Description:	A realist evaluation is focused on analysing the contexts in which activities generate particular outcomes. It is based on realist programme theory (what works for whom in what circumstances through what mechanisms?). The approach recognises that the context – social, economic, geographical, cultural, historical and other observable processes (e.g., class, family, gender, location), and entities (e.g., schools, institutions) - can have an effect on whether programmes work or not. The 'mechanism' by which a programme works depends on how members of the target group respond to the opportunity provided by the programme. Causation is not a linear process and depends on the interactions that take place in different contexts or systems.
Type of evidence:	OfS Type 2 (empirical).
Strengths:	A realist evaluation is concerned with attribution (in the sense of attributing results to causes). It combines aspects of process and outcome evaluation. It's a good approach when thinking about complex interventions, and can help evaluators to understand why results might differ. This approach is also good to identify the conditions in which a programme work best – and in terms of how transferable the approach might be to other contexts. A realist evaluation can be used to test and refine the programme theory. This means it is particularly useful when evaluating new initiatives. Furthermore, the approach can help to make recommendations about how to scale up or mainstream innovative pilot projects. In cases where a programme has had variable results, a realist evaluation would be useful to explain the underpinning factor affecting inconsistent results.
Weaknesses:	Realist evaluation is more concerned with the whys and wherefores rather than the overall/net effect or impact of a programme. Therefore, evaluation designs that 'control for' context with a view to exposing the 'pure' effect would be more appropriate if the focus is on analysing the 'net' effect (although might be limited in understanding how, when and for who the effects occur). A realist evaluation would be less relevant in the case of simple, small-scale activities where the output-impact process is very direct and already well understood.
Mixed Methods:	The evaluation seeks to measure relevant aspects of context, and the mechanisms, as well as the indicators of outcomes.
Expertise:	High.
Requirements:	The evaluation needs to be based on an explicit programme or policy theory, or the evaluation needs to develop one. Realist evaluation might require the evaluator to theorise the different outcomes that might occur in different contexts – which is different to the assumption with other types of impact which are based on changes at one level lead leading to further changes at a higher level irrespective of the context. Realist evaluations usually need a lot of primary data, which can be costly to collect and rely on experts – so are not cheap. Therefore, there needs to be a convincing case to justify the cost of a realist evaluation. The outcomes and impacts need to be well defined and appropriate data needs to be available to measure them. In order to look at the influence of context and mechanisms, the data collection and analysis requires a nuanced approach – i.e. disaggregation according to context or different target groups.
Ethical considerations:	As with any form of primary research with human participants, it is important that evaluations are conducted ethically. One of the challenges with realist evaluation is that it is a dynamic iterative process so it could be difficult to specify the precise methods at the start – for example, topic guides might not be known at the start, and the sampling might depend on how the hypotheses evolves. Pre-empting potential changes may be needed when seeking ethical approval. Plus, the evaluation needs to link data across different contexts – so there needs to be recognition of the implications for research

participants – i.e. data must be stored in such a way that linkages can be made, but extra care will be needed as this may increase the risk of identifying individual participants (one approach could be to anonymise the data first and then use an identifier to link to personally identifying data).

Work planning:

The main stages in a realist evaluation involve:

The first step is to clarify the aims of the evaluation. Usually this is to adapt the programme theory to better understand how, when and for whom good results can be achieved. It could also be to make recommendations about how to adapt the programme for a wider range of groups; or to recommend how the programme could be rolled-out in different contexts.

Secondly, the evaluation needs to be clear on the programme theory that is being tested. This is usually by explicitly setting out a 'context-mechanism outcome' (CMO) hypothesis. The outcomes could be immediate or longer term and could include unintended outcomes.

Data would then be collected, and the data collection tools should allow the data to be disaggregated according to the theory about context – i.e. so that context-mechanisms and outcomes can be linked in the analysis. This is usually easier with quantitative data (e.g., questionnaire survey results which can be disaggregated for different groups). However, qualitative data is also usually important to provide in-depth insights into the contextual factors and mechanisms. Realist evaluations are sometimes based on multiple case studies or narratives. However, realist evaluations do not depend on any specific method – the evaluator can decide what data collection and analysis method best suits the evaluation.

The next stage is to use reasoning to explain what is being observed in the data, in terms of who the programme theory work would work and why; in what context; and by what mechanism; with what result? Inter-linked sets of hypotheses could be identified. For example, the evaluator could develop a matrix to capture different Context-Mechanism-Outcomes situations to show the outcomes achieved in context.

The process of developing recommendations about what did work, and why, for whom in what circumstances would then lead to recommendations emerging about the programme and future interventions.

Analysis:

In a realist evaluation it is likely that quantitative and qualitative data will be analysed together to test the programme theory. This means that qualitative dimensions, such as contextual factors that influence whether mechanisms operate, and description of what the mechanisms are, would be reported alongside data on the different outcomes achieved as a result of the interactions taking place.

Reporting:

The focus of reporting in a realist evaluation would be on the context-mechanisms-outcomes hypotheses (i.e., outcomes for groups, alongside the evidence on how the context and mechanisms explain the outcomes). Based on this, the report should draw out the implications for policy and practice.

Useful links:

Pawson, R and Tilley, N (1997). Realistic Evaluation. London: Sage.

Marchal, B; Van Belle, S and Westhorp, G (2015). Realist Evaluation. https://www.betterevaluation.org/en/approach/realist_evaluation

Graham, A.C., McAleer, S. An overview of realist evaluation for simulation-based education. Adv Simul 3, 13 (2018). https://doi.org/10.1186/s41077-018-0073 https://advancesinsimulation.biomedcentral.com/articles/10.1186/s41077-018-0073-6