Research Article

Reflecting on Reflexivity in Realist Evaluation: A Call to Action

John Downey[1] , **Nicky Harris[2], Sarah Rybczynska-Bunt[3], Elena Golder[4]** , **Katie Shearn[5]** , **and Natasha Bradley[6]**

Abstract
Realist evaluation is increasingly utilised across disciplines due to the value of identifying which mechanisms may explain how, and why, particular outcomes are generated in specific contexts. In theory, realist evaluation provides a tangible way to analyse the inherent complexity in many pressing societal challenges. Realist evaluation encourages a mixed methods approach and choosing a suite of methods that are most relevant within a specific project. Yet, navigating a plurality of methods with abstract philosophical concepts brings operational challenges such that, further methodological guidance is needed. Research processes can be opaque and although the relativist epistemology in realist evaluation is acknowledged, the role of the researcher within the research is often unclear. Reflexivity broadly concerns the overt practice to consider how subjective perspective is intertwined in knowledge production. Reflexivity has been a peripheral consideration in realist evaluation to date and this paper outlines what reflexivity entails, how it enhances integrity to realist principles, and practical ways to exercise it. Through a realist lens the evaluation process is itself subject to the impact of mechanisms and reflexivity provides the apparatus to guard against tunnel vision, undertake robust theory generation and adjudication and increase one's awareness of the influence of personal and organisational entities on research processes and outputs. We conclude with a call to action to the realist community to mobilise reflexivity in a consistent and explicit manner.

Keywords
realist evaluation, reflexivity, research quality, rigour, research methods

Introduction

An appreciation of complexity is becoming more commonplace across disciplines and guidelines for intervention development and evaluation now acknowledge the role of complexity (Skivington et al., 2021). Complexity is a contested concept but essentially posits that interventions are not applied to simply fix a problem, rather interventions disrupt the system of influences resulting in both intended and unintended consequences (Petticrew et al., 2019). The social world is not predictable, controllable or linear, and this recognition requires a greater utilisation of research methods that are epistemologically congruent with the assumptions of complexity science (Greenhalgh & Papoutsi, 2019). Subsequently, the application of realist evaluation has soared due to its emphasis on explaining how and why different outcomes are generated within different contexts (Nielsen et al., 2022).

In this paper, we argue that reflexivity is a crucial yet under-appreciated component of realist evaluation. We begin with an

[1]Centre for Health Technology, Faculty of Health, University of Plymouth, Plymouth, UK
[2]Centre for Health & Clinical Research, University of the West of England, Bristol, UK
[3]Centre for Health and Social Care Innovation, Faculty of Health, University of Plymouth, Plymouth, UK
[4]School of Health and Wellbeing, Plymouth Marjon University, Plymouth, UK
[5]Centre for Health and Social Care Research, Sheffield Hallam University, Sheffield, UK
[6]School of Nursing and Midwifery, Medicine Health and Life Sciences, Queens University Belfast, Belfast, UK

Corresponding Author:
John Downey, University of Plymouth Faculty of Health, Room 202, 6 Kirkby Place, Plymouth PL4 6DT, UK.
Email: john.downey@plymouth.ac.uk

introduction to the underpinning philosophy of realist evaluation and develop this into a discussion of reflexivity as understood through a realist lens. We then present two examples of the practical application of reflexivity to the research process, drawing on two contrasting doctoral projects. Finally, we call on the realist evaluation community to deliberate on the value of reflexivity and incorporate this activity more purposefully and transparently into their work.

A Philosophical Refresher

Realist evaluation is a form of theory driven evaluation that investigates generative causation for its explanatory value (Marchal et al., 2012). Generative causation is coherent with the complexity inherent in the real world because it refutes linear, deterministic, and successionist notions of causality (Figure 1) (see Befani, 2012). Traditionally in realist evaluation mechanisms are seen as generative forces arising from an individual's responses to resources, under certain circumstances, bringing about observable events (Dalkin et al., 2015). However, more recently authors have argued this conceptualisation may stifle our understanding of causal influences because mechanisms operate at different levels of the complex whole (Lemire et al., 2020). The realist cycle identifies links between context, mechanism, and outcome as configurations that produce partial explanations of the phenomena under study.

Unearthing and testing context-mechanism-outcome configurations to explain a particular intervention or programme requires the researcher to unpick and theorise its composite parts identifying the existence of demi-regularities (which are somewhat stable patterns of how things play out) (Pawson, 2006), and proposing conjectures to explain "apparent uniformities" (Pawson, 2013, p. 9). In realist evaluation, it is encouraged to adopt both qualitative and quantitative methods, and draw from a pragmatic common sense to examine how mechanisms manifest and the conditions that catalyse outcomes, to iteratively develop programme theories (Danermark et al., 2005; Gilmore et al., 2019; Lusted, 2018).

The realist evaluation paradigm thus provides a philosophical framework to advance the accumulation of knowledge in complex settings and explain social events (Fletcher, 2017). Methods pluralism is encouraged but the philosophical underpinning is one of realist evaluation's most distinctive elements (Marchal et al., 2012). Realism should be drawn upon for the fruits of its labour whilst privileging empirical data to stress test ideas about causal configurations. Realism should therefore be applied in an instrumental manner, meaning work should be judged on how the integrity to the philosophical framework is demonstrated and the production of worthwhile results, instead of debating the possibility of capturing reality (Hammersley, 2018; Pawson, 2013).

A central principle of this type of realism is that what is real is not limited to what can be known – that is, the separation of ontology and epistemology (Fletcher, 2017). Ontologically,
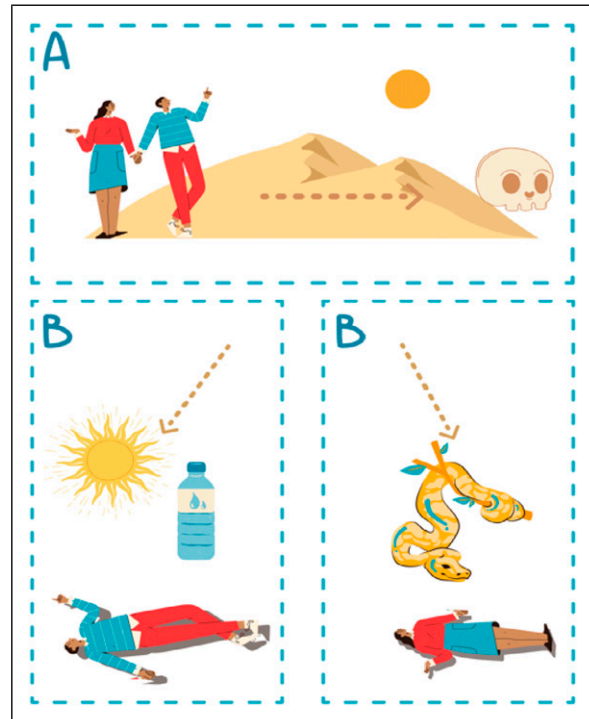

Figure 1. An illustration of a (a) successionist and (b) generative perspective of causality (adapted from Befani, 2012). (a) Represents a successionist view of causality. It is observed that walking in a desert is followed by death. This is routinely demonstrated, showing that walking in the desert causes death, but there is no exploration on what part of the walk leads to death. (b) Represents a generative view of causality. Two separate conditions trigger different causes of death. The two people take separate routes, one is exposed to extreme heat leading to death by thirst and the other takes refuge under a tree where a snake resides, and they die by poisoning.

realism assumes that reality exists independent of human knowledge, observation, or empirical measurement. Depth ontology proposes that reality can be separated into progressively basic layers which provides an ontological map guiding scientific enquiry (Figure 2) (Downey, 2022). Understanding that reality is stratified is helpful in theorising generative causation and furthers the investigation of relationships between what we experience, what actually happens, and the causal forces that produce outcomes (Bhaskar, 2016).

Epistemologically, realism adopts an interpretive lens understanding that human knowledge is socially constructed and theory laden (Easton, 2010) and that accounts of reality are mediated through filters (e.g., meanings, language, social context) (Oliver, 2012). Thus for realists, the world possesses real entities that exist and have causal impact despite our observations and interpretations of them, *and* that the pursuit of these objective entities will always result in fallible or partial knowledge.

For realist research to be intelligible a distinction must be made between the transitive and intransitive dimensions of
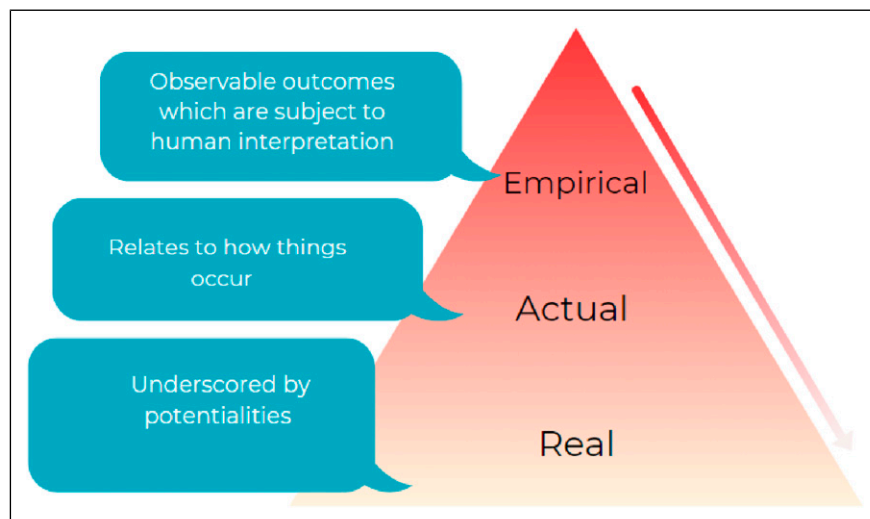
Figure 2. A visual overview of depth ontology or stratified reality, which guides realist enquiry (Downey, 2022).

knowledge. Transitive ideas are contested and changeable because of human interpretation, whereas the intransitive objects of interest remain relatively stable (Lusted, 2018). The transitive dimension of knowledge includes the discourses and hunches adopted to explain the intransitive dimension. Partially stable entities of reality exist regardless of human comprehension which compels the researcher to appreciate and theorise how underlying structures determine the shape of phenomena that can be observed (Wiltshire, 2018). Crucially, the researcher is a part of this reality, and thus their conclusions are also influenced, or even dictated, by underlying real entities.

The application of realist evaluation is at risk of appearing reductionist in journal manuscripts where much of the research process is implicit or pruned for publication. The methods pluralism welcomed in realist evaluation can threaten integrity if methods are selected without consideration of how they align to the philosophical underpinning, challenging if work is "really realist". Having revisited some of the foundational philosophical concepts, we now turn to reflexivity and its value when undertaking realist research.

Defining Reflexivity for Realists

In its simplest form, reflexivity involves understanding how the processes of research determine its outcomes and, crucially, recognising the researchers' roles in these processes (Maxwell, 2012). If research is knowledge production, then reflexivity asks how the involvement of each of the knowledge producers, and the various contexts of knowledge production, will shape the results, which is congruent with the assumptions of realism. Reflexivity is generally recognised as fundamental to qualitative research, but not frequently transparent within publications (Tarrant, 2017). Since research relies on interpretation, reflexivity can be seen as the interpretation of

interpretation - an attempt to explicate and question one's habitual views.

A key characteristic differentiating the social sciences from the physical sciences is that individual people, and groups of people, are capable of self-inquiry and adaptation. Social agents have the capacity to recognise their own socialisation (i.e., the structures within which they exist), and this might lead to change (or resistance to change) in their behaviour (Archer, 2010). Reflexivity is therefore the self-monitoring and responding to thoughts, behaviours, and emotions as the phenomenon under investigation is navigated. Reflexivity is related to, but different from, reflection. Reflection enables observation of practice, as if in a mirror-image, whereas reflexivity is self-referential – it involves "bending back" some thought upon the self (Tauber, 2005, p. 50). By exposing and questioning one's ways of thinking, reflexivity can end up changing the process of reflection.

Researchers of different ontological orientations will approach reflexivity differently, partially related to their different understandings of what it means to be a self (Jamie & Rathbone, 2022; Jamieson et al., 2023). The self can be understood as emerging from the inner dialogue between the "I" and the "Me". The "I" is the active, emotional, impulsive part of the psyche, whereas the "Me" is the learnt repertoire of socialised behaviours and values. Some scholars conceptualise this as a three-way conversation, introducing the "Thou" or the "You" – that is, a future focused concept that has not yet come into being (Archer, 2010).

As a crude example of this inner world: If we imagine that you hear an ice cream van outside. The "I" would love to have a sugary treat. The "Me" is disapproving of this treat before dinnertime. The "Thou" wants to enjoy life while protecting their wallet and their waistline. The interaction between these three perspectives will underpin the response to the sensory stimuli of the ice cream van's melody. Importantly, our

historic, present, and anticipated contexts will inform all three components of the "I-Me-Thou" conversation.

Margaret Archer's work underpinned by critical realism and is at its core about how reflexivity can avoid the conflation of structure and agency accepting their interviewed and reciprocal relationship (Caetano, 2015). It is argued that reflexivity is a diverse process and that the set of practices which will vary depending on the approach and research requirements. Archer (2010) advanced this point and developed four reflexive typologies for how people navigate the social world (1) autonomous: a person who forms conclusions through internal conversations and tends to not be easily swayed by external influences (2) meta-reflexive: describes an individual that rejects traditional values and is critical about their own self-reflection (3) communicative reflexive: is characterised by an inclination to rely on interactions with others to shape or finalise their conclusions (4) fractured reflexive: is demarked by social distress and inability to act purposively.

An appreciation of this standpoint on structure, agency, and the power of internal dialogue to have causal change on external influences has further importance for understanding mechanisms within realist evaluation (Shaw et al., 2018). Engaging in reflexivity can help to reveal the positionality and identity of researchers in relation to the world around them. Reflexive researchers are more able to involve other perspectives in their work - not just as a consultation exercise with a predefined conclusion, but as a genuine collaboration with the potential to transform research practice, values, and institutions. Collective reflexivity has the potential for broadening the conversation, helping individual members work together productively, and reducing the risk of dismissing inconvenient findings in the search for parsimonious theory. Greater awareness of our own tendencies enables further interdisciplinary reflexivity to take place.

Maintaining Integrity in Realist Evaluation

It is recommended to start a realist evaluation with a thorough consideration of the complexity inherent within the programme or object of study which can guide empirical work - see the VICTORE checklist (Pawson, 2013, p. 33). This is a valuable exercise but is insufficient in its own right as one must not forget that the researcher is not a neutral observer. Reflexivity provides a sense check to avoid coveting areas of interest to the team or funders. Reflexivity in realist evaluation is not a new concept and others have highlighted it's role to ameliorate the political and democratic elements of evaluation (Carpenter, 2005); yet the explicit practice of reflexivity remains underutilised in realist work.

Because the evaluator cannot examine every area of interest, it is also encouraged in realist work to map the contours of the phenomena of interest to "concentrate fire" on the "juicy bits" (Astbury, 2018, pp. 73–74). Thus, some areas are given primacy during the research process (Pawson et al., 2005). For realist work to stand up to scrutiny, it must be recognised how

the assumptions, powers, and liabilities of the research team and their key stakeholders form the basis of such decisions. The baggage one carries to the research must be subject to appraisal, which occurs via reflexivity (Emmel, 2013; Maxwell, 2012).

Reflexivity also provides a platform for working through and capturing how the decisions involved in proposing ideas and refining theory occur. This is not to say that it is straightforward or that this self-awareness is possible in its totality (Emmel, 2013; McLachlan & Garcia, 2015), however, it is necessary as a self-audit trail supporting transparency, articulating hunches, mapping the system, and directing pathways for future empirical testing (Layder, 2005).

Furthermore, when attempting to arrange data into causal configurations, creative interpretation of data is needed to allow retroduction to occur, which essentially involves thought experiments to propose "what must be true for X to exist?" (Danermark et al., 2005; Jagosh, 2020; Oliver, 2012). This process relies upon the evaluator's logic which is a culmination of their biography, academic background, coverage of various potential theoretical frameworks and an infinite number of contextual factors including aspects of existing societal power structures. Thus, the analytic process is inherently reflexive, but the introspective work to spell out the evaluator's values and their ideologies about success and failure are rarely articulated.

Realist evaluation often leads to a wide range of potential theories which need curating and adjudication to ensure the work is manageable and of practical value. While realism acknowledges that theories about the social world are inherently propositional, this does not exempt us from acting responsibly with data and demonstrating how we conduct eliminating work to distinguish between competing theories (Bhaskar, 2016). When engaging in realist-informed reflexivity, the articulation of biography is only one key element and there is a need to examine the purposeful theories on how the team came to examine the topic and the consequences this has on subsequent work (Roberts & Sanders, 2005).

As realism staddles paradigms traditional views about rigour require adaptation (Ronkainen & Wiltshire, 2019) (Table 1). For realists, there is an acceptance that something exists which impacts outcomes but that our interpretations of these explanations will be fallible. This creates a challenge when judging validity about conclusions, as enhancing validity to secure objective conclusions from an evaluation is not possible (Maxwell, 2012). Nevertheless, realism posits that there are multiple potential mechanisms explaining outcomes, so validity is not about the accuracy of the account in terms of how it represents a single reality. There is therefore no ideal suite of strategies to judge a realist evaluation, as validity pertains to the conclusions drawn from data, not the methods used. Validity is then about how accounts represent the features they are intending to explain and the plausibility of ideas to explain the actual state of affairs (Ronkainen & Wiltshire, 2019). As the work relies on the investigator's appraisal of

Table 1. The Specific Realist Perspectives on Various Aspects of Validity (Ronkainen & Wiltshire, 2019; Zachariadis et al., 2013).

Validity category	Qualitative labels	Quantitative labels	Realist perspective
Design validity	Descriptive validity Credibility Transferability	Internal validity External validity	Empirical events are a manifestation of unobservable mechanisms. The need to separate and scrutinise each aspect of theory is essential. The work must consider the accuracy of the data generated and execution of the research
Measurement validity	Theoretical validity Dependability Consistency Plausibility	Reliability Construct validity	The work should measure intended mechanisms and their link to outcomes. Legitimacy of claims can be enhanced by employing realist consistent theoretical frameworks and testing theory in partially closed scenarios
Inferential validity	Interpretive validity Confirmability	Statistical conclusion validity	Statistics can provide useful patterns of data, but they are merely descriptions of relationships. There is a need to consider how to ensure interpretations reflect the perceptions and experiences of respondents

relative threats to validity, the process of seeking out rival explanations requires an increased awareness of self, audit trail of decisions, and a process to examine how interpretations of data were made - which makes reflexivity a fundamental practice to enhance the quality of a realist evaluation.

Demonstrating Reflexivity in Realist Evaluation

To illustrate how reflexivity can be used within a realist evaluation we now present two examples showing the various ways reflexivity came to fruition and the rationale as it pertains to realist evaluation.

Example 1: Realist Evaluation of eHealth in Children's Palliative Care

This doctoral study was a mixed method design and employed data collection procedures common in realist evaluation (Harris, 2022; Renmans & Castellano Pleguezuelo, 2023). As mentioned above, the investigator's personal frameworks required acknowledgement; in this case, the research interest derived from a previous clinical role and personal involvement in developing the eHealth intervention. There is a body of work discussing the dual-role issues in research (Coar & Sim, 2006; Hay-Smith et al., 2016; Richards & Schwartz, 2002) which recognises that one way to mitigate the risks is to ensure that reflexivity is embedded throughout (Råheim et al., 2016). Accordingly, the research team adopted a structured approach to reflexivity using a matrix (Rae & Green, 2016) which provided prompts to stimulate discussion, but also required modification to be congruent with a realist approach (Table 2) (Harris, 2022).

At the pre-research stage, the realist RAMESES reporting standards state that a key decision should involve the justification for using realist methodology (Wong et al., 2017). The wording of the research question and the focus on generative causation arise from assumptions brought by the research team, and these need to be articulated in an explicit and

transparent manner. The focus on generative causation can go awry if there is a misalignment of values and anticipated outputs within the research team. Likewise, stakeholder involvement is essential, but it is not feasible to include all possible stakeholders thus selecting which voices to involve will be influential and is usually driven by the priorities of the research team. Structured reflexivity helped to aid this decision-making by raising awareness of motivations and consequences inherent in such decisions.

Data collection and theory development is iterative and initial programme theories evolve over time, adding to the piecemeal development of knowledge (Pawson, 2013). During data collection, the doctoral candidate engaged frequently with reflexivity as a process of self-audit to consider if and how dominant organisational, personal, cultural, or societal agendas were playing out in a semi unconscious manner.

The identification and selection of theoretical frameworks to support abstraction of granular data is a cornerstone of realist inquiry (Westhorp, 2012). Yet, the choice of theories is often based on norms, popularity, and familiarity within a discipline, rather than utility (Davis et al., 2015). Theoretical explanations are fallible, and some dimensions are excluded that may be ripe for further exploration. Structured and frequent reflexivity was helpful in recognising unexamined decisions, surfacing such limitations, and enabling more informed choices between substantive theories.

The innovative application of reflexivity within this eHealth project was documented in the PhD thesis. Although the brevity required by publications leads to the omission of such nuances, these processes added to depth and rigour of the findings.

Example 2: Realist Evaluation of Engagement in Wellbeing Groups for People with Long-Term Conditions

The second example relates the development of programme theory in a PhD exploring engagement in group-based self-

Table 2. A Structured Reflexive Matrix Informed by Rae and Green (2016) to Guide the Phases of a Realist Evaluation.

Position	Overall social space	Within field of specialists	Within academia
Pre-research	How do researchers' broader motivations affect the reason to conduct research in the first place, the choice of topic and research question, and the choice of methodology?	What is the relationship between the researcher and the field? How is the topic relevant to field?	Where do the researchers' interests/conflicts exist in relation to the act of carrying out the research, literature review, and interpretation of the literature/data?
Data collection	What are the shared and divergent understandings between the researcher and those impacted by the research topic? To what extent are these understandings related to social differences (e.g., gender, education, experience)? Is the researcher prepared to undergo change as a result of interaction with the research?	Do the researcher and participants share the same language? Are there any power differentials between the researcher and the participant, based on positions held (present and past), discipline, or education?	Are data collection strategies inadvertently shaped by popular (perhaps fleeting) scholarly opinion?
Data analysis	How are societal assumptions relevant during the interpretation of findings? How does the social world shape the adjudication and elimination of competing theories?	How does the researcher's experience with the field shape the analysis? Are some data dismissed as being commonplace when they may warrant deeper interrogation? Where does analytical authority reside?	How do the researchers moderate the role of academia on the analytical process? Could any drive for timely outputs or specific conclusions produce omissions or inadvertent fabrications?

management programmes for people with long-term health conditions (Golder et al., in preparation). There is an enduring challenge noted in the literature about how to document the realist analysis process (Adams et al., 2016; Bergeron & Gaboury, 2020). The analysis procedures are often not described and examples of how to use software management tools to support this task are lacking (Bergeron & Gaboury, 2020). If there was a clear process to house selective decisions and articulate the iterative process inherent in realist analysis within software, it could help with credibility and the quality of inferences drawn.

In the engagement project reflexivity was documented throughout the theory development process using Microsoft Excel and NVivo. Reflexivity was implemented from the outset to draw boundaries and establish the focus of the evaluation, an important priority early in the realist cycle (Westhorp, 2012). Involving stakeholders in the process of determining the focus can make the findings of the evaluation more useful, which is a key criterion in the RAMESES quality standards for realist research (Wong et al., 2017), and essential for research impact. The decision to investigate engagement was made by triangulating feedback from multiple stakeholder views gathered at the start of the project and considering pragmatic implications of following different lines of enquiry. The conceptualisation of some of the priority concerns of stakeholder was informed by the researcher's prior experience working in education and healthcare. Reflexivity at this early stage was key to transparency, justification of the research focus, and the clarity of the project.

Realist research usually involves analysing large volumes of data from multiple data sources when building theory.

Pawson (2013) uses the analogy of a swamp to describe this conundrum, which has also been used in reflexivity (Finlay, 2002), and offers strategies to progress. These include the use of a conceptual platform, continuous hypothesis selection and shedding, focusing theory selection on points of policy discord, and developing lines of inquiry. These principles were drawn on for focusing both programme theory building and exercising reflexivity in this doctoral study.

Prior to data collection, two complexity-consistent frameworks (Bright et al., 2015; Nobles et al., 2018) were selected to help draw conceptual system boundaries and identify the constructs of engagement that may be relevant to self-management programmes. The resulting conceptual platform provided a framework for the semi-structured interviews with programme architects and practitioners, which were used for initial programme theory building.

Initial programme theories (IPTs) were generated from interviews in the form of "if…then…because" statements, a heuristic used in realist research to facilitate retroductive theorising. Retroduction is (Mukumbang et al., 2021) necessary to move from understanding events from a successional view of causation, to unearthing generative causation (Sayer, 2010). Each "if…then…because" statement was recorded in an Excel table and assigned an IPT number. Interview excerpts were coded to an NVivo node labelled with its corresponding IPT number. Theoretical memos, which function as an ongoing account of the analytical thought process of the researcher (Layder, 2005), were recorded for IPTs using the linked memo function in NVivo. Causal explanations are often not apparent in distinct data excerpts and may be identified in the process of reading larger data excerpts. Extended memos for entire

interview transcripts were used to provide an overview of the data and the researcher's analytical thinking, which in turn provided an opportunity to identify gaps in the data and the need for further theory development (Layder, 2005). Documenting the theorising process using individual IPT and extended linked memos enabled the review of lines of inquiry and the subsequent inclusion and discarding of hypotheses.

Recording IPTs in an Excel table allowed the broad coding and grouping of IPTs into conceptual buckets using the table filter and sort function. This organisation categorisation, where the subject matter is described without exploring what is said, can serve the purpose of sorting data for further analysis (Maxwell, 2012). Viewing "if…then…because" in groups helped to develop lines of inquiry and consolidate partial and similar IPTs into more detailed programme theories for testing later in the evaluation. Further, the process of grouping helped with hypothesis selection and shedding as some broad topics had very little data whereas others had multiple IPTs. Outlier IPTs were archived so they could be revisited, if necessary, along with the reasons for abandoning certain strands of theory. Grouped IPTs were copied and pasted into a new and separate programme theory Excel sheet, which created a trail of the categorisation steps. This process was mirrored in NVivo with IPTs being moved into new parent programme theory nodes. These marked steps provided an opportunity to write reflexive accounts of the grouping and the hypothesis selection and shedding process. Reflexive accounts were recorded using linked memos for the parent programme theory nodes.

Using Excel and Nivo to organise data and ideas during this multistage and iterative theory development phase of this doctoral study offered multiple points to conduct and document reflexivity. Employing a conceptual framework and the process of hypothesis selection and shedding offered a means to navigate both the programme theory and reflexive swamps.

Call to Action

The development of realist evaluation emerged from a critical reflection on existing methodologies, particularly public health evaluators becoming weary with the lack of progress seen when applying context-blind methods (Pawson & Tilley, 1997). Nonetheless, despite the increasing acknowledgment of the complexity inherent in many contemporary challenges; since then, the application of research approaches that explicitly account for complexity has been underwhelming. The popularity of realist evaluation has soared, however, where pragmatism dominates over integrity to the philosophical assumptions, there is the risk of replicating historical issues when researching complex phenomenon. Indeed realist evaluation has been critiqued for producing bureaucratic outcomes that fail to capture the concerns of those affected by the area under investigation (Porter & O'Halloran, 2012). Overcoming these criticisms requires practices to enhance the integrity during the application of an interpretive epistemology. We argue that reflexivity can support auditability, credibility, and

the potential to unearth, interpret, and adjudicate prudent theory (Daniel, 2019) increasing the confidence in realist evaluation by sense checking which voices or agendas are taking precedent. The next generation of realist evaluators should endeavour to delineate their blind-spots, recognise with humility how their own social context shapes their analytical processes, and subsequently work to bring missing perspectives into research teams. Reflexivity is one key process in the repertoire of skills necessary to achieve this, by which we can continue to identify opportunities for progress.

ORCID iDs

John Downey ⓘ https://orcid.org/0000-0001-8534-2437
Elena Golder ⓘ https://orcid.org/0000-0002-2795-1725
Katie Shearn ⓘ https://orcid.org/0000-0001-7209-8404
Natasha Bradley ⓘ https://orcid.org/0000-0003-0768-9819

References

Adams, A., Sedalia, S., McNab, S., & Sarker, M. (2016). Lessons learned in using realist evaluation to assess Maternal and Newborn health programming in rural Bangladesh. *Health Policy and Planning*, *31*(2), 267–275. https://doi.org/10.1093/heapol/czv053

Archer, M. S. (2010). *Conversations about reflexivity*. Routledge.

Astbury, B. (2018). Making claims in realist research. In N. Emmel, J. Greenhalgh, A. Manzano, M. Monaghan, & S. Dalkin (Eds.), *Doing realist research*. Sage Publications Ltd.

Befani, B. (2012). Modes of causality and causal inference. In E. Stern, N. Stame, J. Mayne, K. Forss, R. Davies, & B. Befani (Eds.), *Broadening the range of designs and methods for impact evaluation*. Department for International Development.

Bergeron, D. A., & Gaboury, I. (2020). Challenges related to the analytical process in realist evaluation and latest developments on the use of NVivo from a realist perspective. *International Journal of Social Research Methodology*, *23*(3), 355–365. https://doi.org/10.1080/13645579.2019.1697167

Bhaskar, R. (2016). *Enlightened common sense: The philosophy of Critical Realism*. Routledge.

Bright, F. A. S., Kayes, N. M., Worrall, L., & McPherson, K. M. (2015). A conceptual review of engagement in healthcare and rehabilitation. *Disability & Rehabilitation*, *37*(8), 643–654. https://doi.org/10.3109/09638288.2014.933899

Caetano, A. (2015). Defining personal reflexivity: A critical reading of Archer's approach. *European Journal of Social Theory*, *18*(1), 60–75. https://doi.org/10.1177/1368431014549684

Carpenter, M. (2005). Bridging the gap between collaborative and realist evaluation: A general critique and case study of European union employment policies. *International Journal of Action Research*, *1*(3), 311–338. https://0-search.ebscohost.com.brum.beds.ac.uk/login.aspx?direct=true&db=sih&AN=19434814&site=ehost-live&scope=site

Coar, L., & Sim, J. (2006). Interviewing one's peers: Methodological issues in a study of health professionals. *Scandinavian Journal of Primary Health Care*, *24*(4), 251–256. https://doi.org/10.1080/02813430601008479

Dalkin, S. M., Greenhalgh, J., Jones, D., Cunningham, B., & Lhussier, M. (2015). What's in a mechanism? Development of a key concept in realist evaluation. *Implementation Science*, *10*(1), 1–7. https://doi.org/10.1186/s13012-015-0237-x

Danermark, B., Ekström, M., & Karlsson, J. C. (2005). *Explaining society: Critical realism in the social sciences*. Routledge. https://doi.org/10.4324/9781351017831

Daniel, B. K. (2019). Using the TACT framework to learn the principles of rigour in qualitative research. *Electronic Journal of Business Research Methods*, *17*(3), 118–129. https://doi.org/10.34190/JBRM.17.3.002

Davis, R., Campbell, R., Hildon, Z., Hobbs, L., & Michie, S. (2015). Theories of behaviour and behaviour change across the social and behavioural sciences: A scoping review. *Health Psychology Review*, *9*(3), 323–344. https://doi.org/10.1080/17437199.2014.941722

Downey, J. (2022). *Behaviour change practices in exercise referral practitioners: A realist evaluation of implementation*. Doctoral Disseration, St Mary's University. Retrevied from. https://research.stmarys.ac.uk/id/eprint/5797/

Easton, G (2010) (In this issue). Critical realism in case study research. *Industrial Marketing Management*, *39*(1), 118–128. https://doi.org/10.1016/j.indmarman.2008.06.004

Emmel, N. (2013). *Sampling and choosing cases in qualitative research: A realist approach*. Sage Publications Ltd. https://doi.org/10.4135/9781473913882

Finlay, L. (2002). Negotiating the swamp: The opportunity and challenge of reflexivity in research practice. *Qualitative Research*, *2*(2), 209–230. https://doi.org/10.1177/146879410200200205

Fletcher, A. J. (2017). Applying critical realism in qualitative research: Methodology meets method. *International Journal of Social Research Methodology*, *20*(2), 181–194. https://doi.org/10.1080/13645579.2016.1144401

Gilmore, B., McAuliffe, E., Power, J., & Vallières, F. (2019). Data analysis and synthesis within a realist evaluation: Toward more transparent methodological approaches. *International Journal of Qualitative Methods*, *18*, 1–11. https://doi.org/10.1177/1609406919859754

Golder, E., Dawson, P., Downey, J., & Bloxham, S. (2025). Increasing transparency in programme theory development: An illustrative example of methods. Manuscript in preparation.

Greenhalgh, T., & Papoutsi, C. (2019). Spreading and scaling up innovation and improvement. *BMJ (Online)*, *365*(May), 1–8. https://doi.org/10.1136/bmj.l2068

Hammersley, M. (2018). What is ethnography? Can it survive? Should it? *Ethnography and Education*, *13*(1), 1–17. https://doi.org/10.1080/17457823.2017.1298458

Harris, N. (2022). *Ehealth and the delivery of person-centred care for children with life limiting conditions: A realist evaluation*. Doctoral Thesis, University of the West of England. Retrived from. https://uwe-repository.worktribe.com/output/9458142

Hay-Smith, E. J. C., Brown, M., Anderson, L., & Treharne, G. J. (2016). Once a clinician, always a clinician: A systematic review to develop a typology of clinician-researcher dual-role experiences in health research with patient-participants. *BMC Medical Research Methodology*, *16*(1), 1–17. https://doi.org/10.1186/s12874-016-0203-6

Jagosh, J. (2020). Retroductive theorizing in Pawson and Tilley's applied scientific realism. *Journal of Critical Realism*, *19*(2), 121–130. https://doi.org/10.1080/14767430.2020.1723301

Jamie, K., & Rathbone, A. P. (2022). Using theory and reflexivity to preserve methodological rigour of data collection in qualitative research. *Research Methods in Medicine & Health Sciences*, *3*(1), 11–21. https://doi.org/10.1177/26320843211061302

Jamieson, M. K., Govaart, G. H., & Pownall, M. (2023). Reflexivity in quantitative research: A rationale and beginner's guide. *Social and Personality Psychology Compass*, *17*(4), 1–15. https://doi.org/10.1111/spc3.12735

Layder, D. (2005). *Sociological Practice. Linking theory and social research*. Sage Publications Ltd.

Lemire, S., Kwako, A., Nielsen, S. B., Christie, C. A., Donaldson, S. I., & Leeuw, F. L. (2020). What is this thing called a mechanism? Findings from a review of realist evaluations. *New Directions for Evaluation*, *2020*(167), 73–86. https://doi.org/10.1002/ev.20428

Lusted, J. (2018). A critical realist morphogenetic approach to researching sport policy: Reflections on a large-scale study of policy implementation in grassroots English football. *International Journal of Sport Policy and Politics*, *10*(4), 705–719. https://doi.org/10.1080/19406940.2018.1522658

Marchal, B., van Belle, S., van Olmen, J., Hoerée, T., Kegels, G., van Belle, S., van Olmen, J., Hoerée, T., & Kegels, G. (2012). Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research. *Evaluation*, *18*(2), 192–212. https://doi.org/10.1177/1356389012442444

Maxwell, J. A. (2012). *A realist approach for qualitative research*. Sage Publications Ltd.

McLachlan, C. J., & Garcia, R. J. (2015) (In this issue). Philosophy in practice? Doctoral struggles with ontology and subjectivity in qualitative interviewing. *Management Learning*, *46*(2), 195–210. https://doi.org/10.1177/1350507615574634

Mukumbang, F. C., Kabongo, E. M., & Eastwood, J. G. (2021). Examining the application of retroductive theorizing in realist-informed studies. *International Journal of Qualitative Methods*, *20*(1), 1–14. https://doi.org/10.1177/16094069211053516

Nielsen, S. B., Lemire, S., & Tangsig, S. (2022). Unpacking context in realist evaluations: Findings from a comprehensive review.

Evaluation, *28*(1), 91–112. https://doi.org/10.1177/13563890211053032

Nobles, J. D., Perez, A., Skelton, J. A., Spence, N. D., & Ball, G. D. (2018). The engagement pathway: A conceptual framework of engagement-related terms in weight management. *Obesity Research & Clinical Practice*, *12*(2), 133–138. https://doi.org/10.1016/j.orcp.2017.12.005

Oliver, C. (2012). Critical realist grounded theory: A new approach for social work research. *British Journal of Social Work*, *42*(2), 371–387. https://doi.org/10.1093/bjsw/bcr064

Pawson, R. (2006). *Evidence-based policy: A realist perspective*. Sage Publications Ltd.

Pawson, R. (2013). *The science of evaluation: A realist manifesto*. Sage Publications Ltd.

Pawson, R., Greenhalgh, T., Harvey, G., & Walshe, K. (2005). Realist review – a new method of systematic review designed for complex policy interventions. *Journal of Health Services Research & Policy*, *10*(Suppl 1), 21–34. https://doi.org/10.1258/1355819054308530

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

Petticrew, M., Knai, C., Thomas, J., Rehfuess, E. A., Noyes, J., Gerhardus, A., Grimshaw, J. M., Rutter, H., & McGill, E. (2019). Implications of a complexity perspective for systematic reviews and guideline development in health decision making. *BMJ Global Health*, *4*(Suppl 1), Article e000899. https://doi.org/10.1136/bmjgh-2018-000899

Porter, S., & O'Halloran, P. (2012). The use and limitation of realistic evaluation as a tool for evidence-based practice: A critical realist perspective. *Nursing Inquiry*, *19*(1), 18–28. https://doi.org/10.1111/j.1440-1800.2011.00551.x

Rae, J., & Green, B. (2016). Portraying reflexivity in health services research. *Qualitative Health Research*, *26*(11), 1543–1549. https://doi.org/10.1177/1049732316634046

Råheim, M., Magnussen, L. H., Sekse, R. J. T., Lunde, Å., Jacobsen, T., & Blystad, A. (2016). Researcher-researched relationship in qualitative research: Shifts in positions and researcher vulnerability. *International Journal of Qualitative Studies on Health and Well-Being*, *11*(1), 1–12. https://doi.org/10.3402/qhw.v11.30996

Renmans, D., & Castellano Pleguezuelo, V. (2023). Methods in realist evaluation: A mapping review. *Evaluation and Program Planning*, *97*(July 2022), Article 102209. https://doi.org/10.1016/j.evalprogplan.2022.102209

Richards, H. M., & Schwartz, L. J. (2002). Ethics of qualitative research: Are there special issues for health services research? *Family Practice*, *19*(2), 135–139. https://doi.org/10.1093/fampra/19.2.135

Roberts, J. M., & Sanders, T (2005) (In this issue). Before, during and after: Realism, reflexivity and ethnography. *Sociological Review*, *53*(2), 294–313. https://doi.org/10.1111/j.1467-954X.2005.00515.x

Ronkainen, N. J., & Wiltshire, G. (2019). Rethinking validity in qualitative sport and exercise psychology research: A realist perspective. *International Journal of Sport and Exercise Psychology*, *19*(1), 13–28. https://doi.org/10.1080/1612197X.2019.1637363

Sayer, A. (2010). *Method in social science: A realist approach* (2nd ed.). Routledge.

Shaw, J., Gray, C. S., Baker, G. R., Denis, J. L., Breton, M., Gutberg, J., Embuldeniya, G., Carswell, P., Dunham, A., McKillop, A., Kenealy, T., Sheridan, N., & Wodchis, W. (2018). Mechanisms, contexts and points of contention: Operationalizing realist-informed research for complex health interventions. *BMC Medical Research Methodology*, *18*(1), 1–12. https://doi.org/10.1186/s12874-018-0641-4

Skivington, K., Matthews, L., Simpson, S. A., Craig, P., Baird, J., Blazeby, J. M., Boyd, K. A., Craig, N., French, D. P., McIntosh, E., Petticrew, M., Rycroft-Malone, J., White, M., & Moore, L. (2021). A new framework for developing and evaluating complex interventions: Update of Medical Research Council guidance. *The BMJ*, *374*(2018), 1–11. https://doi.org/10.1136/bmj.n2061

Tarrant, A. (2017). Getting out of the swamp? Methodological reflections on using qualitative secondary analysis to develop research design. *International Journal of Social Research Methodology*, *20*(6), 599–611. https://doi.org/10.1080/13645579.2016.1257678

Tauber, A. I. (2005). The reflexive project: Reconstructing the moral agent. *History of the Human Science*, *18*(4), 49–75. https://doi.org/10.1177/0952695105058471

Westhorp, G. (2012). Using complexity-consistent theory for evaluating complex systems. *Evaluation*, *18*(4), 405–420. https://doi.org/10.1177/1356389012460963

Wiltshire, G (2018) (In this issue). A case for critical realism in the pursuit of interdisciplinarity and impact. *Qualitative Research in Sport, Exercise and Health*, *10*(5), 525–542. https://doi.org/10.1080/2159676X.2018.1467482

Wong, G., Westhorp, G., Greenhalgh, J., Manzano, A., Jagosh, J., & Greenhalgh, T. (2017). Quality and reporting standards, resources, training materials and information for realist evaluation: The RAMESES II project. *Health Services and Delivery Research*, *5*(28), 1–108. https://doi.org/10.3310/hsdr05280

Zachariadis, M., Scott, S., & Barrett, M. (2013). Methodological implications of critical realism for mixed-methods research. *MIS Quarterly: Management Information Systems*, *37*(3), 855–879. https://doi.org/10.25300/misq/2013/37.3.09