



Accounting for complexity in critical realist trials: the promise of PLS-SEM

Heidi Singleton, Sam Porter, John Beavis, Liz Falconer, Jacqueline Priego Hernandez & Debbie Holley

To cite this article: Heidi Singleton, Sam Porter, John Beavis, Liz Falconer, Jacqueline Priego Hernandez & Debbie Holley (2023) Accounting for complexity in critical realist trials: the promise of PLS-SEM, *Journal of Critical Realism*, 22:3, 384-403, DOI: [10.1080/14767430.2023.2217652](https://doi.org/10.1080/14767430.2023.2217652)

To link to this article: <https://doi.org/10.1080/14767430.2023.2217652>



© 2023 Bournemouth University. Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 22 Jun 2023.



Submit your article to this journal [↗](#)



Article views: 136





View related articles [↗](#)



View Crossmark data [↗](#)

Accounting for complexity in critical realist trials: the promise of PLS-SEM

Heidi Singleton ^a, Sam Porter ^a, John Beavis^a, Liz Falconer^a,
Jacqueline Priego Hernandez^b and Debbie Holley^a

^aFaculty of Health and Social Sciences, Bournemouth University, Dorset, UK; ^bFaculty of Humanities and Social Sciences, University of Portsmouth, Hampshire, UK

ABSTRACT

Background: Randomized controlled trials have been criticized for their inability to identify and differentiate the causal mechanisms that generate the outcomes they measure. One solution is the development of realist trials that combine the empirical precision of trials' outcome data with realism's theoretical capacity to identify the powers that generate outcomes. *Main Body:* We review arguments for and against this position and conclude that critical realist trials are viable. Using the example of an evaluation of the educational effectiveness of virtual reality simulation, we explore whether Partial Least Squares Structural Equation Modelling can move statistical analysis beyond correlational analysis to support realist identification of the mechanisms that generate correlations. *Conclusion:* We tentatively conclude that PLS-SEM, with its ability to identify 'points of action', has the potential to provide direction for researchers and practitioners in terms of how, for whom, when, where and in what circumstances an intervention has worked.

KEYWORDS

RCT; critical realist evaluation; critical realism; causality; PLS-SEM

Introduction

The aim of this paper is to explore if and how the application of Partial Least Squares Structural Equation Modelling (PLS-SEM) to randomized controlled trial analysis has the potential to overcome the challenges of applying statistical analysis to supplement and support the complex 'configurational' (Bhaskar 1989, 43) approach to causation espoused by critical realism (CR). This view asserts that 'most social phenomena ... are conjuncturally determined and, as such, in general have to be explained in terms of a multiplicity of causes' (Bhaskar 1989, 43). The claimed strength of regression analysis tools such as PLS-SEM is their ability to increase the capacity of trials to differentiate between variables and identify their relations to a greater specificity. However, the compatibility of such an approach with the critical realist focus on mechanisms is controversial within realist circles.

Considering the possibility of using regression analysis to solve realist problems requires addressing at least two questions. First, are randomized controlled trial

CONTACT Heidi Singleton  hsingleton@bournemouth.ac.uk  Bournemouth University, Bournemouth Gateway Building, 12 St Paul's Ln, Bournemouth BH8 8GP  @blueprintteacher

© 2023 Bournemouth University. Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

methodologies compatible with the epistemological assumptions of critical realism? Second, even if they are, can we use statistical methods to understand mechanisms and their relations, rather than confining them to measuring outcomes?

Before describing the practicalities of applying PLS-SEM, using the example of a randomized controlled trial (RCT) to evaluate the educational effectiveness of a Virtual Reality simulation of a patient with diabetes, we will address these two questions in turn. But first, we wish to situate our argument by outlining the evolution of the relationship between realism and trial methodologies.

Background

Randomized controlled trials are comparative studies in which participants are randomly allocated to two groups: one receiving an intervention, the other receiving standard care or a placebo. The groups are then compared in terms of the outcome of interest, with relevant differences attributed to the impact of the intervention (Blackwood, O'Halloran, and Porter 2010). They have long been regarded as the 'gold standard' method for measuring effectiveness in clinical research. Initially used to evaluate medical and surgical procedures, their remit has expanded to the assessment of 'complex interventions' which involve several interacting components whose effects are influenced by the context in which they are operationalized and the behaviour of those delivering and receiving the intervention. This expansion has facilitated the adoption of RCTs beyond healthcare and into areas such as educational and criminological research. It has also exposed the epistemological weaknesses in RCTs that result *inter alia* from their adherence to a successionist conception of cause and effect, which renders them 'unable to control for the effect of social complexity and the interaction between social complexity and a dynamic system change' (Wolff 2001, 125).

In response to the identification of these weaknesses, there has been a reappraisal of RCT design, which has included an exploration of the possibility of applying realist tenets to RCT methodology. Initially, realist conceptions were introduced implicitly (Medical Research Council 2000; Craig et al. 2006). Later, the possibility of realist RCTs became a matter of explicit discussion and debate. Within the broad church of realism, two main approaches have been adopted by those proposing realist trials. One strand (e.g. Bonell et al. 2012; Jamal et al. 2015) uses realist evaluation, as developed from the work of Pawson and Tilley (1997), as its theoretical base. The other (e.g. Blackwood, O'Halloran, and Porter 2010; Porter, McConnell, and Reid 2017) asserts a critical realist (CR) provenance. This approach argues that RCT methodology has the merit of being able to identify the outcomes resulting from the interaction of a partially controlled configuration of causal mechanisms but does not have the capacity to differentiate mechanisms or determine the consequences of their interaction; that task requires CR analytic techniques.

Several benefits are claimed for the application of CR to RCTs. The first relates to CR's assertion that generative mechanisms are ontologically distinct from the succession of events that they generate, and that, in open systems, patterns of events will typically be co-determined by multiple generative mechanisms. We contend that this enables CR RCTs to move beyond the exclusive concentration of traditional RCTs on the regularity of conjunction to demonstrate the relationship between the hypothesized cause and hypothesized effect (Popper 1975), enabling them to interrogate the processes that generate those relations and to 'identify the components of the intervention, and the

underlying mechanisms by which they will influence outcomes' (Medical Research Council 2000, 3). Moreover, acceptance of the multiplicity of generative mechanisms allows for an expansion of causal explanations from exclusive concentration on intervention mechanisms to include mechanisms embedded in the social and cultural contexts into which an intervention is introduced.

In contrast to traditional RCTs' 'downwards conflation' (Archer and Archer 1995) that places the primary causal onus on external influences upon human behaviour and takes little account of human choice or agency, the CR assertion that 'human individuals are entities with causal powers of their own' (Elder-Vass 2010, 114) mandates the inclusion of agents' causal powers in the explanation of behaviour in addition to the causal powers embedded in the intervention and its context. The contribution of human agency to observed outcomes is consequently an intrinsic component of the explanandum of CR trials. Finally, in line with the emancipatory impulse of CR (Bhaskar 1986), this model adds a normative component to the identification of outcomes by including agents' experiences in addition to rates of behaviour or physical effects.

To reflect this expanded ontology of causal relations, we propose an expansion of the realist evaluation Context + Mechanism = Outcome formula (Pawson and Tilley 1997), in the form of Contextual mechanisms + Programme mechanisms + Agency = Behavioural outcomes + Experiential outcomes (Porter 2015b).

Accepting that RCTs are confined to the important but insufficient function of enumerating aggregate behavioural patterns by means of the conjunction of events, Porter et al.'s solution (2017) has been to propose their supplementation with two additional research strategies. The first involves the development and testing of realist hypotheses about the rules, resources and norms embedded in the intervention and its context, and their relationships. The second involves qualitative investigation designed to uncover how agents exposed to these rules, resources and norms interpret and respond to them, and are experientially affected by them.

While we would argue that strategies designed to develop hypotheses about causal influences and the directions of causality, along with the use of qualitative data to elucidate people's interpretations, responses and experiences, are core to CR evaluation, there remains a disjunction between these mechanisms (and experiential-focussed approaches) and the undifferentiated quantitative approach to behavioural outcomes, which weakens the evidential strength of hypotheses about the causal powers of mechanisms and their contingent configurations. This suggests that, if it were possible, the adoption of statistical procedures capable of numerically measuring the mutual relationships of causal mechanisms to reflect the CR model of causation would significantly improve the coherence of critical realist trials. This is rationale for the adoption of Structural Equation Modelling (SEM). However, before addressing PLS-SEM, it is important to interrogate the assumptions about the relationship between critical realist and trial methodologies upon which its application is based.

Are critical realist trials oxymoronic?

Given the adherence of critical realism to methodological pluralism (Danermark et al. 2019), it might seem counterintuitive to conceive of methodological approaches such as trials that incorporate regression analysis as being beyond the realist pale. However,

the appropriateness of their use has been contested within the tradition of critical realism (Porpora 2001; Næss 2004). There are at least two significant challenges from a critical realist perspective to the very notion of critical realist trials. The first challenge involves the imputation of a categorical distinction between realism's focus on causal mechanisms and correlational statistics' concentration on event regularities. The second challenge questions the scientific capacity to generate knowledge of predictive utility, given Bhaskar's insistence 'that criteria for the rational appraisal and development of theories in the social sciences, which are denied (in principle) decisive test situations, *cannot be predictive* and so must be *exclusively explanatory*' (Bhaskar 1989, 21, italics in original).

Tendencies and events

The first challenge relates to the fact that the focus of randomized controlled trials is on event regularities, which refer 'to the *outcome* or *result* of some acting force', while the critical realist focus is on the tendencies generated by causal mechanisms, which involve 'the *force itself*' (Fleetwood 2001, 216). The relationship between these is far from straightforward because, from a critical realist perspective, causal mechanisms

must be analysed as tendencies of things which may be possessed unexercised and exercised unrealised, just as they may of course be realised unperceived ... Thus, in citing a law one is ... not making a claim about the actual outcome (which will in general be co-determined by the activity of other mechanisms. (Bhaskar 1989, 9–10)

However, this observation does not lead Bhaskar to conclude that the use of correlational statistics is detrimental to the realist conception of science. He argues that in the special circumstances of experimental closure, it is possible to identify specific outcomes generated by specific mechanisms: 'The aim of an experiment is to get a single mechanism going in isolation and record its effects' (Bhaskar 2008, 53). It should be noted that this does not involve the conflation of laws with the succession of events. Because experimental conditions are the result of human manipulation,

the intelligibility of experimental activity presupposes the categorical independence of the causal laws discovered from the patterns of events produced ... [I]n an experiment we produce a pattern of events to identify a causal law, but we do not produce the causal law identified. (Bhaskar 2008, 34)

The argument that experimental science can identify causal laws through the production and analysis of patterns of events is significant, given that the rejection of correlational approaches is central to the arguments of realists who are sceptical about the possibility of realist RCTs. Mahoney argues that causal mechanisms 'are posited relations or processes that the researcher imagines existing; they do not refer to any set of empirical conditions. Consequently, they cannot enter a correlational analysis as an empirical state to be measured' (Mahoney 2001, 581). From a Bhaskarian perspective, Mahoney's position is one of Kantian 'transcendental idealism'. Bhaskar's conception of transcendental realism shares with transcendental idealism the requirement that 'creative model building, in which plausible generative mechanisms are imagined producing the phenomena in question' (Bhaskar 2008, 15). But it insists on a further step 'in which the reality of the mechanisms postulated are subjected to empirical scrutiny' (Bhaskar 2008, 15). It is this step that makes it realist rather than idealist. Nor is this simply a matter of philosophical

nicety. As Hawkins (2016, 273) observes, in the absence of empirical testing, realist explanations are open to ‘the fallacy of hasty generalisation’ (positing explanations for patterns that do not in fact explain anything about the underlying reality that caused them) and, ironically, ‘the attribution error’ (the inability to attribute changes to the intervention rather than other factors).

Causation in the open

While the above discussion may demonstrate that the use of correlational statistics is consonant with CR assumptions about the how causal powers can be demonstrated through closed-system experimentation, that does not necessarily justify their application outside experimental conditions. Indeed, for Bhaskar, experimental production and control were prerequisites for the generation of this type of knowledge: ‘Only if ... the system is closed can scientists in general record a unique relationship between the antecedent and consequent of a law like statement’ (Bhaskar 2008, 53). Given that experimental closure in the manner of natural scientific experiments is not applicable to RCTs, which are applied in semi-open systems that include agents capable of interpretation and choice, can they be accommodated within what appears to be a binary distinction between natural/experimental/closed and social/non-experimental/open modes of scientific activity?

In response, the first point to note is that this is an epistemological problem. In terms of ontology, while Bhaskar’s (1989) conception of causality disarticulates the activity of causal mechanisms from actual outcomes because these are generally co-determined by the activities of multiple mechanisms in various configurations, this does not negate the causal consequences that result from these configurations. Rather than denying that actual outcomes are generated by causal mechanisms, Bhaskar is asserting that in open systems there is no invariable connection between the activity of a specific mechanism and a specific outcome.

The second point to note is that there are natural sciences, such as astronomy, that do not (cannot?) rely on experimental closure (Benton 1988, 19), and there are micro-level social sciences, such as experimental psychology, that come close to doing so (Porter 2015a). Moreover, ‘it is not the case that the only conceivable alternative to well controlled experimentation, or more generally the production of closure, of strict event regularity, is a totally unsystematic, incoherent, random flux’ (Lawson 2013, 149). Indeed, we would suggest that there are good reasons to suppose that neither end of this continuum is achievable, and that it is more fruitful to conceive of different degrees of openness and closure (Næss 2004). Between the two absolutes lies what Lawson has described as demi-regularities.

A demi-regularity, or *demi-reg* for short, is precisely a partial event regularity which *prima facie* indicates the occasional, but less than universal, actualization of a mechanism or tendency, over a definite region of time–space. The patterning observed will not be strict if countervailing factors sometimes dominate or frequently co-determine the outcomes in a variable manner. But where demi-regs are observed there is evidence of relatively enduring and identifiable tendencies in play. (Lawson 2013, 149)

The capacity of social sciences to identify ‘relatively enduring and identifiable tendencies’ stands in contradistinction to Bhaskar’s pessimistic assertion that they cannot, in principle, be predictive.

Lawson (2013, 156–159) proposes the methodological strategy of contrast explanation for the social scientific interrogation of demi-regs and the mechanisms that cause them. This involves (1) identifying ‘contrastive demi-regs’ in contexts where it is not immediately apparent why they differ; (2) positing a mechanism, the action of which could account for the difference; (3) deducing the effects that would follow if the hypothesis about the mechanism were true; (4) empirically testing for the deduced effects; and (5) exploring the conditions under which the mechanism is active and how these affect its actions.

While Lawson has developed contrast explanation as a method to be used in economic science which, because of its macro-social subject matter, cannot aspire to experimental closure, he also notes that contrast explanation is a general scientific strategy that encompasses controlled experimental procedures. Accepting that the level of openness/closure of systems lies along a continuum, Porter, McConnell, and Reid (2017) have argued that RCTs’ focus on meso-level systems enables them to use methodological strategies such as intervention and partial contextual control, randomization, blinding and recruitment of large numbers of participants, to actively produce an increase of the frequency of demi-reg instantiation by reducing the influence of countervailing factors. Given that identifying outcomes is an important part of understanding the powers of mechanisms, RCTs can contribute to stages (1) and (4) of Lawson’s contrast explanation – identification of demi-regs and empirical testing of the deduced effects of mechanisms.

While trials exert a degree of control, the disjunction between the forces themselves and the outcomes of those forces, combined with the impossibility of establishing comprehensive knowledge of (still less control over) all the mechanisms involved in open systems necessarily limits their explanatory power. The outcomes generated by RCTs relate to an undifferentiated configuration of contextual, intervention, and agential mechanisms. This creates several epistemological problems. These include difficulties in identifying the causal contribution of the various mechanisms embedded in the intervention being evaluated. In relation to context, to the extent that researchers control it to focus on the effects of the intervention within a single context, they are unable to account for the consequences of contextual variability. Just as problematic is control over participant variability (through a combination of extensive recruitment and exclusion criteria), which entails inattention to agency. The consequence of these controls and simplifications is that RCTs’ contribution to evaluation is largely confined to establishing the *ceteris paribus* efficacy of the aggregation of intervention mechanisms. Once again, we are drawn back to the need for evaluation strategies capable of identifying salient mechanisms, their configurations, and the responses of human agents to them, to complement the traditional focus of trials on the outcomes produced.

The feasibility of partial least squares structural equation modelling

Having discussed the cogency of the general concept of critical realist trials, the next step of the argument is to interrogate the ability of powerful regression analysis tools such as PLS-SEM to identify mechanisms and their interactions and thereby contribute to CR trial methodology. This depends upon their capacity to meet at least three requirements:

1. They are philosophically compatible with realist assumptions about the nature of mechanisms and their causal powers.

2. They are practically compatible with RCT design.
3. They have the statistical capacity to provide useful and meaningful knowledge about the relationships involved in, and effects of, complex causal configurations of mechanisms embedded in contexts, interventions and agents.

Philosophical compatibility

If we consider Hall and Sammons' (2013) definitions of three major approaches to regression analysis, they demonstrate a striking consonance with realist conceptions of causation. The descriptions of statistical mediation and moderation relate closely to the fundamental realist evaluation question of 'what works for whom in what circumstances?'. Mediation analysis, concerning 'mechanisms of effect' (Hall and Sammons 2013, 268), is designed to uncover what works, while moderation analysis asks 'Under what conditions/for whom/when is a pre-established causal relationship observable?' (268). Finally, their definition of interaction analysis as 'a two-tailed hypothesis implying that two or more concepts, "work together" or, "have a combined effect" in eliciting a third' (2013, 268) demonstrates clear pertinence to the realist conception of configurational causation.

We do not want to push our arguments about compatibility too far. Methods involving probabilistic controls and inferential statistics are associated with deductivist methodologies that conflate causation with stochastic event regularity, which is incompatible with realist conceptions of causation in terms of powers and tendencies. However, we already have an example of the disarticulation of method from methodology in Bhaskar's redescription of experimental procedures in realist rather than positivist terms. Using a similar strategy, rather than assuming that the demi-regularity of outcomes is the result of the stochasticity of event conjunctions (as positivism conceives them), they can be seen as reflecting variation in the configuration of causal mechanisms (as characterized by CR).

If this redescription of the relationship between tendencies and events is accepted, it opens the way to explore the extent to which the use of SEM can expand the contribution of statistical analysis beyond the empirical identification of outcomes that constitute stages (1) and (4) of contrast explanation to support analysis at stage (5) – exploration of the interrelationship of mechanisms and even perhaps to assist with the retroductive work involved in stages (2) and (3) by providing empirical support for the existence of hypothesized mechanisms and demonstrating their effects.

Compatibility with randomized controlled trial design

The scope of statistical analysis is circumscribed by the data it must work with. For regression analysis to be able to make meaningful statements about the causal powers of mechanisms and their configurations, realist trials need to be designed in such way that they facilitate the gathering of quantitative data that reflect these relations of interest. This may sound daunting, but it should be noted that the following recommendations simply involve the adoption of pragmatic trial design recommended by Dal-Ré, Janiaud, and Ioannidis (2018).

It is important that the context in which a trial is conducted should be characterized by organization and practice that conform as closely as possible to that which would pertain in the trial's absence. Measurement of the effects of artificially created and transient contextual mechanisms contributes little to our understanding of real-world processes (Hansen and Jones 2017). The need to take account of contextual complexity implies that realist RCTs should also be conducted on multiple sites. The evaluation of the effect of an intervention on a single site requires control of contextual variation over time, preventing comparative examination of contextual mechanisms. The theoretically informed selection of different sites based on the contextual mechanisms that have been hypothesized to be of most pertinence would allow for cross-site comparison of the configurational contribution of these mechanisms, assuming the intervention mechanisms are held constant across sites. In short, contrary to the objection that 'context ... cannot be allocated' (Hawkins 2016, 279), we argue that it can and should be systematically allocated.

In relation to the interpretation and responses of participants, it is important to be as parsimonious as possible in the application of exclusion criteria. The types of participants included in the trial must be representative of the variety of the population to which the intervention is likely to be applied (Heiat, Gross, and Krumholz 2002). This would facilitate identification of the responses and experiences of different types of people to the external mechanisms to which they are exposed, thus ensuring that agency becomes a significant focus of trials. Once again, this proposal stands in contradistinction to Hawkins's (2016) concern that trials lack sensitivity because they include a wide variety of people solely based on their perceived needs, rather than by sampling according to propensity to benefit. On the contrary, we contend that the application of regression techniques such as PLS-SEM to a large and varied sample population, in combination with realist thematic analysis relating to the 'what works for whom', enables realist trials to empirically confirm the types of people who have a greater or lesser propensity to benefit.

Capacity to provide knowledge about complex causal configurations

The claim that Structural Equation Modelling (SEM) has the capacity to enhance statistical descriptions of event regularities to provide useful and meaningful knowledge to support the realist identification of complex causal configurations of mechanisms embedded in contexts, interventions, and agents rests on its application of advanced multivariate data analysis methods (Hair et al. 2017). SEM is used to concurrently examine and approximate complex causal relationships amongst variables, even when associations are not directly evident (Williams, Vandenberg, and Edwards 2009). Simultaneously uniting factor investigation and linear regression models, SEM permits the scrutiny of associations between variables by computing directly discernible indicator variables (Hair et al. 2014). The crucial question is, however, how variables, as connoted in SEM analysis, relate to mechanisms, as connoted by critical realism.

Variables and mechanisms

The first point to note is that variables are not the same thing as mechanisms, in that variables describe relationships, while mechanisms explain why they are related. To conflate

them undermines the explanatory power of mechanisms (Astbury and Leeuw 2010). Nonetheless, as Astbury and Leeuw also note, there are parallels between them which means that ‘statistical measurement and analysis can help to identify and describe causal relationships’, thereby providing ‘the raw material for elaboration of theoretical models of the mechanisms that explain how statistical associations are generated’ (Astbury and Leeuw 2010, 367–8). We suggest that the use of multivariate analyses such as PLS-SEM bring these parallels closer together. This can be seen if we take the three core attributes of mechanisms identified by Astbury and Leeuw (2010).

The authors describe the first characteristic of mechanisms as hidden, which means that an explanation of outcomes needs to involve going ‘below the “domain of empirical,” surface level descriptions of constant conjunctions and statistical correlations’ (Astbury and Leeuw 2010, 368). One of the significant attributes of multivariate analyses such as PLS-SEM is that they are not confined to addressing the realm of ‘superficial appearances’ (Pratschke 2003). With their ability to identify latent variables, they can be used as tools ‘for discovering mechanisms and factors of influence not manifest in the domain of the actual (that is, as an immediately visible pattern of events)’ (Næss 2004, 150), thereby enhancing the capacity of statistical analysis to describe causal relationships. This mode of SEM analysis, which uses reflective measurement models whereby the putative flow of causation runs from the unobservable to the observable (Ford et al. 2018), is consonant with CR’s causal criterion of reality which deems as real entities with the capacity to bring about changes in material things (Bhaskar 1989), thus providing ‘an additional means of bridging the gap between theoretical and statistical models’ (Pratschke 2003, 23).

A similar observation can be made in relation to the second attribute of mechanisms identified by Astbury and Leeuw – that they are sensitive to variations in context (though it may be more precise to state that it is their effects on outcomes that are sensitive to context). The multivariate capability of PLS-SEM to identify multiple variables and their relationships once again enhances the capacity of statistical analysis to assist in the identification and description of co-determined causal relationships. Indeed, the use of ‘moderator variables’, which are associated with changes in the level or direction of relationship between other variables, allows for the testing of complex causal relationships.

The third characteristic of mechanisms identified by Astbury and Leeuw, that they generate outcomes, reinforces the difference between them and variables in that, while variables denote occurrences that are associated with outcomes, they do not define the forces that generate those outcomes. However, that does not mean that they cannot provide useful evidence about those forces. Most pertinent here are ‘mediator variables’ which come closest to the realist definition of mechanisms. As Baron and Kenny (1986, 1176) put it, ‘Whereas moderator variables specify when certain effects hold, mediators speak to how and why such effects occur’. That said, they remain correlational variables that, rather than defining the mechanisms involved, provide evidence about the effects of hypothesized mechanisms.

The suggestive and confirmatory roles of regression analyses

In the light of the observations above, we suggest that there are two possible roles for regression analysis in CR investigations. The first is suggestive. As Naess (2015, 150) notes: ‘The controlled effects of independent variables may hint at possible causal

relationships that may subsequently be explained by theoretical reasoning and qualitative empirical research’.

The second role is confirmatory, in that they can be used as evidentiary tools to enable the assessment of theoretical explanations that have been previously formulated.

Whether it is used before or after realist theorizing, the key attribute of regression analysis is that it is evidentiary rather than explanatory (Porpora 2001). This is not to downplay the importance of such a function, we contend that the strength of analytic strategies such as PLS-SEM is they can be an important means of providing that control: ‘in open systems, regularities detected by analytical statistics can be as indicative of active mechanisms as are regularities detected in the experimental laboratory. No more actualism is implied in one case than the other’ (Porpora 2001, 262).

Covariance-Based or partial least squares structural equation modelling?

Ours is not the first attempt to interrogate the application of structural equation modelling to realist research. Ford et al. (2018) have used structural equation modelling to quantify Context, Mechanism, Outcome configurations via a covariance-based approach to the secondary analysis of large data sets. This is an important contribution that we have endeavoured to build upon. Covariance-Based SEM enables testing and comparison of alternative theories, particularly where the structural model has circular relationships. However, it has limitations that include the need for the data to be normally distributed and for error terms to require additional specification (Hair et al. 2017).

We believe that PLS-SEM may be more appropriate than Covariance-Based SEM in certain circumstances. Hair et al. (2017) recommends the use of PLS-SEM rather than Covariance-Based SEM when the following criteria are met: a relatively small sample; inclusion of numerous categorical ordinal measured variables with dubious, and probably assumption-violating distributional properties, and where indirectly measured variables are part of the structural model.

With PLS-SEM, structural models can be more complicated and include many directly measurable variables (known as indicators) and indirectly measurable variables (known as constructs). Dissimilar to the covariance-based SEM approaches, which necessitate a multivariate normal distribution of the observed variables, PLS is founded on the resampling measures of bootstrapping, which does not make parametric assumptions (Henseler, Ringle, and Sarstedt 2016). It is thus a preferable method for theory development.

The promise of partial least squares structural equation modelling

PLS-SEM defines the ‘parameters of an equation set in a path model by combining principal component analysis to assess the measurement models with path analysis to estimate the relationships between latent variables’ (Hair et al. 2017, 3). It is a causal-predictive approach to SEM that stresses prediction in estimating statistical models (Avkiran 2018). It allows investigators to model and approximate compound cause–effect association models with both latent and observed variables.

The latent variables exemplify unobserved (i.e. not directly measurable) occurrences such as opinions, attitudes, and intentions. The observed variables (e.g. responses to a survey) are used to indicate the latent variables in a statistical model. PLS-SEM estimates

the relationships between the latent variables (i.e. their strengths) and regulates how well the model explains the target constructs under examination (Hair et al. 2017).

Argues that PLS-SEM can: (1) establish data equivalence via the three stage Measurement Invariance of Composite Models Procedure (MICOM) process (Henseler, Ringle, and Sarstedt 2016), to minimize measurement error, (2) detect the significance and performance of latent antecedent variables to target areas for further research (Hair et al. 2017), and (3) unearth unobserved heterogeneity to facilitate scrutiny of structural and measurement models. Consequently, when considering complex research models, the various flexible analytic possibilities, limited suppositions, and user-friendliness are significant advantages of using PLS-SEM.

The advent of computer software including SMART-PLS V 3.0™ enables automatic computation of functions, including f^2 effect size, multi-collinearity assessment and different types of invariance testing (Henseler, Ringle, and Sarstedt 2016). The latter enables analysis of direction and strength of relationships between variables not only throughout the control or experimental group pathway model, but also through simultaneous analysis across the entire data set. Thus, researchers can spot if specific relationships in pathway models are significant and how these differ between the experimental and control groups.

It has thus far been argued that PLS-SEM is a suitable vehicle for combining with CR evaluation because it enables not only the examination of observed variables but also identification and exploration of latent variables (Hair et al. 2017). However, causal accounts also require the examination of interventions from a systems viewpoint with a case-based (i.e. configurational), rather than a variable-based positioning (Byrne and Uprichard 2012). This is based on the CR focus on the interaction of causal mechanisms, rather than considering them in isolation with the assumption that they do not affect each other and/or that they act independently.

One of the benefits of the PLS-SEM technique is the unlimited integration of latent variables in the path model that incorporates either the reflective or formative measurements models. PLS path models comprise three constituents: the structural model, the measurement model, and the weighting scheme. Path coefficients are the main products of PLS-SEM which measure the hypothesized relationships within the structural model. This unlimited approach to inclusion of variables/mechanisms allows for testing of multiple configurations of mechanisms embedded in context, programme, and agents (see Figure 2). One drawback of using PLS-SEM is that it is only useful for structural models that do not have casual loops (circular relationships), hence only linear models are suitable for this type of statistical analysis.

While structural and measurement models are constituents of all kinds of SEMs with latent constructs, the weighting system specific to the PLS method means that precise points of action can be identified within configurations. We argue that this strengthens the change in focus of CR trials from simply ascertaining whether an intervention is effective to consider why, how, with whom, and in what circumstances it is likely to work certain situations.

We are not asserting that, as a tool for analysing causality in open systems involving differing arrangements of numerous generative mechanisms, PLS-SEM is uniquely compatible with CR. Rather, we are suggesting that it has at least three pragmatic advantages over other SEM models (and indeed other forms of multivariate statistical modelling). First, it enables the testing of the various parts of these configurations; second, by

enabling an unlimited inclusion of variables, it has the power to take account of a multiplicity of mechanisms and their configurations; and third, it has the capacity to measure the direction and strength of the relationships between each aspect of the configuration.

An example of CR PLS-SEM analysis

Our example of combining CR with PLS-SEM involves an RCT conducted to evaluate the educational effectiveness of a Virtual Reality (VR) simulation of the deterioration of a patient with diabetes. Diabetes and its treatment are complex, and for health care professionals, insufficient knowledge about diabetes treatments can result in the serious negative consequences of hypoglycaemia or hyperglycaemia. Simulations are increasingly being offered as part of the educational experience and valued for their more authentic approaches in preparing for live clinical experience (Bayram and Caliskan 2019). A non-immersive (laptop accessed) VR deteriorating diabetes patient simulation was designed and evaluated to address the gap between theory and practice.

Our RCT was designed to answer the research question: 'Does use of VR make a significant difference to students' knowledge, when compared to traditional learning methods in Higher Education?' The trial included Second Year Adult and Mental Health Nursing students at a UK University. The structural model we have selected for illustrative purposes is very simplistic and was chosen to exemplify our argument. It should be noted that RCTs usually have more configurations and might have many more than one outcome, and that PLS-SEM can handle such data. The inclusion of higher-order constructs is also possible. The trial involved three stages of analysis. The first stage consisted of the identification of causal configurations through a realist review. The second stage involved the gathering of empirical data, via a pre and post-test surveys designed to match the configural (or pathway model), and its analysis using PLS-SEM. The third stage involved mapping these results explicitly on to the hypothesized causal configurations of contextual, programme and agential mechanisms and the outcomes these produced. The example worked through here was part of a larger research study that followed a mixed-methods approach that also included both the development of causal theories through realist reviewing and qualitative analysis of people's interpretations and experiences (Singleton et al. 2021; Singleton et al. 2022).

Stage 1: programme theory development

The causal influences and directions of causality that the statistical model was based on were derived from a CR review of the literature (Singleton 2020), which identified several configurations involving relationships between contextual mechanisms, programme mechanisms, agential mechanisms, and outcomes (Porter 2015b). We select one of these for illustrative purposes. Configuration was proposed for the CR-RCT for the current study. The components of the causal configuration are given in Table 1.

Stage 2: empirical testing and analysis using PLS-SEM

The pre-test was formulated using the findings of the CR review. Indicator variables (shown in rectangular boxes on the pathway model, in Figure 1) were linked to each

Table 1. CPAO programme theory.

CPAO configurations	Contextual mechanism	Programme mechanism	Agential mechanisms	Outcome
CPAO	Amount of diabetic nursing EXPERIENCE required to complete exercise and understand complex diabetic concepts.	The use of VR to encourage direct ENGAGEMENT in a non-stressful environment than experienced in real-life encounters.	CONFIDENCE boosting via feedbacks. IMMERSION in VR	Improved KNOWLEDGE of complex diabetic issues

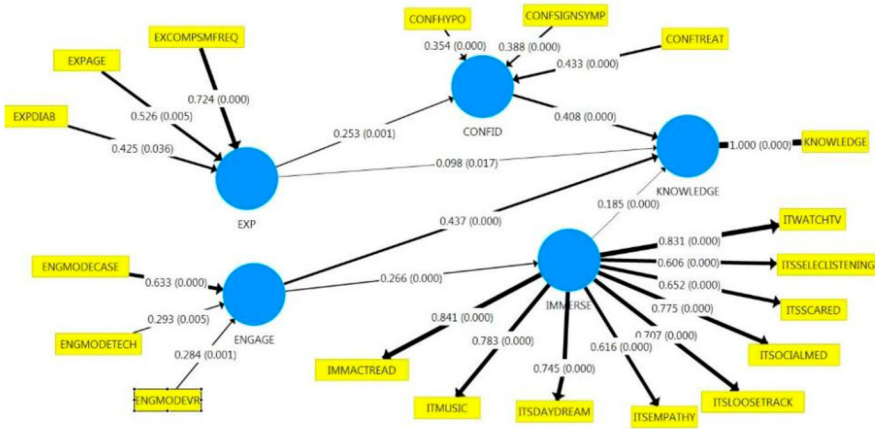


Figure 1. A simple pathway model showing both full and partial mediation of variables.

question of the survey. Following the pre-test, the experimental group ($n = 83$) practiced on the VR simulator patient case study, while controls ($n = 88$) completed a paper-based version. Both groups completed a post-test. Evaluation included Partial Least Square-Structural Equation Modelling (Hair et al. 2016).

As can be seen in Table 2 both Configural and Compositional invariances were established. MICOM Step 2, demonstrated measurement model invariance between groups; in other words, both groups tackled the survey in the same way conceptually. The permutation test results (5000 permutations) demonstrate that the mean value and variance of a composite in the Experimental Group do significantly diverge from the results in the Control Group. The permutation test reliably controls for Type 1 errors when the assignment of observations occurs randomly, as is the case when using SMART-PLS 3.0. A two-tailed 95% permutation-based confidence interval was created. According to Hair et al. (2018, 154) if the original difference (d) of the group-specific path coefficient estimates does not fall into the confidence interval, it is statistically significant. The engage→immersion pathway had an original difference of 0.278; as this does not fall between the confidence intervals of -0.255 and 0.262 , it is statistically significant.

Our main findings were that prior computing and diabetic nursing experience were invariant between groups. Mean scores for the experimental group were higher than for the control group, for the latent variables of ‘Confidence’, ‘Engagement’, and ‘Immersion’. The experimental group was found to be significantly more knowledgeable about the correct use of the hypoglycaemia treatment box than the control group.

Table 2. MICOM results.

Summary of the MICOM via PLS-SEM results

MICOM Step 1 – Configural invariance				
Configural invariance established? YES				
MICOM Step 2 – Compositional invariance established- YES				
Composite	Correlation c	5% quartile of the empirical distribution of c_{μ}	P value	Compositional invariance established?
Confidence	0.993	0.842	0.876	Yes
Engage	0.969	0.874	0.497	Yes
Experience	0.975	0.506	0.860	Yes
Immerse	0.994	0.968	0.842	Yes
Knowledge	Singlet item measure – hence outer relationship is 1 by design.			
MICOM Step 3 – Assessment of equality of composite means and variances				
	Difference of the composites mean value (=0)	Permutation p value	Equal mean values?	
Confidence	1.347	<0.001	No	
Engage	1.224	<0.001	No	
Experience	0.268	0.062	Yes	
Immerse	-0.455	0.003	No	
Knowledge	1.166	<0.001	No	
	Logarithm of the composite's variances ratio (=0)	Permutation p value	Equal variances?	
Confidence	-0.694	<0.001	No	
Engage	0.034	0.827	Yes	
Experience	0.247	0.367	Yes	
Immerse	0.003	0.992	Yes	
Knowledge	0.248	0.179	Yes	

By using SMART-PLS 3.0 software, we were also able to extrapolate the following more nuanced findings. Across the data set (post-test survey) all pathways were found to be significant. Hence, the experience (including prior computing and diabetic nursing experiences) to knowledge pathway was fully mediated by confidence. This means that for participants involved in the study, the more previous experience they had, the more confidence they had when using either diabetic intervention (either paper or VR) and then the higher their knowledge scores post intervention. This upper section of the structural model maps to the causal configuration.

In the lower half of the diagram, across the data set the engagement to knowledge pathway was partially mediated by immersion, and again these were significant pathways across the whole data set. Across groups, however, the only significant pathway in the conceptual model was the engagement to immersion pathway. This might indicate that the improved knowledge scores for the Experimental Group have resulted from this 'key pathway', which could be deemed to be the 'action point' of the PLS-SEM model. Moreover, engagement can be viewed as being the 'driver' construct in this conceptual model. In other words, one possible theory is that for the Experimental Group, students became more engaged in their learning and therefore more immersed in their learning than those learners who did not use VR (the Control Group); and this resulted in them attaining higher knowledge scores (as measured by a set of ten MCQs) than the Control Group.

The latent variable prior experience was not found to differ between groups statistically significantly. This means that when using the VR simulation, a students' prior experience, including if they had used VR before, etc., did not impact on their MCQ scores. This

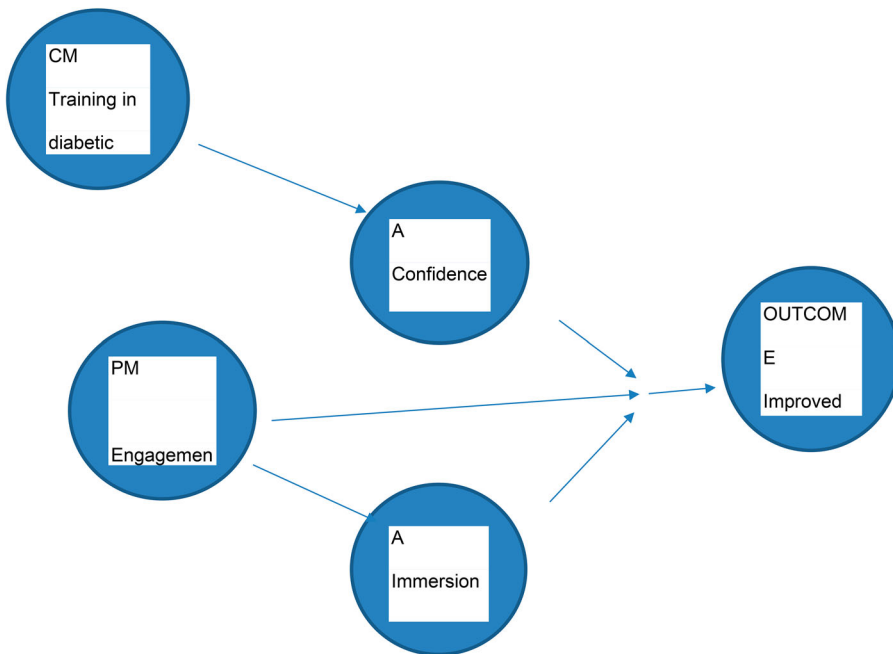


Figure 2. Programme theory including theorized causal relations.

indicates that the VR simulation is an inclusive learning tool, regardless of their age, computing experience or diabetic nursing experience. It means that it is a suitable learning method for all students (provided they are not susceptible to nausea or migraines).

Stage 3: conceptualizing results in CR terms

The third stage involved back-translating the PLS-SEM variables to CR mechanisms. In our example, PLS-SEM enabled quantification and testing (using our primary data) of our hypothesized causal configuration to support our inferences about combinations and strengths of relationships, and the direction of causal flow. Merging our Smart-PLS diagram with our CR approach (see Figure 2), we have developed the following representation. Note that this amalgamated diagram captures the interactions and effect of programme mechanisms and hence the fuller concept of causality. Thus, the contextual affordances of previous training in diabetic care had little direct effect on knowledge, but were indirectly effective by fostering participants' confidence, the agential mechanism that directly led to improved knowledge. In the lower segment, while the programme mechanism of generating engagement had a powerful direct effect upon knowledge levels, these were further empowered through interaction with the level of immersion that participants experienced. In sum, the outcome observed was the result of the interaction between a contextual mechanism, a programme mechanism and two agential mechanisms.

Discussion

This paper adds to the debate about the possibility of conducting realist RCTs and proposes one method for conducting them by positing PLS-SEM as a method for the

quantification of causal configurations. We contend that PLS-SEM moves the argument forward in relation to the capacity of statistical analysis to embrace 'holistic causality'. Most significantly, the unique pathway weighting approach of PLS-SEM enables the evaluation of the strengths and directions of causal relationships. This weighting technique facilitates the pinpointing of statistically significant 'points of action' across the pathway model and between trial arms (e.g. control versus experimental). This very precise evaluation has the potential to confirm direction for researchers and practitioners in terms of how an intervention has worked but moreover, for 'whom, when, where and in what circumstances' that intervention has worked. This aspect also has the capabilities to pinpoint any 'driver constructs' (or mechanisms). In future research, we recommend that more complex models are tested that incorporate more causal configurations, and that structural models that comprise higher-order constructs are tested, to establish whether these types of models can be aligned with CR configurations.

While our central aim was to consider the possibility of critical realist RCTs, analysed via PLS-SEM, as a viable approach to the evaluation of complex interventions in open systems, several cautions require to be addressed. Some of these relate to the possibility of categorical errors. For example, we suspect that independent variables frequently correspond to the intervention under examination rather than the mechanism or mechanisms that are embedded in it. Van Belle et al.'s (2016) observation, that the mislabelling of interventions as mechanisms is a common confusion, indicates the need for realist analysis to ensure clear demarcation between them when using statistical data as a tool for the identification of mechanisms.

Other cautions relate to the specific technicalities of using PLS-SEM. First, it is appreciated that not all trials will be suitable for data analysis via this approach (Hair et al. 2017). Second, it is conceded that Covariance-Based SEM may better enable testing and comparison of alternative theories than PLS-SEM (which has a more developmental function). However, this is counterbalanced by Covariance-Based SEM's restriction to the analysis of multivariate normal distribution of the observed variables, which restricts its real-world application. More investigation is required into whether the pragmatic advantages of PLS-SEM outweigh its epistemological constraints.

A more fundamental caution needs to be raised in relation to the danger of fetishizing mathematical strategies like PLS-SEM by bestowing on the numerical data that they generate an exactitude that is not merited when describing social processes (Ron 2002). To counter such tendencies, we need to keep in mind that, at an ontological level, the regularities manifest in partly open/partly closed systems are 'of degree, ranging from the very weak to the very strong' (Karlsson 2011, 161), and, at an epistemological level, regression analysis needs to be used in a theory-informed way that takes account of human reflexivity (Næss 2004).

Notwithstanding these caveats, it has been our intention to contribute to the methodological conversation about the possibility of critical realist trials as a form of contrast explanation (Lawson 2013), and to illuminate the contribution that statistical analysis in the form of PLS-SEM might make to aiding (rather than replacing) the retroductive work and qualitative analysis of critical realist evaluation. While PLS-SEM's role is evidentiary rather than explanatory, the importance of the evidence it provides should not be underestimated, in that, when combined with realist strategies, it can contribute to the development 'empirically tested, refined theoretical explanations of complex

phenomena' (Brown and Singleton 2023, 502). By exercising empirical control over realist causal theories, regression analysis, in combination with qualitative strategies to uncover agents' responses and experiences, enhances confidence in the prediction of outcomes. The conception of prediction that we are using is 'modest' (Næss 2004), in that we accept that predictions in partially open/closed systems are circumscribed both by the contingency of causal configurations, and by the reflective capacity and agential powers of the people involved. However, the fact that the predictive capacity of critical realist trials is modest, relating to tendencies rather than constant conjunctions, makes them no less useful in providing guidance for future action.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Data availability statement

Please contact the corresponding author to view the available data.

Notes on contributors

Dr Heidi Singleton, Having trained as a Children and Young People's Nurse, she worked as a Research Assistant in the field of nursing. Next, obtaining a Post Graduate Certificate of Education and a Master's Degree in Education, and has just completed Doctoral studies at Bournemouth University. Heidi is now a lecturer, working across a range of research projects all of which focus on either digital therapeutics (virtual reality distraction and virtual reality exposure therapy as psychological interventions) or technology enhanced education (e.g. virtual reality deteriorating patient simulations).

Professor Sam Porter is a sociologist by academic training and a nurse by profession. His wide research interests reflect this combination. His main areas of interest are palliative and end-of-life care; supportive care for cancer survivors and carers; maternal and child health; the sociology of health professionals; and the use of arts-based therapies. Sam has adopted a wide range of methods in his research, ranging from randomized controlled trials to the use of qualitative approaches such as ethnography. He is a strong advocate of realist approaches to knowledge, which provide a robust rationale for the use of mixed methods. In his advocacy of a realist philosophy of science, he has enjoyed a number of contentious debates.

Dr John Beavis is a retired academic formally working at Bournemouth University. His area of speciality was Archaeology and statistics. John provided the PLS-SEM expertise for this project. He has published several books including *Communicating Archaeology* (1999) by John Beavis and Alan Hunt.

Professor Liz Falconer, After more than 30 years working in Higher Education in the UK, Liz is now a consultant in the use of virtual and augmented reality technologies in heritage and archaeology, with a particular emphasis on the enhancement of public understanding of ancient sites. She is a Visiting Fellow in the Centre for Archaeology and Anthropology at Bournemouth University, UK. She has particular research interests and practitioner expertise in 3D virtual reality, virtual worlds, video and machinima, and the affordances these environments offer for a wide range of simulations

Dr Jacqueline Priego Hernandez is a social psychologist whose expertise lies at the intersection between education and healthcare interventions for the psychosocial wellbeing of youth in socially excluded communities. Having worked in a number of research projects with communities in Latin America, India and the UK, Jacqueline has a keen interest in the use of participatory and creative methodologies for the study of psychosocial issues. She is also interested in systematic review methods and approaches to the systematization of multi-country qualitative datasets.

Professor Debbie Holley is Professor of Learning Innovation at Bournemouth University. Her expertise lies with blending learning to motivate and engage students with their learning inside /outside the formal classroom, at a time and place of their own choosing. This encompasses the blend between learning inside the classroom and within professional practice placements, scaffolding informal learning in the workplace. Debbie writes extensively the affordances of technologies such as Augmented Reality, Virtual/ Immersive Realities and Mobile Learning.

ORCID

Heidi Singleton  <http://orcid.org/0000-0002-3487-1029>

Sam Porter  <http://orcid.org/0000-0002-6950-9388>

References

- Archer, Margaret S., and Margaret Scotford Archer. 1995. *Realist Social Theory: The Morphogenetic Approach*. Cambridge University Press.
- Astbury, Brad, and Frans L. Leeuw. 2010. "Unpacking Black Boxes: Mechanisms and Theory Building in Evaluation." *American Journal of Evaluation* 31 (3): 363–381. doi:10.1177/1098214010371972
- Avkiran, N. K. 2018. "Rise of the Partial Least Squares Structural Equation Modeling: An Application in Banking." *Partial Least Squares Structural Equation Modeling: Recent Advances in Banking and Finance*: 1–29.
- Baron, Reuben M., and David A. Kenny. 1986. "The Moderator–Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations." *Journal of Personality and Social Psychology* 51 (6): 1173. doi:10.1037/0022-3514.51.6.1173
- Bayram, S. B., and N. Caliskan. 2019. "Effect of a Game-Based Virtual Reality Phone Application on Tracheostomy Care Education for Nursing Students: A Randomized Controlled Trial." *Nurse education today* 79:25–31.
- Benton, T. 1988. "Realism and Social Science: Some comments on Roy Bhaskar's 'The Possibility of Naturalism'." In *Critical Realism* (pp. 297–312). Routledge.
- Bhaskar, R. 1986. *Scientific Realism and Human Emancipation*. London: Verso.
- Bhaskar, R. 1989. *The Possibility of Naturalism: A Philosophical Critique of the Contemporary Human Sciences*. 2nd edn. Hemel Hempstead: Harvester Wheatsheaf.
- Bhaskar, R. 2008. *A Realist Theory of Science*. 2nd edn. London: Verso.
- Blackwood, Bronagh, Peter O'Halloran, and Sam Porter. 2010. "On the Problems of Mixing RCTs with Qualitative Research: The Case of the MRC Framework for the Evaluation of Complex Healthcare Interventions." *Journal of Research in Nursing* 15 (6): 511–521. doi:10.1177/1744987110373860
- Bonell, Chris, Adam Fletcher, Matthew Morton, Theo Lorenc, and Laurence Moore. 2012. "Realist Randomised Controlled Trials: A new Approach to Evaluating Complex Public Health Interventions." *Social Science & Medicine* 75 (12): 2299–2306. doi:10.1016/j.socscimed.2012.08.032
- Brown H., and H. J. Singleton. 2023. "Atopic Eczema and the Barriers to Treatment Adherence for Children: A Literature Review." *Nursing Children and Young People* 35(2).
- Byrne D., and E. Uprichard. 2012. "Useful Complex Causality." *The Oxford handbook of philosophy of social science* 109.
- Craig, P., P. Dieppe, S. Macintyre, S. Michie, I. Nazareth, and M. Petticrew. 2006. "Developing and Evaluating Complex Interventions: New Guidance." Medical Research Council. 2017.
- Dal-Ré, Rafael, Perrine Janiaud, and John PA Ioannidis. 2018. "Real-world Evidence: How Pragmatic are Randomized Controlled Trials Labelled as Pragmatic?" *BMC Medicine* 16: 1–6. doi:10.1186/s12916-017-0981-7
- Danermark, B. 2019. "Applied Interdisciplinary Research: A Critical Realist Perspective." *Journal of Critical Realism*. 18 (4): 368–82.
- Elder-Vass, E. 2010. *The Causal Power of Social Structures: Emergence, Structure and Agency*. Cambridge: Cambridge University Press.

- Fleetwood, Steve. 2001. "Causal Laws, Functional Relations and Tendencies." *Review of Political Economy* 13 (2): 201–220. doi:[10.1080/09538250120036646](https://doi.org/10.1080/09538250120036646)
- Ford, J. A., A. Jones, G. Wong, A. Clark, T. Porter, and N. Steel. 2018. "Access to Primary Care for Socio-Economically Disadvantaged Older People in Rural Areas: Exploring Realist Theory Using Structural Equation Modelling in a Linked Dataset." *BMC Medical Research Methodology* 18: 57. doi:[10.1186/s12874-018-0514-x](https://doi.org/10.1186/s12874-018-0514-x)
- Hair, Joe, Marko Sarstedt, Lucas Hopkins, and Volker G. Kuppelwieser. 2014. "Partial Least Squares Structural Equation Modelling (PLS-SEM) An Emerging Tool in Business Research." *European Business Review* 26 (2): 106–121. doi:[10.1108/EBR-10-2013-0128](https://doi.org/10.1108/EBR-10-2013-0128)
- Hair, Joseph F., Marko Sarstedt, Christian M. Ringle, and Siegfried P. Gudergan. 2017. *Advanced Issues in Partial Least Squares Structural Equation Modeling*. Sage Publications.
- Hall, James, and Pamela Sammons. 2013. "Mediation, Moderation & Interaction: Definitions, Discrimination & (Some) Means of Testing." In *Handbook of Quantitative Methods for Educational Research*, 267–286. Brill.
- Hansen, Anders Blædel Gottlieb, and Allan Jones. 2017. "Advancing 'Real-World' trials That Take Account of Social Context and Human Volition." *Trials* 18: 1–3.
- Hawkins, Andrew J. 2016. "Realist Evaluation and Randomised Controlled Trials for Testing Program Theory in Complex Social Systems." *Evaluation* 22 (3): 270–285. doi:[10.1177/1356389016652744](https://doi.org/10.1177/1356389016652744)
- Heiat, Asefeh, Cary P. Gross, and Harlan M. Krumholz. 2002. "Representation of the Elderly, Women, and Minorities in Heart Failure Clinical Trials." *Archives of Internal Medicine* 162 (15).
- Henseler, J., C. M. Ringle, and M. Sarstedt. 2016. "Testing Measurement Invariance of Composites Using Partial Least Squares." *International Marketing Review* 33: 405–431. doi:[10.1108/IMR-09-2014-0304](https://doi.org/10.1108/IMR-09-2014-0304)
- Jamal F., A. Fletcher, N. Shackleton, D. Elbourne, R. Viner, and C. Bonell. 2015. "The Three Stages of Building and Testing Mid-Level Theories in a Realist RCT: A Theoretical and Methodological Case-Example." *Trials* 16: 446.
- Karlsson, Jan. 2011. "People Can not Only Open Closed Systems, They Can Also Close Open Systems." *Journal of Critical Realism* 10 (2): 145–162. doi:[10.1558/jcr.v10i2.145](https://doi.org/10.1558/jcr.v10i2.145)
- Lawson, Tony. 2013. "Economic Science Without Experimentation." In *Critical Realism*, 144–169. Routledge.
- Mahoney, James. 2001. "Beyond Correlational Analysis: Recent Innovations in Theory and Method." In *Sociological Forum*, 575–593. Eastern Sociological Society.
- Medical Research Council (Great Britain). 2000. *A Framework for Development and Evaluation of RCTs for Complex Interventions to Improve Health*. Medical Research Council. Health Services and Public Health Research Board..
- Næss, Petter. 2015. "Critical Realism, Urban Planning and Urban Research." *European Planning Studies* 23 (6): 1228–1244. doi:[10.1080/09654313.2014.994091](https://doi.org/10.1080/09654313.2014.994091)
- Næss, Petter. 2004. "Prediction, Regressions and Critical Realism." *Journal of Critical Realism* 3 (1): 133–64.
- Pawson, R., and N. Tilley. 1997. *Realistic Evaluation*. London: Sage.
- Popper, Karl Raimund. 1975. *Objective Knowledge: An Evolutionary Approach*. Vol. 49. Oxford: Clarendon Press.
- Porpora, Douglas. 2001. "Do Realists run Regressions." *After Postmodernism: An Introduction to Critical Realism*, 260–268.
- Porter, Sam. 2015a. "The Uncritical Realism of Realist Evaluation." *Evaluation* 21 (1): 65–82. doi:[10.1177/1356389014566134](https://doi.org/10.1177/1356389014566134)
- Porter, Sam. 2015b. "Realist Evaluation: An Immanent Critique." *Nursing Philosophy* 16 (4): 239–251. doi:[10.1111/nup.12100](https://doi.org/10.1111/nup.12100)
- Porter, Sam, Tracey McConnell, and Joanne Reid. 2017. "The Possibility of Critical Realist Randomised Controlled Trials." *Trials* 18: 1–8.
- Pratschke, Jonathan. 2003. "Realistic Models? Critical Realism and Statistical Models in the Social Sciences." *Philosophica* 71.
- Ron, Amit. 2002. "Regression Analysis and the Philosophy of Social Science: A Critical Realist View." *Journal of Critical Realism* 1 (1): 119–142. doi:[10.1558/jocr.v1i1.119](https://doi.org/10.1558/jocr.v1i1.119)

- Singleton, Heidi. 2020. "Virtual Technologies in Nurse Education: The Pairing of Critical Realism with Partial Least Squares Structural Equation Modelling as an Evaluation Methodology." PhD diss., Bournemouth University.
- Singleton, Heidi, Janet James, Liz Falconer, Debbie Holley, Jacqueline Priego-Hernandez, John Beavis, David Burden, and Simone Penfold. 2022. "Effect of non-Immersive Virtual Reality Simulation on Type 2 Diabetes Education for Nursing Students: A Randomised Controlled Trial." *Clinical Simulation in Nursing* 66: 50–57. doi:[10.1016/j.ecns.2022.02.009](https://doi.org/10.1016/j.ecns.2022.02.009)
- Singleton, Heidi, Janet James, Simone Penfold, Liz Falconer, Jacqueline Priego-Hernandez, Debbie Holley, and David Burden. 2021. "Deteriorating Patient Training Using non-Immersive Virtual Reality: A Descriptive Qualitative Study." *CIN: Computers, Informatics, Nursing* 39 (11): 675–681. doi:[10.1097/CIN.0000000000000787](https://doi.org/10.1097/CIN.0000000000000787)
- Van Belle, Sara, Geoff Wong, Gill Westhorp, Mark Pearson, Nick Emmel, Ana Manzano, and Bruno Marchal. 2016. "'Can Realist' Randomised Controlled Trials be Genuinely Realist?"
- Williams, L. J., R. J. Vandenberg, J. R. Edwards. 2009. "12 Structural Equation Modeling in Management Research: A Guide for Improved Analysis." *The Academy of Management Annals* 3 (1): 543–604.
- Wolff, Nancy. 2001. "Randomised Trials of Socially Complex Interventions: Promise or Peril?" *Journal of Health Services Research & Policy*, 123–126. doi:[10.1258/1355819011927224](https://doi.org/10.1258/1355819011927224)