

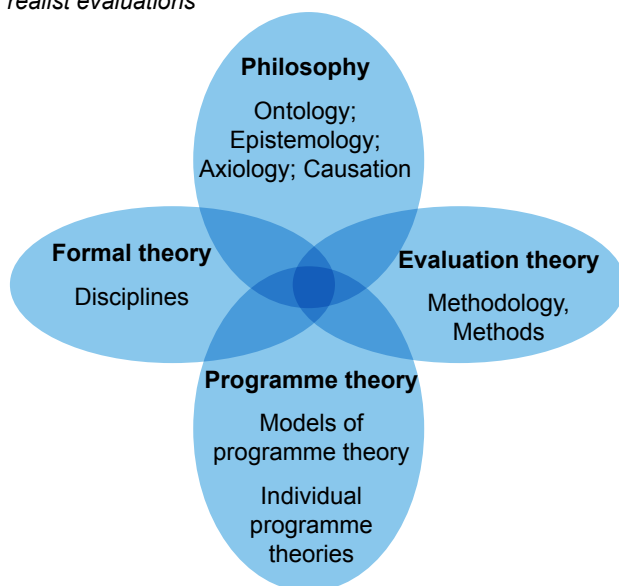
# “Theory” in realist evaluation

## The RAMESES II Project

Introductions to realist evaluation commonly start by noting that realist evaluation is one of the family of theory-based evaluations. But what exactly do we mean by ‘theory’ and how and why is it used in realist evaluation?

The term ‘theory’ means several different things in evaluation. Figure 1 presents four different categories of theory: philosophy, formal theory, evaluation theory and programme theory. Each of these has specific roles to play in evaluation practice.

Figure 1: Different types of theory that maybe relevant in realist evaluations



### Philosophy

The first type of theory is philosophy, which deals with the ‘fundamental natures’ of things at such a level of abstraction that the ideas cannot be disproved. Different branches of philosophy consider different things: ontology deals with the nature of reality; epistemology with the nature of knowledge; axiology with the nature of value. Different philosophies also make different assumptions about how things are caused. These

four questions – what we consider to be ‘real’, what we can know about it, how we might go about valuing it, and how we think programs or policies cause change – are fundamental questions for evaluators.

For realist evaluation, the underlying philosophy is a realist philosophy of science. (See ‘Philosophies and Evaluation Design’ in this series.)

### Evaluation theory

The second category is evaluation theory. These are theories about evaluation itself, relating to, for example:

- what evaluation is or isn’t (e.g.: is it different from research? In what ways and why does that matter?)
- methodologies (the practical application of epistemological assumptions. See ‘Philosophy and Evaluation Design’ in this series for more detail);
- methods for data collection, analysis and interpretation. (Different schools of evaluation have different theories about what methods should be used, how and why they should be done and why each methods decision matters);
- the roles of evaluators (Independent? In house? Participant evaluators? Providing recommendations or just findings? Why?),
- the political stances that should or should not be taken (ranging from the politics of empowerment evaluation to that of randomised controlled trials).

Evaluation theories help us to “make good judgments about what kind of methods to use, under what circumstances, and toward what forms of evaluation influence” (Marks, 2005), learn from the past, and understand and contribute to debates in the sector.

Theoretical assumptions are apparent even at the level of specific methods. For example, many of the statistical methods commonly used in evaluation assume linear causation, even though there is

increasingly widespread acceptance that causation is not linear. Evaluators have to be able to justify and defend the evaluation theories and perspectives that they select and use for particular evaluations.

For realist evaluators, the overarching evaluation theory will be realist evaluation; and that will hold implications for every aspect of evaluation design and for each of the issues in the dot points above.

## Formal theory

The third category of theory is ‘formal theory’ or ‘substantive theory’: the theories that operate in different domains or disciplines. Examples might include incentives theory in economics, attachment theory in human development or constructivist learning theory in education. Formal theories are often ignored in programme theory but probably should not be, for two reasons. The first is that programmes grounded in formal theories are sometimes more effective than those that are not. The second is that formal theories provide a bridge to a wealth of existing research and knowledge about a topic, which can inform several steps in many evaluation approaches: building programme theory, shaping questions for evaluations, interpreting findings and so on. In realist evaluation, formal theory is often used to identify mechanisms and features of context, and to explain how overall (often apparently disparate) sets of findings fit together. Theories written about a relatively specific phenomenon, such as Bandura’s Social Learning Theory or Merton’s Reference Group Theory, are usually more useful for this purpose than grand overarching theories such as Marx’s theory of capitalism.

## Programme theory

The final theoretical category is programme theory. This is what is usually referred to as being ‘theory’ in theory based evaluation. It is the description, in words or diagrams, of what is supposed to be done in a policy or programme (theory of action) and how and why that is expected to work (theory of change). Even here, traps await the unwary. Funnell and Rogers’ book *Purposeful Program Theory* identified at least 15 different models of programme theory and 22 ways it is described (Funnell and Rogers, 2011, especially Chapter 9). The models have different purposes, contain different categories, imply different data needs, and justify different sorts of evaluative decisions.

“Realist programme theory can be developed using relatively straightforward questions.”

For realists, programme theory will explain how and why different outcomes are generated in different contexts.

Realist programme theory can be developed using relatively straightforward questions. The table below demonstrates.

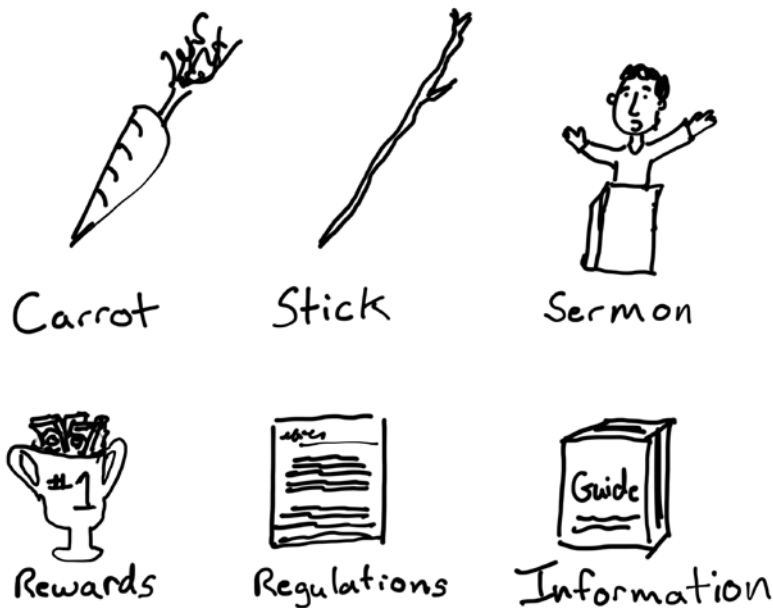
Realist assumption	Focus Questions	Realist programme theory
1. Programmes intend to cause a change.	What change (outcome) does it intend to create?	Identifies <i>intended</i> outcomes.
2. The programme intends someone to do something different, or differently, to cause the change.	<i>Who</i> does it intend to do <i>what</i> differently? (May be several groups).	Identifies whose decision-making should be examined.
3. Programmes provide resources or opportunities, or change environments, to enable the different choice/ behaviour.	What does it provide to enable that choice or behaviour?	Identifies the ‘resource’ (in ‘reasoning and resources’). Most programmes provide multiple resources each of which can trigger different reasoning.
4. Programme staff and participants make active choices and respond differently to resources.	How might different sub-groups of staff and participants respond to the resource?	Identifies various sets of ‘reasoning’ (in ‘reasoning and resources’). Contributes to identifying ‘for whom’ programs may / may not work.
5. Participants’ choices depend on/are affected by context.	What features of context affect how people respond to the resources? In what ways do those features affect responses?	Identifies what it is about context that matters. Contributes to identifying ‘for whom’ programmes may / may not work.
6. Context affects more than participant choices.	Which other features of context will affect whether & how the programme ‘works’?	Identifies features of (e.g.) implementation, organisation, capacity, history that affect whether or not programmes work.
7. The choices that ‘participants’ make lead to different outcomes.	What outcomes would be generated by different decisions?	Identifies a wider range of potential outcomes (and thus indicators). Can include negative outcomes.
8. To evaluate programmes, collect information about context, mechanism and outcome.	What information will be needed and could be collected about contexts? Mechanisms? Outcomes?	Identifies the data necessary to test the programme theory.

Answering these questions provides both an overall programme theory and information about the data needed to test it. The overall programme theory comprises a number of sub-theories, any combination of which might be the focus of a specific evaluation. It is never possible to test all the sub-theories in a programme theory in one evaluation.

Finally, a point to note on the term 'middle-range theory'. These are theories that "... involves abstraction, of course, but they are close enough to observed data to be incorporated in propositions that permit empirical testing." (Merton, 1967). This term is often used in realist research and at times is confused as being programme theory or formal theory – or an additional theory required as well as both of these. However, the

“The term “middle-range” is an adjective describing the level of abstraction of a theory. It is not another category or type of theory.”

term “middle-range” is an adjective describing the level of abstraction of a theory. It is not another category or type of theory. Apart from philosophy, any of the types of theory mentioned above can (and probably should) be written as a middle-range theory. CMOs should also be written at a middle level of abstraction: specific enough to clearly explain the phenomenon, and general enough to apply across cases of the same type.



## References

Funnell S and Rogers P (2011) *Purposeful program theory*. San Francisco: Jossey-Bass.

Marks M (2005) Evaluation Theory or What Are Evaluation Methods for?

[www.hfrp.org/evaluation/the-evaluation-exchange/issue-archive/evaluation-methodology/evaluation-theory-or-what-are-evaluation-methods-for](http://www.hfrp.org/evaluation/the-evaluation-exchange/issue-archive/evaluation-methodology/evaluation-theory-or-what-are-evaluation-methods-for)

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

---

## Project team:

Professor Trish Greenhalgh  
UNIVERSITY OF OXFORD  
[www.phc.ox.ac.uk/team/trish-greenhalgh](http://www.phc.ox.ac.uk/team/trish-greenhalgh)

Professor Ray Pawson  
UNIVERSITY OF LEEDS  
[www.sociology.leeds.ac.uk/people/staff/pawson/](http://www.sociology.leeds.ac.uk/people/staff/pawson/)

Dr Geoff Wong  
UNIVERSITY OF OXFORD  
[www.phc.ox.ac.uk/team/geoffrey-wong](http://www.phc.ox.ac.uk/team/geoffrey-wong)

Dr Gill Westhorp  
CHARLES DARWIN UNIVERSITY  
[www.cdu.edu.au/northern-institute/our-teams/603/5928#gillian-westhorp](http://www.cdu.edu.au/northern-institute/our-teams/603/5928#gillian-westhorp)

Dr Joanne Greenhalgh  
UNIVERSITY OF LEEDS  
[www.sociology.leeds.ac.uk/people/staff/greenhalgh](http://www.sociology.leeds.ac.uk/people/staff/greenhalgh)

Dr Ana Manzano  
UNIVERSITY OF LEEDS  
[www.sociology.leeds.ac.uk/people/staff/manzano/](http://www.sociology.leeds.ac.uk/people/staff/manzano/)

Dr Justin Jagosh  
UNIVERSITY OF LIVERPOOL  
[www.liv.ac.uk/cares](http://www.liv.ac.uk/cares)

This project was funded by the National Institute of Health Research Health Services and Delivery Research Programme (project number 14/19/19).

Professor Trish Greenhalgh's salary is part-funded by the Oxford Biomedical Research Centre, NIHR grant number BRC-1215-20008.

The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the HS&DR programme, NIHR, NHS or the Department of Health.

