Nurse Education Today 162 (2026) 107075

Contents lists available at ScienceDirect

Nurse Education Today

journal homepage: www.elsevier.com/locate/nedt

Beyond "what works": Realist evaluation as a transformative approach to building evidence in nurse education☆

Gemma Stacey *

School of Social Sciences, Nottingham Trent University, 50 Shakespeare Street, Nottingham, NG1 4FQ, United Kingdom

ARTICLE INFO

Keywords:
Realist evaluation
Critical realism
Context Mechanism Outcome
Programme theory
Nurse education research
Evidence based education

ABSTRACT

The persistent "what works" question in nurse education research conceals a conceptual trap. We seek universal interventions that reliably produce outcomes regardless of context, building an evidence base where promising interventions repeatedly fail to transfer across settings. Critical realism and its methodological offspring, realist evaluation, offers a transformative alternative from the social sciences. Rather than asking "does this work," realist evaluation asks: for whom, in what circumstances, through what mechanisms, producing which outcomes? This Big Ideas paper introduces the Context Mechanism Outcome (CMO) configuration framework as analytical architecture for understanding how educational interventions operate. Using Resilience Based Clinical Supervision as an exemplar, I demonstrate how realist approaches generate middle range theories - conditional, contextualised knowledge about what tends to work for whom in what circumstances. Nurse education's characteristic features (diverse student populations, constitutive context, complex interventions, multiple stakeholders) make realist approaches essential rather than optional. The transformation required is epistemological: abandoning false universalism to build cumulative knowledge about how mechanisms operate under varying conditions. This produces an evidence base adequate to nursing education's complexity, offering practitioners sophisticated conceptual resources for contextually intelligent implementation rather than simplistic prescriptions.

1. The conceptual trap

The question "what works?" has driven decades of research seeking transferable "best practices" (Pawson, 2013). Yet this deceptively simple question embodies a problematic assumption: that interventions work universally, independent of context. We have built an evidence base reflecting this preoccupation, and its limitations are increasingly impossible to ignore.

Consider the familiar pattern: simulation based learning shows promise in one institution, generates enthusiasm and publications, then disappoints when scaled or transferred to different settings. Problem based learning transforms student engagement in one curriculum but produces resistance and confusion in another. Reflective practice interventions create meaningful professional development in one workplace but feel like bureaucratic box ticking in another. We typically blame "implementation failure" or "poor fidelity" (Greenhalgh et al., 2017), but rarely interrogate whether our fundamental approach to building evidence is fit for purpose.

Nurse education, with its contextual complexity, diverse student populations, varied clinical environments, and multiple interacting systems, demands a more sophisticated epistemological foundation. Our students range from school leavers to experienced healthcare assistants returning to study. Our practice placements span university teaching hospitals, small rural units, care homes, and community settings. Our educational interventions operate within curriculum frameworks, regulatory requirements, resource constraints, and workplace cultures that vary enormously. Yet we continue seeking evidence of "what works" as if these complexities could be controlled away.

Critical realism, developed primarily through Roy Bhaskar's philosophy (Bhaskar, 2008) and operationalised by Ray Pawson and Nick Tilley's realist evaluation methodology (Pawson and Tilley, 1997), offers precisely the epistemological foundation we need. Yet despite its uptake in public health and social policy research, realist approaches remain underutilised in nurse education, a missed opportunity to build evidence that is simultaneously more rigorous and more practically useful.

☆ This article is part of a Special issue entitled: 'Novel Methodologies in Nurse Education Research' published in Nurse Education Today.

* Corresponding author.
E-mail address: gemma.stacey@ntu.ac.uk.

https://doi.org/10.1016/j.nedt.2026.107075
Received 29 January 2026; Received in revised form 16 February 2026; Accepted 9 March 2026
Available online 16 March 2026

2. Critical realism: a philosophical foundation

Critical realism rejects both naïve realism (the world is exactly as we observe it) and radical constructivism (there is no reality beyond interpretation). Instead, it posits a stratified ontology: the real (underlying mechanisms and structures), the actual (events that occur when mechanisms activate), and the empirical (what we observe and measure) (Bhaskar, 2008). This philosophical position has particular relevance for healthcare education research (Clark et al., 2008).

For nurse education, this means educational processes involve real causal mechanisms – cognitive, social, emotional, relational – but these mechanisms operate contingently, activated or suppressed by contextual conditions. A reflective practice intervention does not "work" or "not work" in some absolute sense. Rather, it offers resources that may trigger particular mechanisms (cognitive reappraisal of challenging experiences, collective sense making that resists individualised blame, professional identity development, emotional processing) in practitioners who possess certain characteristics (readiness to engage in reflection, psychological safety in relationships, workload capacity for protected time, prior experience with reflective approaches) within specific contexts (team culture that values learning, organisational support for protected time, facilitation quality, alignment with assessment requirements).

Pawson and Tilley's (1997) central insight is that programmes are theories incarnate, every educational intervention embodies assumptions about how change happens, even when these assumptions remain implicit. The evaluator's task is not determining whether the programme "works" in some binary sense, but testing and refining the programme theory: for whom does this work, in what circumstances, through what mechanisms, to produce which outcomes?

3. CMO configurations: new grammar for evidence

Context Mechanism Outcome configurations provide realist evaluation's analytical architecture (Pawson and Tilley, 1997). Unlike conventional evaluation frameworks that treat context as noise to be controlled or a barrier to overcome (Fig. 1), CMO configurations position context as fundamental to causal explanation (Dalkin et al., 2015).

Consider problem based learning: traditional approaches might compare PBL versus lecture based learning on exam scores, attempting to control for confounding variables to establish whether PBL "works." Realist evaluation asks fundamentally different questions: for which students (those with particular learning orientations, levels of prior knowledge, self regulation capacities, comfort with ambiguity), in what circumstances (small group composition and dynamics, tutor facilitation approaches, curriculum structure, assessment alignment, resource availability), does PBL trigger what mechanisms (collaborative knowledge construction, metacognitive development, intrinsic motivation, peer accountability), leading to which outcomes (deep learning, clinical reasoning ability, professional identity formation, retention of knowledge)?

This generates portable, nuanced knowledge. Rather than claiming "PBL works" or "PBL does not work" in some universal sense, we develop what Merton (1968) termed middle range theories: PBL appears to support deep learning for students with moderate prior knowledge when groups achieve psychological safety and tutors facilitate without dominating, by triggering collaborative sense making that integrates conceptual and experiential learning. When prior knowledge is very limited, students may lack the foundation for productive peer learning. When tutors dominate discussions, the collaborative mechanism is suppressed. When assessment rewards memorisation rather than application, the intrinsic motivation mechanism weakens.

This form of evidence – conditional, contextualised, mechanism

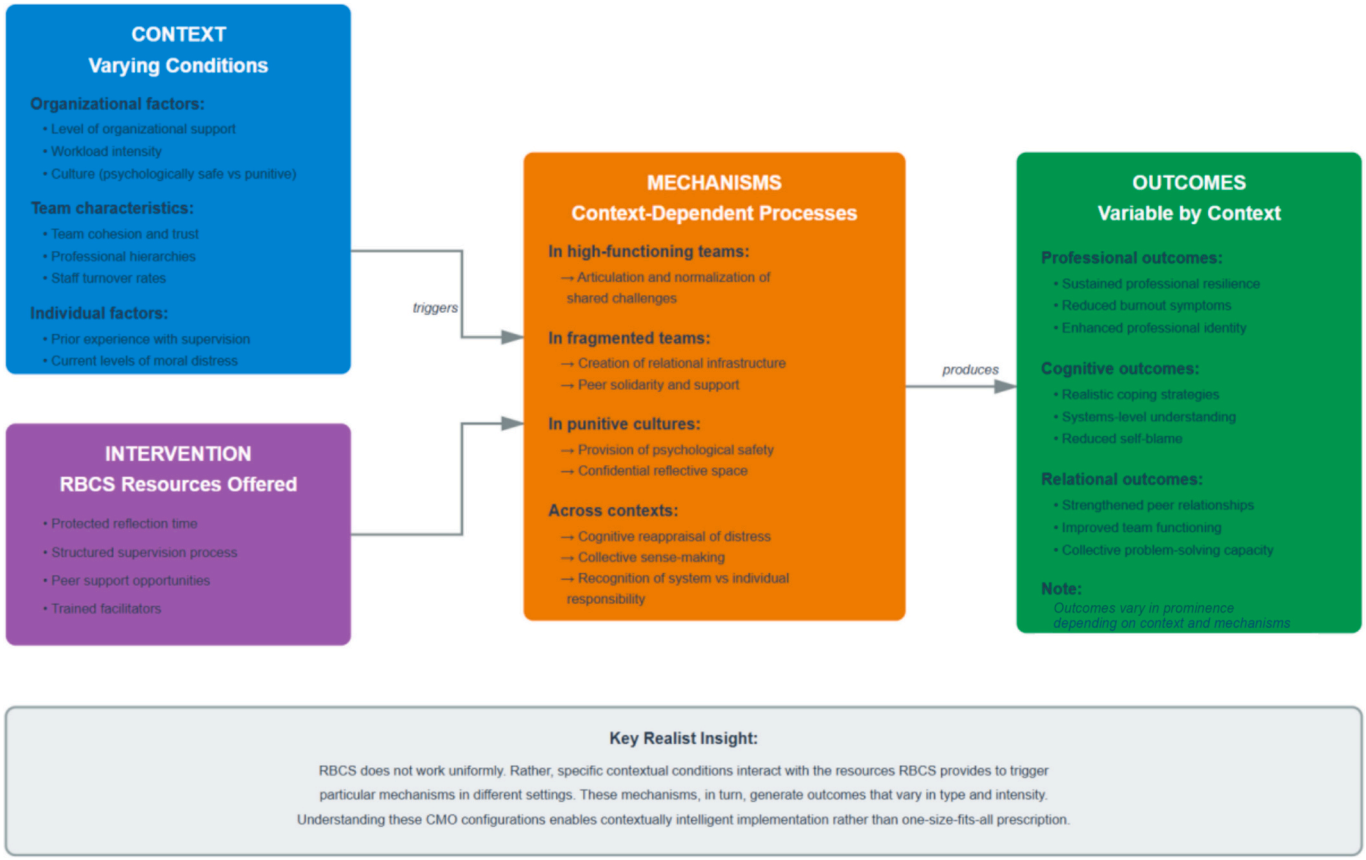

Fig. 1. CMO configuration for resilience based clinical supervision.

focused – reflects what practitioners need for intelligent implementation. It explains not just whether something worked in one setting, but why it worked and under what conditions similar mechanisms might be activated elsewhere.

4. Why nurse education needs realist approaches

Nursing education presents particular features that make realist approaches essential rather than just useful. First, our student population is irreducibly diverse: traditional school leavers entering higher education for the first time, mature learners returning to education after years away, international students navigating linguistic and cultural transitions, nursing degree apprentices balancing full time employment with academic study. Any intervention encounters this diversity, and mechanisms fire differently depending on what students bring to the learning situation.

Second, context is not peripheral but constitutive in nurse education. What succeeds in a well resourced university teaching hospital with established education cultures may fail completely in a small rural placement with limited supervision capacity. A reflective practice intervention implemented in a psychologically safe team with supportive leadership operates through entirely different mechanisms than the same intervention implemented in a toxic workplace culture characterised by blame and defensiveness. We cannot, and should not, attempt to control this variation away, it is fundamental to understanding how education works.

Third, nurse education involves complex interventions with multiple interacting components that rarely operate in isolation. Curriculum redesign intersects with practice assessment changes, which interact with digital learning technologies, which are shaped by regulatory requirements. Realist evaluation provides frameworks for understanding how components interact and which combinations are necessary or sufficient for particular outcomes (Wong et al., 2017).

Finally, we have multiple stakeholders with legitimately different outcome priorities. Students care about manageable workload and employability. Practice supervisors prioritise patient safety and workforce readiness. Academic staff value intellectual development and critical thinking. Service users want compassionate, competent care. Employers need staff who can function effectively from day one. Regulators focus on public protection and professional standards. CMO configurations can accommodate this multiplicity, asking: for which stakeholders, under what conditions, does this intervention produce valued outcomes?

5. Exemplar: resilience based clinical supervision

My experience developing and evaluating Resilience Based Clinical Supervision (RBCS), which has now been implemented across varied healthcare contexts internationally (Stacey et al., 2020), exemplifies how realist thinking illuminates programme theory and guides adaptive implementation.

Conventional evaluation would ask a simple question: does RBCS improve retention rates for newly qualified nurses? Realist evaluation enables far more sophisticated interrogation of how, why, for whom, and under what conditions the intervention works.

The programme theory makes explicit what RBCS offers and how it might work: RBCS provides resources including protected reflection time, structured supervision process, peer support through group sessions, and trained facilitators skilled in managing psychological safety. These resources are implemented within contexts characterised by varying levels of organisational commitment and support, team cohesion and pre-existing relationships, workload intensity and staffing pressures, and professional hierarchies and power dynamics.

When these resources meet receptive contexts and responsive participants, they may trigger particular mechanisms: cognitive reappraisal that helps practitioners make sense of distressing experiences without personalising system level failures; collective sense making that resists narratives of individual inadequacy by recognising shared challenges; explicit recognition of system level constraints versus personal performance issues; development of contextually appropriate coping strategies that acknowledge workplace realities.

Crucially, RBCS operates through different mechanisms in different contexts, and this matters for implementation. In high functioning teams with pre-existing trust and psychological safety, RBCS primarily enables articulation and normalisation of shared challenges - the confidential group space allows people to voice difficulties they already suspected were common. In fragmented teams where staff barely know each other, the peer support mechanism becomes primary - RBCS actively creates relational infrastructure that did not previously exist, with trust developing through the regular sessions. In organisations with punitive cultures characterised by blame and defensiveness, the confidential space itself becomes the critical mechanism, offering psychological safety that is notably absent in everyday work.

This understanding, derived from realist evaluation across multiple implementations, generates genuinely actionable knowledge. Rather than simplistic recommendations to "implement RBCS" with fidelity to a fixed model, we can provide contextually intelligent guidance: RBCS requires sufficient organisational commitment to protect time consistently, as irregular sessions undermine the trust building mechanism. Facilitators need training not just in process facilitation but specifically in creating and maintaining psychological safety, particularly in challenging group dynamics. In high turnover contexts, attention to group continuity becomes critical because the peer support mechanism depends fundamentally on trust that develops only over time.

This exemplifies how realist evaluation moves beyond traditional implementation fidelity concerns toward adaptive, context sensitive application that maintains programme integrity through explicit attention to mechanism activation rather than rigid adherence to protocol.

6. Toward a more applied evidence base

Realist evaluation demands explicit theorising about how change happens, systematic data collection across multiple sources, transparent reasoning about causal pathways, and honest acknowledgment of scope conditions and limitations (Pawson, 2013). It rejects false universalism in favour of conditional, contextualised knowledge that practitioners can apply intelligently to their own situations.

Without realist approaches, we perpetuate an evidence base that is simultaneously too weak and too rigid. Too weak because findings do not transfer across contexts, we have systematically stripped away the very situational mechanisms that enabled interventions to succeed in their original settings. Too rigid because we keep seeking universal solutions to problems that are inherently contextual, then expressing frustration when "evidence based" interventions fail in practice (Greenhalgh et al., 2017).

Realist methodology offers middle range theories (Merton, 1968), transferable insights about what tends to work, for whom, in what circumstances, through what processes. This is precisely what educators need: not simplistic prescriptions claiming universal applicability, but sophisticated conceptual resources for contextually intelligent practice.

Conducting realist evaluation requires methodological eclecticism guided by theoretical clarity (Pawson and Tilley, 1997). It begins with articulating programme theory, making explicit the often implicit assumptions about how interventions generate change. This involves engaging stakeholders, reviewing literature through a realist lens, and developing initial CMO configurations as testable hypotheses.

Data collection is iterative and explicitly theory testing, using multiple sources to illuminate different aspects of CMO configurations (Rycroft Malone et al., 2012). Qualitative data from interviews and observations reveals mechanism triggering and the contextual conditions that activate or suppress mechanisms. Quantitative data maps outcome patterns and identifies which configurations appear more

prevalent. Comparative case studies examine how varying contexts produce different mechanism outcome pathways. Longitudinal approaches trace how CMOs evolve as contexts shift or participants change.

Analysis focuses on configurational causation: which combinations of context and mechanism produce which outcomes? This requires moving beyond traditional variable centred analysis toward case based comparison, asking not "does X correlate with Y?" but "under what circumstances does X trigger mechanisms leading to Y?" (Dalkin et al., 2015).

Rigour differs from but equals other approaches, following RAMESES quality standards (Wong et al., 2017): explicit programme theories articulated as testable propositions; systematic and transparent data collection; clear reasoning chains linking evidence to theoretical claims; acknowledgment of alternative explanations and scope conditions; iterative theory refinement through empirical testing.

7. The epistemological shift

The transformation required is epistemological as much as methodological. We must abandon the seductive but ultimately unhelpful premise of context free best practices and embrace the more challenging work of building cumulative knowledge about how educational mechanisms operate under varying conditions.

Realist evaluation can enhance rather than replace existing research approaches. It complements experimental designs by explaining how and why effects occur or fail to occur in different contexts, directly addressing the persistent "black box" problem in evaluation research. It enriches implementation science by explicitly theorising the mechanisms through which implementation strategies themselves work (Greenhalgh et al., 2017). It provides coherent frameworks for mixed methods integration, where different data types illuminate different aspects of CMO configurations (Rycroft Malone et al., 2012). It transforms evidence synthesis by treating heterogeneity of findings as analytically valuable rather than problematic.

For nurse education researchers considering realist approaches, integration might proceed incrementally. Begin by articulating programme theories explicitly even when using conventional evaluation designs. Add targeted qualitative components that explore mechanism activation alongside outcome measurement. Analyse existing data through realist lenses, examining outcome variation as evidence of context mechanism interactions rather than merely noise to control. Report findings in CMO format alongside traditional results, making conditional claims about what works for whom in what circumstances. Contribute to building cumulative knowledge by explicitly positioning findings within emerging middle range theories rather than presenting them as isolated effects.

The fundamental question shifts from "what works?" to "what works, for whom, in what circumstances, why, and how?" Answer that question rigorously and systematically, and we build an evidence base worthy of nursing education's complexity- evidence that respects both methodological rigour and contextual reality, offering practitioners genuine guidance for intelligent implementation rather than false promises of universal solutions that inevitably disappoint.

CRediT authorship contribution statement

Gemma Stacey: Writing – review & editing, Writing – original draft, Conceptualization.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author used Claude (Anthropic) to support structuring arguments and refining language for clarity. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the published article.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

Professor Gemma Stacey is Editor-in-Chief of Nurse Education Today. Given this journal affiliation, Professor Stacey has had no involvement in the peer-review of this article and has no access to information regarding its peer-review. Full responsibility for the editorial process for this article has been delegated to another journal editor to ensure the integrity and independence of the peer review process.

The author declares no other competing interests, financial or otherwise, that could inappropriately influence or bias this work. The paper has not been submitted elsewhere.

References

Bhaskar, R., 2008. A Realist Theory of Science, 3rd ed. Verso, London.
Clark, A.M., Lissel, S.L., Davis, C., 2008. Complex critical realism: tenets and application in nursing research. Adv. Nurs. Sci. 31 (4), E67–E79.
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. Implement. Sci. 10, 49.
Greenhalgh, T., Wherton, J., Papoutsi, C., Lynch, J., Hughes, G., A'Court, C., Hinder, S., Fahy, N., Procter, R., Shaw, S., 2017. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale up, spread, and sustainability of health and care technologies. J. Med. Internet Res. 19 (11), e367.
Merton, R.K., 1968. Social Theory and Social Structure, Enlarged ed. Free Press, New York.
Pawson, R., 2013. The Science of Evaluation: A Realist Manifesto. Sage Publications, London.
Pawson, R., Tilley, N., 1997. Realistic Evaluation. Sage Publications, London.
Rycroft Malone, J., McCormack, B., Hutchinson, A.M., DeCorby, K., Bucknall, T.K., Kent, B., Schultz, A., Snelgrove Clarke, E., Stetler, C.B., Titler, M., Wallin, L., Wilson, V., 2012. Realist synthesis: illustrating the method for implementation research. Implement. Sci. 7, 33.
Stacey, G., Cook, G., Aubeeluck, A., Stranks, B., Long, L., Krepa, M., Lucre, K., 2020. The implementation of resilience based clinical supervision to support transition to practice in newly qualified healthcare professionals. Nurse Educ. Today 94, 104564.
Wong, G., Westhorp, G., Manzano, A., Greenhalgh, J., Jagosh, J., Greenhalgh, T., 2017. RAMESES II reporting standards for realist evaluations. BMC Med. 15, 96.