set out to study would of course produce a significant challenge to any review team. However we suspect that the overly restrictive inclusion process also contributed to limiting the data that could have been available for the realist review component of this review.

Case study 2

Selecting documents based on methodological hierarchies (in particular restricting sources to RCTs) restricts the range of evidence available, to the detriment of how informative the findings can be.

Kane et al. specifically wanted to explore what kind of a contribution randomised controlled trials (RCTs) can make to a realist review.

"Since randomised control trials (RCTs) have high internal validity, in this paper we review RCTs of interventions involving CHWs for improving child health in LMIC from a realist perspective with the aim to see if the RCTs can yield insight into the working of the CHWs [community health workers]."(23)

Their conclusion is self-explanatory and clearly highlights the limitations they found in only including RCTs in a realist synthesis:

"The RCTs under review offered a fair amount of information about the interventions, only some information about context - allowing us to formulate only generic hypotheses. Disentangling context from intervention elements was a daunting task, particularly when doing this across RCTs. …

Authors seldom described or discussed the mechanisms that explained their study outcomes. We realise that the RCT design, the exacting reporting requirements and word limits of journals, restrict authors from sharing all their operational experiences. In addition RCTs tend to report average effects and not differential effects of interventions, and less so of the context and rarely of the mechanisms triggered by their interactions. This makes the RCTs less useful for answering the questions regarding how interventions work. These generic hypotheses seem to be recurring in the literature, however they have not been explicitly tested across contexts."

Case study 3

The article, 'Known Knowns, Known Unknowns, Unknown Unknowns: The Predicament of Evidence-Based Policy' provides an opportunity to look in more detail at how the selection and appraisal of documents works in a realist review (21). This paper provides a more methodological analysis of the paper by Wong et al. mentioned above (see Case study 5 section 5.3). The purpose of their review was to develop a, "…framework of threats to the program theory of public health legislation…" and then to test this framework using the case example of banning smoking in private vehicles carrying children.

Their program theory was expressed in a series of questions (see Box 2 page 520 in the 'Known Knowns…' article).

Their second question and its sub-questions were:

"2. Is there likely to be public support for such a law?
  2.1. What is the overall magnitude of support for such a law?
  2.2. What are the levels of support among smokers?
  2.3. What is the motivation behind public support?
  2.4. Does endorsement depend on the extent and success of previous smoking bans?"

We can see that the data needed to address each of these questions is very different, because of the nature of the question asked. For question 2.1, data from surveys was used.

However, whilst relevant, such data are not without their problems - the issue of rigour being raised here. Pawson et al. offer warnings about the challenges of survey data:

"There are two familiar problems with such materials—attitudinal responses on health matters can be unreliable and the data, perforce, provide only a snapshot of

opinion at particular time and place. Survey responses can never be taken entirely at face value. Well-known technical problems exist due to the slipperiness of question wording. … questions carry subtle differences of emphasis that might shape the willingness to support a ban. Public compassion might well differ for ''children,'' ''children under ten'' and ''preschool' children, not to mention the ''elderly,'' ''pregnant women,'' ''nonsmokers,'' and so on. Probably, even more of a threat in the present case is the ''social desirability effect.'' Respondents, naturally enough, prefer to be on the side of the angels and thus often ''fake good'' when confronted by a stranger asking questions about sensitive topics [Reference x1]. Put in a nutshell, the problem is that smoking addicts, who suffer routine stigma on top of slow poisoning, may well choose to dissemble." (p528)

Thus while survey data may be relevant, caution needs to be exercised as to their credibility and trustworthiness. The implication is that inferences made on the basis of evidence have to take into account the rigour by which the data were generated. In the example we have used, the authors use the survey data not as definitive data that settles any argument, but more as part of an explanation building process:

"A sizable number of studies have shown significant levels of support for a ban in smoking in cars carrying children. …. In this case, the solidity of smokers' support is attested in further evidence on the grounds for that support, namely, their beliefs about the vulnerability of children, their sentiments of regret about taking up smoking and their acknowledgment that public sympathy for the smoking habit has declined under incremental legislation." (p531)

6.4 Learning Activity for Sections 5 and 6

*This activity is designed to provide practice in developing a search strategy **and** selecting and appraising documents for a realist review.*

Searching
Return to the rough program theory that you developed in learning activity 4.4 (or any other program theory that is of interest and available to you).

Choose one 'strand' (or sub-section) of that program theory.

Write a list of search terms that you could use to construct a search to test that strand of theory. Note that these terms will relate to the elements within the strand of theory, not just the name or type of program that the review deals with.

Selection for relevance
Make a list of the kinds of evidence that might be used to test that strand of theory. Might it be evidence from opinion surveys? Ethnographic studies? Program evaluations? Census data? Administrative data from programs? Align each type of evidence with the particular element of the strand of program theory for which it would be appropriate.

Assessment of rigour

For each of the types of evidence you could use, make a few brief notes about the issues that might affect data quality or rigour that you would need to take into account during appraisal of the documents.

6.5 Reflection Activity for Sections 5 and 6

To assist you in developing a suitable search strategy and in selecting and appraising documents appropriately, we have developed a list of questions a reviewer / review team might like to ask themselves. These questions are based on the quality standards in Tables 3 and 4 and are listed in Box 4. We suggest that a reviewer might like to go through the questions in Box 4 to work out if the questions are relevant to their review and then how each question might be addressed.

Box 4: Questions to assist developing a search strategy and selection and appraisal of documents

Developing a search Strategy
- How will you ensure that your search process is such that it would identify data to enable the review team to develop, refine and test program theory or theories?
- Is the necessary searching expertise available to you? If not, what will you do to remedy this?
- Will your search be piloted and refined?
- Will further searching be undertaken if more data are needed?

Selection and appraisal of documents
- Are *relevance* and *rigour* being used to guide the selection and appraisal process? If not, why not?

Section 7. Applying realist principles in analysis

Learning objectives for this section

- Apply realist principles in analysis
- Deal with differential patterns of outcomes in analysis
- Define what constitutes good practice in applying realist philosophy and realist logic in analysis

7.1 The implications of realist philosophy for analysis

Realism is not one unified philosophy of science, but most schools within it subscribe to a number of core assumptions. These include generative causation (that is, mechanisms generate outcomes), the contingency of mechanisms (that is, mechanisms are context sensitive), the existence of many mechanisms operating at all levels of reality (which may or may not act and may or may not interact depending on the context) and a stratified reality.

The notion of 'stratified reality' has two meanings in realist philosophy. The first is that all systems have sub-systems and all systems are part of larger systems. That is, all systems are embedded within other systems. Causation operates both upwards and downwards through these systems. The second meaning of 'stratified reality' lies in Roy Bhaskar's philosophical constructs of three domains in reality: the real (the domain in which mechanisms exist, whether or not they are observed and whether or not they are currently operating), the actual (events) and the empirical (what is observed by humans) (24). This second meaning informs realist understanding of how causation works and distinguishes between what exists and what we can know about it. The first meaning – open, embedded, interactive systems – is central to the process of analysis in actual cases.

These assumptions constitute a realist philosophical 'lens' and lie at the heart of a realist analysis. It is this that distinguishes a realist review from all other review types.

A realist philosophical lens has distinct implications for the nature and process of analysis in a realist review. These include:
- the purpose of a review
- the basic explanatory structure used within a realist review
- the analytic tasks that follow from that explanatory structure
- the logic of comparison within realist analysis
- the relationship of evidence to theory
- the conceptual tools used to deal with evidence, and in particular, apparently conflicting or contradictory evidence

Some of these have been discussed earlier in these materials but are summarised briefly here in order to demonstrate 'the logic of analysis' overall.

Purpose

The purpose of a realist review is explanatory, not simply descriptive. It seeks to explain how and why programs generate different outcomes in different contexts. The aim is to assist policy makers and practitioners to make decisions about whether to use particular kinds of programs to achieve

their intended goals in particular contexts and how to adapt them to those contexts to increase their chances of success.

Explanatory Structure

The basic explanatory structure in realism is generative causation - that is, an outcome (O) of interest was generated by relevant mechanism(s) (M) being triggered in specific context(s) (C). The implication for a realist review is that realist program theory takes this structure. It comprises sets of context-mechanism-outcome configurations (CMOCs). Each CMO can be 'read as a sentence' (i.e. 'In this context, that mechanism generates this outcome'.)

Outcomes can lie at different levels of systems, at different points along an implementation chain, and at different points in time (that is, they can be sequenced, with earlier outcomes necessary before later outcomes can be achieved). Wherever an outcome lies, it will be the result of one or more mechanisms operating within particular contexts.

Analytic tasks

The basic analytic task in a realist review is to **find and align the evidence** to demonstrate that particular mechanisms generate particular outcomes and to demonstrate which aspects of context matter. Working from the basic analytic structure described above, it follows that relevant mechanisms cannot be identified without reference to outcomes (mechanisms are what cause outcomes) and that relevant aspects of context cannot be identified without reference to mechanisms. An 'ideal' realist synthesis provides evidence for outcomes, evidence to support the existence of the hypothesised mechanisms, evidence that those mechanisms cause those outcomes, evidence that features of context exist and evidence that those features of context affect whether and which mechanisms fire.

It is because these different kinds of evidence are likely to be found in different kinds of sources that realist review casts its evidentiary net as widely as it does.

The logic of comparison in realist synthesis

In realist evaluation and realist review, the basic logic of comparison is "intra-program, inter-context comparison, on the basis of the program theory". In realist evaluation, the term 'intra-program' refers to the program being evaluated. In realist review, it refers to making comparisons within the pool of literature that has been included as relevant to the intervention(s) under study.

'Inter-context' means making comparing across contexts – be that population groups, cultures, geographic locations or organisational settings. The features of context that might be important are initially hypothesised to be important on the basis of program theory. Comparisons are then made on the basis of those features of context. Are the outcomes indeed different in these different sub-groups? Box 5 offers and example of these types of comparisons.

Box 5: An example of intra-program, inter-context comparisons based on program theory

<div style="border:1px solid #000; background:#d6e8ef; padding:10px;">

Intra-program, inter-context comparisons based on program theory: An example

Perhaps a program aims to increase employment outcomes for unemployed people. It is hypothesised to work by building bridging capital between unemployed people and employers. The relevant social theory is social capital theory; the intended mechanism is bridging capital. It might therefore be hypothesised that the program is most likely to be effective in urban population centres, where there are large numbers of employers and unemployed people who are not already linked. In country towns and villages, already strong existing social networks may mean that there is little chance for the program to build social capital and therefore little change in employment outcomes should be expected. In really remote regions where there are few opportunities to strengthen linkages, no change in employment outcomes would be expected. Based on this program theory, comparisons would be made across urban, rural and remote settings.

Evidence for increased bridging capital might include strengthened networks between unemployed people and employers, perhaps collected using social network analysis. Evidence for outcomes would relate to employment of program participants. Evidence that bridging capital operates as a mechanism in employment programs might be found in research about the ways in which people find jobs. Evidence that the bridging capital mechanism fired in this program and accounted for employment outcomes might include the proportion of newly-employed participants whose employment was with employers 'reached' through network-building activities. It might also come from interviews with employers, program participants and/or program staff. Both changes in the strength of networks and employment outcomes would be compared across urban, rural and remote settings to test the program theory about contextual differences affecting the bridging capital mechanism and the outcomes achieved.

The realist review might therefore ask: How much of this evidence is available, and in what sorts of research or evaluation might it be found? Is it possible to disaggregate program outcomes for urban, rural and remote settings? Are there case studies that examine in depth how the program worked in different settings?

</div>

The relationship of evidence to theory

Realists do not believe that there is such a thing as final truth, knowledge or 'Truth with a capital T'. All knowledge is partial; all theories remain theories that can be refined or disproved as new evidence comes to light.

In a realist review, the task is to align the 'nuggets' of evidence drawn from different sources against the element of program theory to which they refer, and then to ask: What does this evidence suggest about this aspect of our theory? Does it support it? Does it disprove it? Does it suggest an amendment to it?

Amendments to the theory will take the form of a new CMO configuration. Perhaps one or more of the hypothesised CMOs are removed altogether. Perhaps it has been demonstrated that a particular mechanism requires an alignment of two or more features of context. Perhaps the target group(s) for which an intervention is effective have been refined.

This stage of the work provides both summary and analysis, but it does not yet provide a full *synthesis* of the findings. The final step of the process involves one further level of abstraction – making sense of the pattern of findings. This is most commonly done using existing formal theory in the field in which the analysis is undertaken (e.g. some form of learning theory in education, sociological theory, economic theory, political science, organisational theory…). For an example, see Pawson's use of Merton's theory of reference group behaviour (in Chapter 7) to explain the range of outcomes in interventions that aim to 'name' and shame' (1).

A full realist analysis addresses both these levels and attempts to make sense of the relationship between these two levels. Syntheses that address only one level may also be considered realist syntheses, assuming that they apply and demonstrate application of a realist philosophy of science. The level(s) of analysis chosen will depend on the review's focus. The theories used may have been developed and/or refined from the data and/or be refinement of existing substantive theory.

Conceptual tools

Realist review cannot rely solely on quantitative analytic techniques to achieve its ends (although it can use them where the available data allows). Its purpose is explanatory and it needs to be able to explain the conflicting patterns of outcomes that are almost invariably found in reviews of social programs. Pawson's book 'Evidence Based Policy' (1) suggested what some of these tools might be:

- **juxtaposing** ("for instance, when one study provides the process data to make sense of the outcome pattern noted in another")

- **reconciling** (identifying differences which explain apparently contradictory sets of findings)

- **adjudicating** between studies (based on the quality of research);

- **consolidating** (building 'multi-faceted explanations of success')

- **situating** ('this mechanism in context A, that one in context B")

The key analytic process in realist review involves iterative testing and refinement of theoretically based explanations, using empirical findings in data sources. Reviewers may draw on any appropriate analytic techniques to undertake this testing. When reporting a review, explanation and justification for the choice of techniques should be provided.

7.2 Quality standards for understanding and applying the underpinning principles of realist reviews

Table 5 below sets out our recommended quality standards for understanding and applying the principles of realist review overall.

Table 5: Quality standards for understanding and applying the underpinning principles of realist reviews.

	Inadequate	Adequate	Good	Excellent
The review demonstrates understanding and application of realist philosophy and realist logic which underpins a realist analysis.	Significant misunderstandings of realist philosophy and/or logic of analysis are evident. Common examples include: • program/intervention activities or strategies are confused with mechanisms • no attempts are made to uncover mechanisms • outcomes are assumed to be caused by the program/intervention • relationship(s) between an outcome, its causal mechanism(s) and context(s) are not explained • some theory is provided but this is not explicitly linked to outcome(s)	Some misunderstandings of realist philosophy and/or logic of analysis exist, but the overall approach is consistent enough that a recognisably realist analysis results from the process.	The review's assumptions and analytic approach are consistent with a realist philosophy at all stages of the review. Where necessary a realist program theory is developed and tested.	Good plus: Review methods, strategies or innovations used to address problems or difficulties within the review are consistent with a realist philosophy of science.

7.3 Examples from the literature

Case study 1

One of the most difficult challenges in conducting a realist review is to be able to make a clear distinction between what is an intervention and a mechanism. Through the realist lens, interventions do not cause outcomes, mechanisms do. An example of this confusion may be found in the review by Daykin et al.:

> *"Mechanisms for effective PPI [patient and public involvement] are likely to include adequate provision of resources and information systems to monitor the impact of PALS [Patient Advisory and Liaison Services]. In addition, they may include appropriate models of PPI as well as support for deliberative processes and methods of addressing discursive strategies of engagement with service users' agendas."* (15)

Here, "adequate provision of resources" and "information systems" are (wrongly in our view) described as mechanisms. In realist logic, the provision of something is 'something an intervention does'. Provision of a resources does not cause change, but the reasoning in response to the resource provided does – the reasoning in response to a resource being one of the commonly used definitions of a program mechanism (3) (and see section 2.3 A Glossary of Terms).

Case study 2

Jackson et al.'s realist review on Moving To Opportunity illustrates how interventions, programs and strategies might be found within documents in a realist review (19). Our post-hoc analysis of this review (for educational purposes) can be found in Box 6. The square brackets indicate where we have inserted what we have identified as the mechanism in operation for the outcome of interest.

Box 6: A post-hoc analysis of Jackson et al.'s MTO review to illustrate realist concepts

A post-hoc analysis of Jackson et al.'s MTO review to illustrate realist concepts

Moving to Opportunity (MTO) was a program where the *"... central goal ... was to move families living in public housing in high-poverty neighborhoods to public or private housing in lower-poverty neighborhoods in order to provide 'better opportunities' (e.g., employment, education and housing) for families."* (p962)

In this program the main program strategy was to relocate families. However the program's designers also felt that additional strategies were need to enable relocation to occur, for example:

"In each city partnerships were formed with local public housing authorities which administered the rental assistance, and one or more local, non-profit counseling organizations which provided counseling on how to find rental units and work with landlords when appropriate." (p962)

MTO set out to change multiple outcomes for impoverished families. In Jackson et al.'s review, they focussed on *"... improvement in mental health for adult women, children and adolescent girls"* and provide *"... some insights about the mechanisms and contexts through which the intervention appears to have impacted mental health."* (p963)

The way in which the MTO program strategy changed context on one level is clear to see - it moved families from one context to another, a high-poverty neighbourhood to one with lower-poverty. However, it was not the move (the program strategy) that necessarily caused the mental health of families to improve. As Jackson et al. explain:

"... moving from high-poverty neighborhoods 'worked' for many adults, adolescent girls, and children in terms of mental health outcomes. Statistically significant improvements in mental health for adults (mainly women), female youth, and girls appear to have been related to moving to a better physical and social environment and especially reduced levels of violence in new neighborhoods. Indeed, moving appeared to create an immediate 'resolution' to the stress [mechanism] *that many adult women, female youth, and children were experiencing because of the violence within their previous neighborhoods. Children and female youth who moved appear to have been less afraid to leave their homes and/or spend time outside* [mechanism – reduced fear of violence]*, and were thus able to participate in a broader social life including after- school activities."* (p967)

Conversely for adolescent males, their mental health was not on the whole changed for the better:

"Moving to a new neighborhood was reportedly not as positive for many adolescent males as for females. The lack of significant positive changes in mental health for adolescent males in the MTO program may have been related to the fact that the fear of violence [mechanism] *was less significant for this sub-population, and not enough of a 'push' factor to warrant leaving valued relationships to people or place. At least some male youth may have moved involuntarily, resulting in no improvement in mental health in some evaluations, and decreased mental health in others. These outcomes may be related to the lack of social integration* [mechanism] *into new neighborhoods. Many adolescent males appear to have kept ties with their old neighborhoods. At least two factors are potentially at play: a desire to maintain existing relationships and contacts* [mechanism] *with familiar places in old neighborhoods and a response to feelings and experiences of discrimination* [mechanism]*, including racial discrimination, in new neighborhoods."* (p967)

As might be expected for such a complex intervention as MTO, other 'competing' mechanisms were at play and the mere act of moving (changing one context) resulted in the triggering of desirable mechanisms for some, and other less desirable mechanisms with unanticipated outcomes for others.

Case study 3

A common problem occurs when reviews stop at the level of description of the data or thematic analysis, rather than moving on to offer causal explanation. This does not mean that the findings of such reviews are not valuable, but it does imply that they cannot call upon the realist warrant of their findings being transferable because the same mechanism(s) may be in operation.

McLean et al.'s review "... sought to identify which TBI [traumatic brain injury] behavioral interventions may be most successful within the context of a nursing home." (25). Their realist review provided detailed summaries of the data on two themes - 'caregiving context of a nursing home' (with six subthemes) and 'behavioural interventions for individuals with TBI' (with five subthemes). From their data analysis and synthesis, a matrix was produced that matched the "fit" between 'Nursing home contextual factors' and 'TBI behavioral interventions' (see Table 1 on page 20).

The final model presented by the authors was their 'Practice model of behavioral caregiving' (see figure 1 on page 21). They summarise their model as follows:

> *"There are 3 key factors relating to context, social interaction, collaboration, and everyday activities and routines, and 3 key components of behavioral intervention, antecedent strategies, meaningful activity, and pharmacotherapy. These 6 elements are not isolated components of behavioral caregiving. They are connected to each other and to the 4 stages of the care planning process and the central players, the care aide and resident. This analysis led to a Model of Behavioral Caregiving for TBI (see Figure 1)."*

This analysis is not realist (in the sense used in these training materials) for two reasons. Firstly, whilst context is reported, it is not clear how and which outcome(s) they affect. Secondly, the analysis does not identify and test and mechanisms.

Case study 4

When analysing data using a realist lens, it can be a challenge identifying whether data should be classified and conceptualised as context, mechanism or outcome. Jagosh et al.'s review, 'Uncovering the Benefits of Participatory Research: Implications of a Realist Review for Health Research and Practice', demonstrates the value of clearly naming each element and the role it plays in the explanatory process (6).

Their findings indicated that the outcomes of using a participatory research approach could be explained by a theory of partnership synergy. Synergy was defined as "…combining the perspectives, resources, and skills of a group of people to "create something new and valuable together—a whole that is greater than the sum of its individual parts."

The authors provide numerous examples in their article. Here is one example, which also clearly labels context, mechanisms and outcome in their explanation of the data sample they provided:

> *"In Messengers for Health, recruitment to the advisory board was accelerated by the good reputation and connectedness of an initial community partner:*
>
>> *"The initial partnership in Messengers for Health proved critical in gaining the trust of extended community partners because A.K.H.G.M. [an initial community partner . . . and a parent who lost a child to cancer] is a member of the tribe, is fluent in her language, and is a well-respected individual in the community. At an interview training session one year into funding, community women stated that they were interested in the project because this person was involved. (Reference x1)"*
>
> *Although community members had reason to mistrust outside researchers (context), they felt willing to participate because they trusted (mechanism) the judgment of a well-respected and long-standing community member who was already involved. Trust, respect, and consequent synergy were established from this initial partnership, propelling subsequent stages of program planning (outcome)."*

Jagosh et al.'s review also illustrates another interesting methodological point. They demonstrate that the outcome of one CMO configuration can become the context for another outcome.

> *"Our realist analysis provided evidence that synergy has the potential to build over time when the partnership's activities repeatedly produce successful outcomes. This evidence was synthesized by identifying the outcome of one CMO configuration as forming part of the context in the next phase of research along a chain of planning and implementation stages—what we call a "$C^1M^1O^1$ -C^2" pattern, in which outcome^1 becomes a contributor to context^2. This demonstrates how partnerships alter elements of context over time, leading to enhanced outputs and outcomes. ...*
>
> *Barriers to conducting a randomized community trial included community resistance and the demands placed on them given the complex and structured research protocol (context^1). A decision was made at the outset to hire only African-Americans familiar with the community as project staff (context^1). Because of their prior history in the community, the project staff were glad to assist community members beyond the scope of the study (mechanism^1). This led to the staff's greater investment in the project (outcome^1 → context^2). The staff's deepening investment increased the community members' trust in the project (mechanism^2), resulting in closer interactions between the project staff and the community members (outcome^2 → context^3). Because of the greater sense of trust and safety (mechanism^3) due to the previously described trust-building processes, some participants revealed their desire to enroll in the project even though their children had not participated in the school-based asthma program (outcome^3). This led to new methods of recruitment being developed (outcome^3 → context^4), and new recruitment methods led to higher than expected enrollment (context^4). This added to the project stakeholders' desire to overcome attrition obstacles (mechanism^4). As a result, a new capacity to retain*

participants and prevent attrition in a complex clinical trial was created in a mobile population by addressing problems as they arose and through the project stakeholders' increasing sense of motivation, trust, and co-ownership of the project (outcome[4])." (p329-330)

7.4 Learning activities

These activities are designed to provide practice in identifying realist concepts and in realist analysis.

Learning activity 1

Choose any published realist review. Check that you understand the question that it aims to answer. Then turn to the summary of findings section.

To what extent is it possible to identify what has been classified as outcomes, what as context, and what as mechanism? If the authors have not labelled the elements of their findings in this way, are you able to do so from the text?

Now go back to the evidence within the review. Is evidence clearly aligned against a program theory? Is the theory clearly refined in the light of the evidence?

Imagine that you had to provide feedback to the reviewers. What recommendations about improving the 'realist nature' of their review might you provide?

Learning activity 2

This activity is designed to be undertaken by two or more members of a review team. (If you are not undertaking a realist review currently, you can still do this task by choosing a topic of interest to you and working with a colleague).

Select any three articles that will be included in your review. Independently from your colleague, read each of them, highlighting outcomes, mechanisms (or clues about potential mechanisms) and features of context that appear to affect outcomes.

Write or sketch a summary of your analysis based on these three articles, in realist terms.

Now meet with your colleague and compare findings. How are your interpretations of the data similar? How are they different? How would you go about seeking additional data to resolve any differences in interpretation that you might have?

7.5 Reflection activity

To assist you in applying realist principles in analysis, we have developed a list of questions a reviewer / review team might like to ask themselves. These questions are based on the quality standards in Table 5 and are listed in Box 6. We suggest that a reviewer might like to go through the questions in Box 6 to work out if the questions are relevant to their review and then how each question might be addressed.

Box 6: Questions to assist understanding and applying the underpinning principles of realist reviews

- Does the review team understanding underpinning principles of realist reviews?
- Does the review team know how to apply the underpinning principles of realist reviews in their analysis of their data?
    - If 'no' to either question, what steps are you taking to ensure you have sufficient methodological expertise? For example:
        - Recruiting realist review expertise
        - Organising training
        - Organising ongoing methodological support
- What opportunities have been built into the review process to enable the review team to discuss, analyse and/or synthesise the data together?

Section 8. Further reading and resources

VARITIES OF REALIST RESEARCH

Maxwell J. A Realist Approach for Qualitative Research. London: Sage; 2012.

McKelvey B. Toward a Campbellian Realist Organization Science. In: Baum J, McKelvey B, editors. Variations in Organization Science: In Honor of Donald T. Campbell.Thousand Oaks: Sage; 1999. p. 383-411.

Olsen W. Realist Methodology. London: Sage; 2010.

Sayer A. Method in Social Science: A Realist Approach. 2nd ed. Abingdon, Oxon: Routledge; 1992.

MECHANISMS

Astbury B, Leeuw F. Unpacking Black Boxes: Mechanisms and Theory Building in Evaluation. American Journal of Evaluation 2010;31(3):363-81.

Gerring J. Causal mechanisms: Yes, But...
http://people.bu. edu/jgerring/documents/CausalMechanisms_Extended pdf 2013
[accessed 2013 Jul 7]

Hedström P, Ylikoski P. Causal Mechanisms in the Social Sciences. Annual Review of Sociology 2010;36:49-67.

THEORY

Friedman V. Designed Blindness: An Action Science Perspective on Program Theory Evaluation. American Journal of Evaluation 2001;22(2):161-81.

Pawson R. Middle Range Theory and Program Theory Evaluation: From Provenance to Practice. In: Vaessen J, Leeuw F, editors. Mind the Gap: Perspectives on Policy Evaluation and the Social Sciences.New Brunswick, New Jersey: Transaction Publishers; 2010. p. 171-202.

Rogers P, Petrosino A, Huebner T, Hacsi T. Program theory evaluation: Practice, promise, and problems. New Directions for Evaluation 2004;2000(87):5-13.

Weiss C. Which links in which theories shall we evaluate? New Directions for Evaluation 2013;2000(87):35-45.

REALIST REVIEW

The main text for realist synthesis is **Pawson, R: Evidence-Based Policy: A Realist Perspective (Sage, 2006)**

For underlying principles and assumptions: see Chapter 2
For Pawson's critique of existing review methods: see Chapter 3

For the stages in conducting a review: see Chapter 4

For an overview: 'Realist Synthesis: an introduction' – available on line and free at:
http://www.ccsr.ac.uk/methods/publications/documents/RMPmethods2.pdf

For the origins and influences on Pawson's interpretation of realism, see Pawson, R. (2013) The Science of Evaluation: A Realist Manifesto (Part 1). Part 2 of this book provides Pawson's explanation of complexity from a realist perspective.

Rycroft-Malone R, McCormack B, Hutchinson A, DeCorby K, Bucknall T, Kent B, et al. Realist synthesis: illustrating the method for implementation research. Implementation Science 2012;7:33.

WEBSITES

The RAMESES Project
http://ramesesproject.org/

The realist hive
http://blogs.exeter.ac.uk/realisthive/

Theory-driven inquiry for health systems research
http://www.itg.be/internet/ds/tde/index.html

Realist synthesis: The website
http://www.leeds.ac.uk/sociology/realistsynthesis/

Section 9. References

(1) Pawson R. Evidence-based policy: A realist perspective. London: Sage; 2006.

(2) Astbury B, Leeuw F. Unpacking Black Boxes: Mechanisms and Theory Building in Evaluation. American Journal of Evaluation 2010;31(3):363-81.

(3) Pawson R, Tilley N. Realistic evaluation. London: Sage; 1997.

(4) Merton R. On Theoretical Sociology. Five Essays, Old and New. New York: The Free Press; 1967.

(5) Pawson R. The Science of Evaluation: A Realist Manifesto. London: Sage; 2013.

(6) Jagosh J, Macaulay AC, Salsberg J, Bush PL, Henderson J, Sirett E, et al. Uncovering the Benefits of Participatory Research: Implications of a Realist Review for Health Research and Practice. Milbank Quarterly 2012;90(2):311-46.

(7) Leedy P, Ormrod J. Practical research: Planning and design. 7 ed. Thousand Oaks: Sage; 2001.

(8) Wikigren M. Critical realism as a philosophy and social theory in information science? Journal of Documentation 2005;61(1):11-22.

(9) Archer M. Realist Social Theory. Cambridge: Cambridge University Press; 1995.

(10) Klee R. Introduction to the philosophy of science. Cutting nature at its seams. New York: Oxford University Press; 1997.

(11) Wong G, Greenhalgh T, Pawson R. Internet-based medical education: a realist review of what works, for whom and in what circumstances. BMC Med Educ 2010;10:12.

(12) O'Campo P, Kirst M, Tsamis C, Chambers C, Ahmad F. Implementing successful intimate partner violence screening programs in health care settings: Evidence generated from a realist-informed systematic review. Social Science & Medicine 2011;72:855-66.

(13) Jagosh J, Pluye P, Macaulay A, Salsberg J, Henderson J, Sirett E, et al. Assessing the outcomes of participatory research: protocol for identifying, selecting, appraising and synthesizing the literature for realist review. Implementation Science 2011;6(24).

(14) Funnell S, Rogers P. Purposeful Program Theory: Effective use of theories of change and logic models. San Francisco, CA: Jossey-Bass; 2011.

(15) Daykin N, Evans D, Petsoulas C, Sayers A. Evaluating the impact of patient and public involvement initiatives on UK health services: a systematic review. Evidence & Policy 2007;3(1):47-65.

(16) McMahon T, Ward P. HIV among immigrants living in high-income countries: a realist review of evidence to guide targeted approaches to behavioural HIV prevention. Systematic Reviews 2012;1:56.

(17) Connelly J, Duaso M, Butler G. A systematic review of controlled trials of interventions to prevent childhood obesity and overweight: A realistic synthesis of the evidence. Public Health 2007;121:510-7.

(18) Morgan G. Evidence-based health policy: A preliminary systematic review. Health Educ J 2010;69(1):43-7.

(19) Jackson L, Langille L, Lyons R, Hughes J, Martin D, Winstanley V. Does moving from a high-poverty to lower-poverty neighborhood improve mental health? A realist review of 'Moving to Opportunity'. Health & Place 2009;15:961-70.

(20) Banning smoking in vehicles with children: A realist review. University of East Anglia, UK: SAPC; 2010.

(21) Pawson R, Wong G, Owen L. Known Knowns, Known Unknowns, Unknown Unknowns: The Predicament of Evidence-Based Policy. American Journal of Evaluation 2011;32(4):518-46.

(22) Carr S, Lhussier M, Forster N, Geddes L, Deane K, Pennington M, et al. An evidence synthesis of qualitative and quantitative research on component intervention techniques, effectiveness, cost-effectiveness, equity and acceptability of different versions of health-related lifestyle advisor role in improving health. Health Technol Assess 2011;15(9):1-284.

(23) Kane S, Gerretsen B, Scherpbier R, Dal Poz M, Dieleman M. A realist synthesis of randomised control trials involving use of community health workers for delivering child health interventions in low and middle income countries. BMC Health Services Research 2010;10(286).

(24) Bhaskar R. A Realist Theory of Science. London: Verso; 1978.

(25) McLean AM, Koppang J. Behavioral caregiving for adults with traumatic brain injury living in nursing homes: developing a practice model. J Theory Construct Test 2010;14(1):17-22.