



Theory-based evaluation and programme theories in nursing: A discussion on the occasion of the updated Medical Research Council (MRC) Framework

Martin Wallner^{a,b,*}, Hanna Mayer^a, Laura Adlbrecht^c, Anna Louisa Hoffmann^{d,e}, Anne Fahsold^{d,e},
Bernhard Holle^{d,e}, Adelheid Zeller^c, Rebecca Palm^d

^a Karl Landsteiner University of Health Sciences, Division of Nursing Science with focus on Person-Centred Care Research, Krems, Austria

^b University of Vienna, Vienna Doctoral School of Social Sciences, Department of Nursing Science, Vienna, Austria

^c OST Eastern Switzerland University of Applied Sciences, Department of Health, Competence Center Dementia Care, St. Gallen, Switzerland

^d Witten/Herdecke University, Faculty of Health, School of Nursing Science, Witten, Germany

^e Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE), site Witten, Witten, Germany

ARTICLE INFO

Article history:

Received 6 October 2022

Received in revised form 13 January 2023

Accepted 26 January 2023

Keywords:

Complex interventions

Discussion

Methodology

Nursing

Program theory

Realist evaluation

Theory-based evaluation

ABSTRACT

Developing and evaluating health interventions for the benefit of patients is notoriously difficult. This also applies to the discipline of nursing, owing to the complexity of nursing interventions. Following significant revision, the updated guidance of the Medical Research Council (MRC) adopts a pluralistic view to intervention development and evaluation, including a theory-based perspective. This perspective promotes the use of program theory, aiming to understand how and under what circumstances interventions lead to change.

In this discussion paper, we reflect the recommended use of program theory in the context of evaluation studies addressing complex nursing interventions. First, we review the literature by investigating the question whether and how evaluation studies targeting complex interventions used theory and to what extent program theories may contribute to enhance the theoretical foundations of intervention studies in nursing. Second, we illustrate the nature of theory-based evaluation and program theories. Third, we argue how this may impact theory building in nursing in general. We finish by discussing which resources, skills and competencies are necessary to fulfill the demanding task of undertaking theory-based evaluations.

We caution against an oversimplified interpretation of the updated MRC guidance regarding the theory-based perspective, e.g. by using simple linear logic models, rather than articulating program theories. Instead, we encourage researchers to embrace the corresponding methodology, i.e. theory-based evaluation.

With the prevailing perspective of knowledge production in crisis, we might be on the verge of a paradigm shift in health intervention research. Viewed through this lens, the updated MRC guidance could lead to a renewed understanding of what constitutes useful knowledge in nursing. This may facilitate knowledge production and, thereby, contribute to improve nursing practice for the benefit of the patient.

Tweetable abstract: The latest iteration of the MRC Framework for developing and evaluating complex healthcare interventions could lead to a renewed understanding of what constitutes useful knowledge for nursing.

© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

What is already known

- The Medical Research Council (MRC) Framework is the state of the art for developing and evaluating complex interventions.
- The latest iteration of the framework adopts a pluralistic view on research perspectives, including a theory-based perspective advocating

the use of program theories in intervention development and effectiveness research.

What this paper adds

- A rapid review on intervention studies published in the International Journal of Nursing Studies indicates that the effectiveness/efficacy perspective prevails and the use of program theory is scarce.
- An exploration about program theories, their philosophical foundations and the value of the realist methodology compared to other methodologies.

* Corresponding author at: Karl Landsteiner University of Health Sciences, Division of Nursing Science with focus on Person-Centred Care Research, Dr.-Karl-Dorrek-Strasse 30, 3500 Krems, Austria.

E-mail address: martin.wallner@kl.ac.at (M. Wallner).

- A discussion on how the utilization of theory-based evaluations and program theorizing may reinvigorate theory development in nursing.

1. Introduction

Intervention effectiveness research plays an essential role in knowledge production for an evidence-based and safe nursing practice (Sidani and Braden, 2021). Regarding the selection of adequate methodologies to examine the effectiveness of nursing interventions, researchers face major difficulties (Balzer et al., 2012; Dichter, 2022; Höhmann and Bartholomeyczik, 2013; Moers et al., 2011). This is in large part due to the widely agreed fact that nursing interventions are complex interventions (Corry et al., 2013; Höhmann and Bartholomeyczik, 2013). They consist of multiple interacting components, thereby addressing a range of behavior changes and involving several target groups as well as settings (Craig et al., 2008).

In recent years, the development and testing of complex interventions has been the subject of extensive discussion in nursing research (Möhler et al., 2015; Paterson et al., 2009). Multiple methodological guidelines are available with descriptions of different phases ranging from intervention development to testing and evaluation (Corry et al., 2013; Sidani and Braden, 2021). Grounding the intervention in a theoretical framework is considered key in the development phase of nursing interventions (Corry et al., 2013). In this regard, several ways of theorizing the intervention have been proposed, including theory-based evaluations (De Silva et al., 2014; Fletcher et al., 2016; Höhmann and Bartholomeyczik, 2013).

The Medical Research Council (MRC) Framework is considered as the state of the art for developing and evaluating complex interventions in nursing. It guides researchers in designing research projects and funders in deciding acceptance or rejection of research proposals. For universities, the MRC Framework provides orientation for compiling their curricula of research methods. The first publication in 2000 introduced different phases characterizing the development and evaluation of complex interventions: 'Theory' (pre-clinical), 'Modelling' (phase I), 'Exploratory trial' (phase II), 'Definite RCT' (phase III), and 'Long term implementation' (phase IV) (Campbell et al., 2000). A revision in 2008 resulted in the description of four phases: 'Development', 'Feasibility/Piloting', 'Evaluation', and 'Implementation' (Craig et al., 2008). The 2008 version clarified that the phases are not linear but interacting with each other. The latest iteration (Skivington et al., 2021) addressed two key aspects of the 2008 version that were in need of rethinking:

1. The definition of a complex intervention taking into account a broader understanding of complex contexts in which interventions are delivered and evaluated.
2. The design of evaluation studies intending not only to minimize bias but also considering how to maximize the usefulness of evidence for decision making.

Skivington et al. (2021) advocate to widen the perspective of effectiveness/efficacy that was characteristic for most complex health intervention research so far. They recommend to adopt a theory-based perspective aiming to understand how and under what circumstances interventions lead to a change. Furthermore, Skivington et al. (2021) suggest to complement the four phases (as one of the six core elements of the new MRC Framework) by developing, refining and (re)testing a program theory.

In this discussion paper, we aim to reflect the recommended use of a program theory in the context of evaluation studies addressing complex nursing interventions. First, we review the literature by investigating the question whether and how evaluation studies targeting complex interventions used theory and to what extent program theories may contribute to enhance the theoretical foundations of intervention studies in nursing. Second, we illustrate the nature of theory-based evaluation and

program theories. Third, we argue how this may impact theory building in nursing in general. We finish by discussing which resources, skills and competencies are necessary to fulfill the demanding task of undertaking theory-based evaluations.

2. The use of program theories in nursing intervention research

We screened studies addressing complex interventions to find out which research perspectives were apparent and whether a theory-based perspective was among them. As a source of high-impact studies we chose the *International Journal of Nursing Studies* (IJNS) since it has the highest impact factor (6.612 in 2021) in nursing (Thomson Reuters, 2022) (Journal Citation Reports). In January and December 2022, we performed a systematic search in the Medline database for intervention studies published in IJNS during the last five years (search syntax can be obtained from the authors). We identified 299 articles in the initial search (01/2022) and 72 in the update (12/2022) and screened their full text to find out which of them followed the MRC Framework. In our opinion, seven studies applied the MRC Framework as a guidance for their studies (Abraham et al., 2019; Duxbury et al., 2019; Jiang et al., 2021; Lin et al., 2023; Rahn et al., 2018; van Duinen-van den Ijssel et al., 2019; Xiao et al., 2022). These studies evaluated complex nursing interventions consisting of multiple components. The interventions aimed at changing individual behavior, mood or well-being as well as practice, culture or policy by means of education, therapy, training, decision coaching, information repositories (e.g. Wikis, guidelines), structured support, policy statements, structured assessments, and evaluations. The interventions outlined in these studies were based on a broad empirical basis encompassing knowledge about the prevalence of the problem to be solved, influencing factors, qualitative evidence documenting the view of patients, staff or relevant others as well as evidence on the effectiveness of interventions. This empirical basis was built on several studies applying different methodologies. Researchers of all included studies have comprehensively described how often and why the respective problem occurs and what may help to reduce its occurrence. The interventions were carefully developed, took into consideration different perspectives and were evaluated mainly with regard to feasibility and practicability. In four studies (Duxbury et al., 2019; Lin et al., 2023; van Duinen-van den Ijssel et al., 2019; Xiao et al., 2022), the intervention had been tested previously in a different context (e.g. country or target group). Four research teams (Jiang et al., 2021; Lin et al., 2023; Rahn et al., 2018; Xiao et al., 2022) described the theoretical basis of their interventions (the train-the-trainer principle based on the teaching approach, the six steps of shared decision-making, the situation-specific theory of heart failure, the Progressively Lower Stress Threshold model and a combined model of the eighth stage of Erikson's model of psychosocial development and the dignity model). They designate their interventions or programs as theory-based but not their evaluation approach. With regard to prevailing research perspectives, we concluded that the effectiveness/efficacy perspective is certainly the strongest in the papers that we reviewed. One exception was the study from Lin et al. (2023) who performed a qualitative inquiry in addition to a randomized controlled trial to explore how and why the treatment worked from the participants' perspective.

While effectiveness/efficacy was the dominant perspective, all included studies used process evaluations that provided answers to many questions reaching beyond the dichotomous question of effectiveness. Abraham et al. (2019) as well as van Duinen-van den Ijssel et al. (2019) acknowledged that they probably missed active components and their impact on the intervention in the new context. Abraham et al. (2019) discussed that context conditions have a strong influence on the intervention and examined center variations. These conditions were not known before the study and may have played an important role (Abraham et al., 2019). But neither did the subgroup-analysis explain why the intervention worked in some centers but not

in others, nor did the process evaluation reveal which active ingredient was needed to make the intervention work in every nursing home. From the process evaluation, Abraham et al. (2021) concluded that different nursing homes need different approaches to change the practice but that they did not identify what predicts the responsiveness of nursing homes. Xiao et al. (2022) point out that their process evaluation revealed difficulties in applying an intervention that was originally developed in a different context and that high attrition rates may reduce the reliability of the results.

None of the studies, however, developed or presented a program theory in order to explain how the active components of the intervention were expected to work and which role the context played. Developing program theory might help to address such difficulties beforehand.

Following the updated MRC Framework, the explanatory content and usefulness of intervention studies will increase when reductionist approaches to address complexity are abandoned. The updated MRC recommendations support research adequately accounting for, rather than controlling for the complexity of the research object. Thus, the focus is not only on the intervention logic (how an intervention works), but on how the system might affect the intervention and vice versa. Therefore, it seems insufficient to focus on multiple empirical sources as the basis for developing interventions and conducting a pilot phase, followed by a traditional effectiveness design (e.g., RCT) and supplemented by an accompanying process evaluation (focusing on feasibility issues). If the aim is to 'let the complexity genie out of the bottle' (Pawson, 2022, paragraph 5), it is necessary to view program theorizing not just as an additional step or component of intervention research, but to consider and, indeed, embrace the corresponding methodology (i.e., theory-based evaluation) as a whole. If one considers program theorizing as one additional formal step that has to be done, there is the danger of stepping again into a 'reductionist trap'.

3. Theory-based evaluation and the nature of program theories

Theory-based evaluation relies on a program theory explaining how an intervention is expected to generate results. A program theory is used to design the evaluation, to implement the intervention, and/or to interpret the results (Coryn et al., 2011). Program theory describes how an intervention (e.g., a project or program) contributes to a chain of intermediate results, and finally to the intended outcomes (Funnell and Rogers, 2011). By doing so, evaluators claim to open the 'black box' of interventions' relationship between cause and effect (Funnell and Rogers, 2011).

Theory-based evaluation is considered an umbrella term covering a variety of methodologies and concepts. Funnell and Rogers (2011) listed over 20 terms associated with theory-based evaluation, among others theory-driven evaluation, theory-guided evaluation, realistic evaluation/realist evaluation, program theory, theory of action, theory of change, intervention theory, and logic model. The heterogeneity of terms is criticized in methodological discourses due to inconsistent application and lack of distinctness (Funnell and Rogers, 2011; von Werthern, 2020).

Theory-based evaluations emerged as a response to criticism and shortcomings of traditional evaluation methodologies following a positivist understanding of science and preferring (quasi-) experimental research designs (von Werthern, 2020). This traditional type of evaluation provides not enough information enabling stakeholders to better understand and to improve a program (Chen, 2015). Contrary to traditional 'black box' evaluations (Astbury and Leeuw, 2010), the intention of theory-based evaluations is not only to understand *whether* a program works, but also *how* it works to mitigate a social problem (Funnell and Rogers, 2011; Höhmann and Bartholomeyczik, 2013). The common ontological, epistemological and methodological foundations of theory-based evaluations are rooted in critical realism (Brouselle and Buregeya, 2018). Critical realism (analog to realism) assumes that outside of human consciousness there is a real world

accessible to human perception. However, according to critical realism, the knowledge of this world remains uncertain (Döring et al., 2016). Critical realism is based on a stratified ontology, wherein reality is divided into the domains of the real, the actual, and the empirical. Underlying powers and structures in the real domain cause events in the actual domain. These events are known but may not always be observable. The empirical domain refers to human perceptions and interpretations of observable events (Bhaskar, 2013; Clark et al., 2008; Walsh and Evans, 2014).

von Werthern (2020) described five characteristics shared by different approaches to theory-based evaluation: First, the development of a program theory is intrinsic to a theory-based evaluation and reaches beyond a logic model. Second, a program theory is embedded in the methodological/conceptual framework of the evaluation. It guides the epistemological interest, the formulation of research questions, the methodology, and the choice of methods for data collection and analysis. Third, theory-based evaluation can serve the purpose of a final evaluation or of a continuous development process to further understand and improve the program. Fourth, theory-based evaluation is considered pragmatic, methodologically neutral, and methodologically pluralist. The choice of methods should follow the epistemological interest. Thus, theory-based evaluation cannot be assigned to either the experimental or the constructivist research paradigm. It is placed on a continuum between the two poles. Fifth, the utility of theory-based evaluations extends beyond the use of evaluation results (von Werthern, 2020).

Theory-based evaluation has its main roots in public health and sociology publications of the second half of the last century. After a first boom in the 1980s/90s, it arrived in the mainstream of evaluation in the 2000s (Funnell and Rogers, 2011; von Werthern, 2020). Over the past twenty years, theory-based evaluation has undergone continuous development in practice and theory (Astbury and Leeuw, 2010; Chen, 2015; Coryn et al., 2011; Funnell and Rogers, 2011; Giel, 2013; Haubrich, 2009; Stame, 2017; Vaessen and Leeuw, 2017; von Werthern, 2020). Currently, it is increasingly discussed and applied in health care research (Blamey et al., 2013; De Silva et al., 2014; Fromer, 2017; Gentile and Marzinski, 2020; Palm and Hochmuth, 2020; Renmans et al., 2017; Skivington et al., 2021).

In the following, two approaches to theory-based evaluation are explained in more detail: theory-driven evaluation (Chen, 2015) and realistic evaluation (Pawson and Tilley, 1997). Theory-driven evaluation is well-known, widely discussed and generally used in evaluation research. Realistic evaluation has been increasingly discussed and applied in health care and nursing science in recent years (Palm and Hochmuth, 2020). A non-sensitive, rapid overview search for program theory in nursing in PubMed (December 2022) resulted in 38 realistic evaluations, 56 realistic reviews or syntheses, one theory-driven evaluation, seven publications using other approaches to theory-based evaluation and eight publications on program theories without reference to theory-based evaluation. A reference list of the studies can be found in the Supplementary material.

Chen developed the theory-driven approach to evaluation in 1990. Its centerpiece is a two-part program theory including a prescriptive theory (theory of action), and a causative theory (theory of change). Their linkage is operationalized in the 'Action and Change Model' schema. The 'Action Model' contains prescriptive components considered as necessary for program success. The 'Change Model' includes the descriptive assumptions underlying the program's mechanisms of impact (Chen, 2015). To this day, the two-part program theory is widely used also in the context of other approaches to theory-based evaluation (Funnell and Rogers, 2011). In addition, Chen further developed the typology of theory-driven evaluations and designed a matrix of process and outcome evaluations (assessing processes, such as implementation, vs. assessing the impact and effect of a program) as well as constructive (information to improve the program) and concluding evaluations (information to assess the value of the program) (Chen, 2015).

Realistic evaluation – nowadays usually referred to as ‘realist evaluation’ – was developed by Pawson and Tilley and first published in 1997 (Pawson and Tilley, 1997). In contrast to other approaches to theory-based evaluation, they explicitly position themselves in ‘scientific realism’ (von Werthern, 2020). Pawson and Tilley call for evaluations intending to identify why a program works or does not work. They emphasize the importance of considering the influence of the context on interventions’ effect. Therefore, the central realist question is: What works for whom in what circumstances? (Pawson and Tilley, 1997, p. 125). This question is answered in context–mechanism–outcome configurations. A wide variety of methods serves to develop these configurations. The approach is intended as a heuristic meta-framework rather than a methodological instruction for action (von Werthern, 2020).

4. Constructing program theories

Assuming a realist view of causation based on a stratified, generative ontology (Clark et al., 2008), theory-based evaluations aim to identify the generative mechanisms of a program. The effects produced by these mechanisms can be observed empirically and are articulated/elaborated in a program theory (von Werthern, 2020). Program theory is the most important constituent feature of theory-driven evaluations (Coryn et al., 2011; von Werthern, 2020). It commonly consists of two components: a theory of change and a theory of action (Funnell and Rogers, 2011). As a program theory stays close to the research object – the program – it should be distinguished from scientific theory (formal theory), such as social or psychological theory. However, it can be informed by scientific theories (Coryn et al., 2011; De Silva et al., 2014; Giel, 2013; Haubrich, 2009; Reith and Kelle, 2016).

Program theories’ level of abstraction is considered as middle or low (Burns and Grove, 2011; Shearn et al., 2017), because they usually articulate a limited number of concepts pertaining to a specific situation in a specific context lending itself to empirical investigation (Meleis, 2018). Program theories are commonly represented by using visual aids, such as logic models, including, for instance, pipeline logic models, outcomes chain logic models, or realist matrices contributing to communicate program ideas among stakeholders (Funnell and Rogers, 2011). However, it is important to note that program theories should be described thoroughly in written form, since most logic models are reductionist and thus unlikely to articulate a program theory in a sufficient manner (Skivington et al., 2021). This aspect is particularly important for program theories following the realist methodology and explicating context–mechanism–outcome configurations (Pawson and Tilley, 1997).

Various data sources are commonly used to develop program theories, including, for instance, a review of program documents, a review of the scientific literature, interviews with stakeholders, workshops or interviews with staff performing the intervention (Dössegger et al., 2017; Funnell and Rogers, 2011). Although there is no one best way to elaborate a program theory (Funnell and Rogers, 2011), workshops and interviews are among the most common methods for explicating stakeholders’ assumptions about how a program works (Dössegger et al., 2017; Funnell and Rogers, 2011). The literature offers some guidance on collecting relevant data for the construction of program theories and logic models, such as semi-structured interview guides (Funnell and Rogers, 2011; Gugiu and Rodríguez-Campos, 2007).

While qualitative methods are recommended to collect data on stakeholders’ assumptions, the literature rarely provides methodological recommendations concerning theory construction (von Werthern, 2020). A recently published model indicates how to proceed when constructing a program theory (von Werthern, 2020). It offers some guidance and emphasizes the concept of abductive reasoning as a ‘missing link’ between deduction and induction for developing program theories (von Werthern, 2020). The model proposes to facilitate abductive reasoning using, for example, thought experiments, such as the ‘why heuristic’ (Jaccard and Jacoby, 2020). It contributes to identifying program mechanisms and relationships among program components.

5. Theory-based evaluations and program theories in nursing research and theory building

We would like to emphasize that the updated MRC Framework does not reinvent the wheel of theory-based evaluation and program theories. In nursing, program theories were considered early on as intervention theories (Burns and Grove, 2011; Sidani and Braden, 1998, 2021). They were conceived as an alternative paradigm for evaluating an intervention in clinical research traditionally using experimental research designs (McEvoy and Richards, 2003; Sidani and Braden, 1998; Sidani et al., 2003). Critical realism is considered ‘a natural fit’ for nursing research due to its emphasis on context, complexity and the interplay between social structures and human agency (Clark et al., 2008; McEvoy and Richards, 2003; Schiller, 2016; Williams et al., 2017). Moreover, the recent nursing and healthcare literature shows a considerable increase in research using a realist methodology, such as realist evaluation and realist synthesis philosophically drawing on critical realism (Palm and Hochmuth, 2020; Williams et al., 2017).

The history of the MRC Framework reflects the growing interest in alternative research designs and in critical realism for evaluating interventions (Craig et al., 2008; Skivington et al., 2021). While emphasizing the importance of understanding an intervention theoretically from the outset (Medical Research Council, 2000), it took nearly two decades of interdisciplinary debate and several revisions before arriving at a clear methodological commitment to theory-driven intervention research and evaluation. This is evident in the most recent update of the MRC Framework (Skivington et al., 2021). Thus, it can be argued that theory-based methodologies have finally entered the main stage of health intervention research, including nursing research. This is evidenced by an increasing number of publications reviewing the use of theory-based evaluations and reporting on program theory development (see for instance Adlbrecht et al., 2018; Dichter, 2022; Höhmann and Bartholomeyczik, 2013; Kohler et al., 2020; Palm et al., 2021; Palm and Hochmuth, 2020).

Nursing has traditionally drawn on knowledge from a variety of disciplines and sources (Meleis, 2018; van Dulmen et al., 2017). Combined with an emergent nature predicated by its process orientation, this comes at the cost of nursing struggling to articulate its unique contribution to health and wellbeing of people (Bender, 2018), despite its clear and undisputable societal mandate (Kirkevel, 2002). Program theorizing offers one way of unraveling the operating principles of nursing interventions. It explicates the mechanisms of impact and demonstrates how a nursing intervention contributes to positive health trajectories for patients. Provided that this is pursued with the necessary methodological rigor and reflexivity, program theorizing may contribute to strengthening the theoretical basis of nursing in the long run. It may also relegate returning debates on the legitimacy of nursing as a scientific discipline (Grace and Zumstein-Shaha, 2020).

Since program theorizing is now considered as a core process in impact research, it may contribute to reviving theory construction in nursing. This has been regarded as a desideratum in various ways (Grace et al., 2016; Roy, 2019). First, program theorizing addresses various meanings of theory as well as processes of theory construction (von Werthern, 2020). Second, it requires competencies and procedures fundamentally similar to those necessary for formal theory development. In this respect, our understanding of knowledge production may be enhanced. Furthermore, competencies reaching beyond traditional research skills (e.g., qualitative or quantitative methods) may be developed. Thus, program theorizing offers a chance to approach constitutive elements of nursing interventions via concrete empirical objects. Syntheses of program theories may be a suitable means of enhancing our understanding of mechanisms mediating the effect of nursing practice in general. It may also improve our knowledge about how individual components interact. Focusing on so called ‘program archetypes’, i.e. classes of interventions with generic outcomes chains that can be used as program theory building blocks (Funnell and Rogers, 2011), as

well as building capacity and developing competence in program theorizing (Lemire et al., 2019) may be promising. Developing a heuristic of nursing interventions outlining its basic logic could provide valuable insights. Furthermore, this heuristic may be adapted to specific contents. Therefore, program theorizing may form an initial step toward a theory of nursing interventions.

The iterative nature of program theorizing has important implications with regard to resources that are already scarce, including time, finances, and capacity (Lemire et al., 2019). To improve the efficiency of program theory construction, Lemire et al. (2019) have suggested to adopt long-term perspectives regarding knowledge accumulation, and to demonstrate the value of program theorizing to decision makers. This will undoubtedly require a great deal of persuasiveness (Pawson, 2022). Seriously implementing theory building requires funding that allows researchers to follow a topic for more than the typical project duration of three years. However, if one considers how many resources are invested in basic research, this demand is not excessive but reasonable.

Effective and efficient program theorizing also requires fundamental theory building skills. With regard to competence development in scientific knowledge production, academic training is mostly focused on methodical and technical aspects of empirical research. "Whereas most graduate programs in the social sciences require multiple courses in research methodology to ensure that students become equipped with the tools to test theories empirically, the same can not be said for theory construction" (Jaccard and Jacoby, 2020, p. 3). This results in a one-sided understanding of knowledge production, namely focusing on empirical research in the sense of producing evidence (Mayer, 2019).

Working with and on program theories can be considered as building a bridge between empirical research and theory construction. Program theory is mostly based on empirically collected data followed by several iterations until the initial theory has been developed. This combination of empirical data and theory building can help to lose the fear of creative processes and to understand the value of theory in the context of empirical research once again from a different perspective. An important prerequisite is the ability to analyze, to synthesize, and to think 'out of the box' (Mayer, 2019).

6. Concluding thoughts

According to the logic of the updated MRC Framework, our understanding of efficacy and effectiveness of an intervention increasingly gives way to questions regarding mechanisms of action and the interplay between the intervention and the context/system. Therefore, the updated MRC guidance could represent an important step leading to a paradigm shift in nursing intervention/effectiveness research.

However, some aspects are not yet entirely methodologically convincing from a strictly philosophical point of view. Historically, the efficacy/effectiveness perspective as well as the theory-based and systems perspective are considered antagonistic and, indeed, incommensurable (Pawson, 2022). Hence, the combination of randomized controlled trials with realist evaluation has led to a lively academic discourse in the past (Bonell et al., 2012; Bonell et al., 2016; Fletcher et al., 2016; Marchal et al., 2013; Porter et al., 2017; Van Belle et al., 2016). In German-speaking countries, theory-based evaluation has attracted attention in the context of a methodological debate on the development and evaluation of complex nursing interventions that has been ongoing for nearly a decade. This methodological debate is entangled with a discussion regarding the status of nursing theory construction and development (Balzer et al., 2012; Meyer, 2010; Moers et al., 2011). As a result of pursuing evidence-based nursing, theory construction and development has arguably come to a halt. The updated MRC Framework may resuscitate this academic discussion.

According to Kuhn (1962), researchers engaging in philosophical discussion may be a sign of a paradigm in crisis. As with the movement of evidence-based medicine (Greenhalgh et al., 2014), this might also be the case in evidence-based nursing. Moving away from the paradigm of

cause and effect toward an understanding of how, why and under which conditions interventions work, could bring nursing research one step closer to understanding nursing interventions. This, in turn, may contribute to the development of useful, empirically supported nursing interventions. In terms of the studies that we examined at the beginning, a theory-based evaluation approach may have enabled the researchers to give important recommendations that are indispensable for implementation of their respective interventions in different contexts. It may also have improved their understanding why interventions did not work as expected.

Rather than considering theory-based evaluation as incompatible with designs associated with evidence-based nursing, it could be viewed as leading to a renewed understanding of what constitutes useful knowledge in nursing. This may facilitate knowledge production and, thereby, contribute to improve nursing practice for the benefit of the patient.

Funding

The authors received funding from the German Research Foundation to build up an international research collaboration. The discussion paper was created within this collaboration.

CRediT authorship contribution statement

Martin Wallner: Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Hanna Mayer:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Supervision. **Laura Adlbrecht:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Anna Louisa Hoffmann:** Conceptualization, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **Anne Fahsold:** Conceptualization, Writing – review & editing. **Bernhard Holle:** Conceptualization, Writing – review & editing. **Adelheid Zeller:** Conceptualization, Writing – review & editing. **Rebecca Palm:** Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ijnurstu.2023.104451>.

References

- Abraham, J., Bake, M., Berger-Höger, B., Köpke, S., Kupfer, R., Meyer, G., Möhler, R., 2021. Process evaluation of a multicomponent intervention to prevent physical restraints in nursing homes (IMPRINT): a mixed methods study. *J. Adv. Nurs.* 77 (3), 1465–1477. <https://doi.org/10.1111/jan.14694>.
- Abraham, J., Kupfer, R., Behncke, A., Berger-Höger, B., Icks, A., Haastert, B., Möhler, R., 2019. Implementation of a multicomponent intervention to prevent physical restraints in nursing homes (IMPRINT): a pragmatic cluster randomized controlled trial. *Int. J. Nurs. Stud.* 96, 27–34. <https://doi.org/10.1016/j.ijnurstu.2019.03.017>.
- Adlbrecht, L., Bartholomeyczik, S., Mayer, H., 2018. Entwicklung einer programmtheorie als grundlage für die evaluation einer demenzstation. *Pflege* 31 (3), 125–134. <https://doi.org/10.1024/1012-5302/a000610>.
- Astbury, B., Leeuw, F.L., 2010. Unpacking black boxes: mechanisms and theory building in evaluation. *Am. J. Eval.* 31 (3), 363–381. <https://doi.org/10.1177/1098214010371972>.
- Balzer, Köpke, Langer, Meyer, Behrens, 2012. Theorieerne Evidenzbasierung? Replik zum Beitrag von Moers et al., *Pflege*, Dezember 2011, Heft 6, Themenschwerpunktheft Theorieentwicklung in der Pflege im 21. Jahrhundert. *Pflege* 25 (2), 137–141. <https://doi.org/10.1024/1012-5302/a000189>.
- Bender, M., 2018. Re-conceptualizing the nursing metaparadigm: articulating the philosophical ontology of the nursing discipline that orients inquiry and practice. *Nurs. Inq.* 25 (3), e12243. <https://doi.org/10.1111/nin.12243>.

- Bhaskar, R., 2013. *A Realist Theory of Science*. Routledge.
- Blamey, A.A.M., MacMillan, F., Fitzsimons, C.F., Shaw, R., Mutrie, N., 2013. Using programme theory to strengthen research protocol and intervention design within an RCT of a walking intervention. *Evaluation* 19 (1), 5–23. <https://doi.org/10.1177/1356389012470681>.
- Bonell, C., Fletcher, A., Morton, M., Lorenc, T., Moore, L., 2012. Realist randomised controlled trials: a new approach to evaluating complex public health interventions. *Soc. Sci. Med.* 75 (12), 2299–2306. <https://doi.org/10.1016/j.socscimed.2012.08.032>.
- Bonell, C., Warren, E., Fletcher, A., Viner, R., 2016. Realist trials and the testing of context-mechanism-outcome configurations: a response to Van Belle et al. *Trials* 17, 478 England.
- Brousselle, A., Buregeya, J.-M., 2018. Theory-based evaluations: framing the existence of a new theory in evaluation and the rise of the 5th generation. *Evaluation* 24 (2), 153–168. <https://doi.org/10.1177/1356389018765487>.
- Burns, N., Grove, S.K., 2011. *Understanding Nursing Research: Building an Evidence-Based Practice*. 5 ed. Elsevier Saunders, Maryland Heights.
- Campbell, M., Fitzpatrick, R., Haines, A., Kinmonth, A.L., Sandercock, P., Spiegelhalter, D., Tyrer, P., 2000. Framework for design and evaluation of complex interventions to improve health. *BMJ* 321 (7262), 694–696. <https://doi.org/10.1136/bmj.321.7262.694>.
- Chen, H.T., 2015. *Practical Program Evaluation: Theory-driven Evaluation and the Integrated Evaluation Perspective*. 2 ed. SAGE, Thousand Oaks, California.
- Clark, A.M., Lissel, S.L., Davis, C., 2008. Complex critical realism: tenets and application in nursing research. *ANS Adv. Nurs. Sci.* 31 (4), E67–E79. <https://doi.org/10.1097/01.ANS.0000341421.34457.2a>.
- Corry, M., Clarke, M., While, A.E., Llor, J., 2013. Developing complex interventions for nursing: a critical review of key guidelines. *J. Clin. Nurs.* 22 (17–18), 2366–2386. <https://doi.org/10.1111/jocn.12173>.
- Coryn, C.L.S., Noakes, L.A., Westine, C.D., Schröter, D.C., 2011. A systematic review of theory-driven evaluation practice from 1990 to 2009. *Am. J. Eval.* 32 (2), 199–226. <https://doi.org/10.1177/1098214010389321>.
- Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., Petticrew, M., 2008. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 337, a1655. <https://doi.org/10.1136/bmj.a1655>.
- De Silva, M.J., Breuer, E., Lee, L., Asher, L., Chowdhary, N., Lund, C., Patel, V., 2014. Theory of change: a theory-driven approach to enhance the Medical Research Council's framework for complex interventions. *Trials* 15 (1), 267. <https://doi.org/10.1186/1745-6215-15-267>.
- Dichter, M.N., 2022. Logische modelle und der theory of change-ansatz zur theoretischen fundierung pflegewissenschaftlicher interventionen. *Pflege* 35 (1), 1–3. <https://doi.org/10.1024/1012-5302/a000855>.
- Döring, N., Bortz, J., Pöschl, S., Werner, C.S., Schermelleh-Engel, K., Gerhard, C., Gäde, J.C., 2016. *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. 5 ed. Springer, Berlin, Heidelberg.
- Dössegger, A., Weibel, D., Frei, K.M., Wissmath, B., Hense, H., 2017. *Entwicklung eines wirkmodells für die evaluation des programms jugend und sport*. *Z. Eval.* 16 (1), 97–120.
- Duxbury, J., Baker, J., Downe, S., Jones, F., Greenwood, P., Thygesen, H., Whittington, R., 2019. Minimising the use of physical restraint in acute mental health services: the outcome of a restraint reduction programme ('RESTRRAIN YOURSELF'). *Int. J. Nurs. Stud.* 95, 40–48. <https://doi.org/10.1016/j.ijnurstu.2019.03.016>.
- Fletcher, A., Jamal, F., Moore, G., Evans, R.E., Murphy, S., Bonell, C., 2016. Realist complex intervention science: applying realist principles across all phases of the Medical Research Council framework for developing and evaluating complex interventions. *Evaluation* 22 (3), 286–303. <https://doi.org/10.1177/1356389016652743>.
- Fromer, R.F., 2017. Theory-driven integrative process/outcome evaluation of a concept-based nursing curriculum. *Nurs. Educ. Perspect.* 38 (5), 267–269. <https://doi.org/10.1097/01.Nep.0000000000000140>.
- Funnell, S.C., Rogers, P.J., 2011. *Purposeful Program Theory: Effective Use of Theories of Change and Logic Models*. Jossey-Bass, San Francisco.
- Gentile, D., Marzinski, S.J., 2020. Theory-driven evaluation of a multisite nursing professional practice model. *J. Nurs. Adm.* 50 (7/8), 419–425. <https://doi.org/10.1097/nnn.0000000000000910>.
- Giel, S., 2013. *Theoriebasierte Evaluation. Konzepte und methodische Umsetzungen*. Waxmann, Berlin.
- Grace, P.J., Willis, D.G., Roy, C., Jones, D.A., 2016. *Profession at the crossroads: a dialog concerning the preparation of nursing scholars and leaders*. *Nurs. Outlook* 64 (1), 61–70. <https://doi.org/10.1016/j.outlook.2015.10.002>.
- Grace, P.J., Zumstein-Shaha, M., 2020. Using Ockham's razor to redefine "nursing science". *Nurs. Philos.* 21 (2). <https://doi.org/10.1111/nup.12246>.
- Greenhalgh, T., Howick, J., Maskrey, N., Evidence Based Medicine Renaissance, G., collab, 2014. Evidence based medicine: a movement in crisis? *BMJ* 348, g3725. <https://doi.org/10.1136/bmj.g3725>.
- Gugiu, P.C., Rodríguez-Campos, L., 2007. Semi-structured interview protocol for constructing logic models. *Eval. Program Plann.* 30 (4), 339–350. <https://doi.org/10.1016/j.evalprogplan.2007.08.004>.
- Haubrich, K., 2009. *Sozialpolitische Innovation ermöglichen: Die Entwicklung der rekonstruktiven Programmtheorie-Evaluation am Beispiel der Modellförderung in der Kinder- und Jugendhilfe*. Waxmann, Berlin.
- Höhm, U., Bartholomeyczik, S., 2013. *Komplexe Wirkungszusammenhänge in der pflege erforschen: konzepte statt rezepte*. *Pflege Ges.* 18 (4), 293–312.
- Jaccard, J., Jacoby, J., 2020. *Theory Construction and Model-Building Skills, Second Edition: A Practical Guide for Social Scientists*. The Guilford Press, New York, London.
- Jiang, Y., Koh, K.W.L., Ramachandran, H.J., Nguyen, H.D., Lim, S., Tay, Y.K., Wang, W., 2021. The effectiveness of a nurse-led home-based heart failure self-management programme (the HOM-HEMP) for patients with chronic heart failure: a three-arm stratified randomized controlled trial. *Int. J. Nurs. Stud.* 122, 104026. <https://doi.org/10.1016/j.ijnurstu.2021.104026>.
- Kirkevold, M., 2002. *Pflegewissenschaft als Praxisdisziplin*. Huber, Bern.
- Kohler, M., Mayer, H., Kesselring, J., Saxer, S., 2020. Urinary incontinence in stroke survivors – development of a programme theory. *J. Clin. Nurs.* 29 (15–16), 3089–3096. <https://doi.org/10.1111/jocn.15345>.
- Kuhn, T., 1962. *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago.
- Lemire, S., Whynot, J., Montague, S., 2019. How we model matters: a manifesto for the next generation of program theorizing. *Can. J. Program Eval.* 33 (3), 414–433. <https://doi.org/10.3138/cjpe.53070>.
- Lin, R.S.Y., Yu, D.S.F., Chau, P.H., Li, P.W.C., 2023. Effects of an empowerment-based educative psycho-behavioral program on neuropsychiatric symptoms among persons with mild cognitive impairment: a mixed methods study. *Int. J. Nurs. Stud.* 137, 104381. <https://doi.org/10.1016/j.ijnurstu.2022.104381>.
- Marchal, B., Westhorp, G., Wong, G., Van Belle, S., Greenhalgh, T., Kegels, G., Pawson, R., 2013. Realist RCTs of complex interventions – an oxymoron. *Soc. Sci. Med.* 94, 124–128. <https://doi.org/10.1016/j.socscimed.2013.06.025>.
- Mayer, H., 2019. Was Hänschen nicht lernt, lernt Hans nimmermehr? Ein Beitrag über die Rolle der lehre im zusammenhang mit der theorieentwicklung in der pflege. *Pflege Ges.* 24 (2), 110–122.
- McEvoy, P., Richards, D., 2003. Critical realism: a way forward for evaluation research in nursing? *J. Adv. Nurs.* 43 (4), 411–420. <https://doi.org/10.1046/j.1365-2648.2003.02730.x>.
- Medical Research Council, 2000. A framework for development and evaluation of RCTs for complex interventions to improve health. Retrieved from <https://mrc.ukri.org/documents/pdf/rcts-for-complex-interventions-to-improve-health/>.
- Meleis, A.I., 2018. *Theoretical Nursing: Development and Progress*. Wolters Kluwer.
- Meyer, G., 2010. Wohin des weges, Pflegeforschung? *Pflege* 23 (3), 151–153. <https://doi.org/10.1024/1012-5302/a000034>.
- Moers, M., Schaeffer, D., Schnepf, W., 2011. Too busy to think? Essay über die spärliche theoriebildung der deutschen pflegewissenschaft. *Pflege* 24 (6), 349–360. <https://doi.org/10.1024/1012-5302/a000151>.
- Möhler, R., Köpke, S., Meyer, G., 2015. Criteria for reporting the development and evaluation of complex interventions in healthcare: revised guideline (CREDECI 2). *Trials* 16, 204. <https://doi.org/10.1186/s13063-015-0709-y>.
- Palm, R., Fahsold, A., Roes, M., Holle, B., 2021. Context, mechanisms and outcomes of dementia special care units: an initial programme theory based on realist methodology. *PLOS ONE* 16 (11), e0259496. <https://doi.org/10.1371/journal.pone.0259496>.
- Palm, R., Hochmuth, A., 2020. What works, for whom and under what circumstances? – Using realist methodology to evaluate complex interventions in nursing: a scoping review. *Int. J. Nurs. Stud.* 103601. <https://doi.org/10.1016/j.ijnurstu.2020.103601>.
- Paterson, C., Baarts, C., Launso, L., Verhoef, M.J., 2009. Evaluating complex health interventions: a critical analysis of the 'outcomes' concept. *BMC Complement. Altern. Med.* 9, 18. <https://doi.org/10.1186/1472-6882-9-18>.
- Pawson, R., 2022. Rapid response to: a new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *BMJ* 374, n2061. <https://doi.org/10.1136/bmj.n2061>.
- Pawson, R., Tilley, N., 1997. *Realistic Evaluation*. SAGE Publications, Thousand Oaks.
- Porter, S., McConnell, T., Reid, J., 2017. The possibility of critical realist randomised controlled trials. *Trials* 18, 133. <https://doi.org/10.1186/s13063-017-1855-1>.
- Rahn, A.C., Köpke, S., Backhus, I., Kasper, J., Anger, K., Untiedt, B., Heesen, C., 2018. Nurse-led immunotreatment Decision coaching in people with multiple sclerosis (DECIMS) – feasibility testing, pilot randomised controlled trial and mixed methods process evaluation. *Int. J. Nurs. Stud.* 78, 26–36. <https://doi.org/10.1016/j.ijnurstu.2017.08.011>.
- Reith, F., Kelle, U., 2016. Subjektive programmtheorien als heuristisch-analytische rahmenkonzepte zur semi-operationalisierung in der organisationsforschung. In: Liebig, S., Matiaske, W. (Eds.), *Methodische Probleme in der Empirischen Organisationsforschung*. Springer Fachmedien Wiesbaden, Wiesbaden, pp. 85–109.
- Renmans, D., Holvoet, N., Criel, B., 2017. Combining theory-driven evaluation and causal loop diagramming for opening the 'Black Box' of an intervention in the health sector: a case of performance-based financing in Western Uganda. *Int. J. Environ. Res. Public Health* 14 (9). <https://doi.org/10.3390/ijerph14091007>.
- Roy, C., 2019. Nursing knowledge in the 21st century: domain-derived and basic science practice-shaped. *ANS Adv. Nurs. Sci.* 42 (1), 28–42. <https://doi.org/10.1097/ans.0000000000000240>.
- Schiller, C.J., 2016. Critical realism in nursing: an emerging approach. *Nurs. Philos.* 17 (2), 88–102. <https://doi.org/10.1111/nup.12107>.
- Shearn, K., Allmark, P., Piercy, H., Hirst, J., 2017. Building realist program theory for large complex and messy interventions. *Int J Qual Methods* 16 (1). <https://doi.org/10.1177/1609406917741796>.
- Sidani, S., Braden, C.J., 1998. *Evaluating Nursing Interventions: A Theory-Driven Approach*. SAGE Publications, Thousand Oaks.
- Sidani, S., Braden, C.J., 2021. *Nursing and Health Interventions: Design, Evaluation, and Implementation*. 2 ed. Wiley Blackwell, Hoboken, NJ.
- Sidani, S., Epstein, D.R., Moritz, P., 2003. An alternative paradigm for clinical nursing research: an exemplar. *Res. Nurs. Health* 26 (3), 244–255. <https://doi.org/10.1002/nur.10086>.
- Skivington, K., Matthews, L., Simpson, S.A., Craig, P., Baird, J., Blazeby, J.M., Moore, L., 2021. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *BMJ* 374, n2061. <https://doi.org/10.1136/bmj.n2061>.
- Stame, N., 2017. *US sociology and evaluation: issues in the relationship between methodology and theory*. In: Vaessen, J., Leeuw, F.L. (Eds.), *Mind the Gap. Perspectives on Policy Evaluation and Social Sciences*. Routledge, pp. 29–44.
- Thomson Reuters, 2022. 2021 Journal citation reports®. Retrieved from <https://jcr.clarivate.com/uni-wh.idm.oclc.org/jcr/browse-journals>.

- Vaessen, J., Leeuw, F.L., 2017. *Mind the Gap. Perspectives on Policy Evaluation and Social Sciences*. Routledge.
- Van Belle, S., Wong, G., Westhorp, G., Pearson, M., Emmel, N., Manzano, A., Marchal, B., 2016. Can "realist" randomised controlled trials be genuinely realist? *Trials* 17, 313. <https://doi.org/10.1186/s13063-016-1407-0>.
- van Duinen-van den IJssel, J.C.L., Bakker, C., Smalbrugge, M., Zwijsen, S.A., Appelhof, B., Teerenstra, S., Koopmans, R., 2019. Effects on staff outcomes from an intervention for management of neuropsychiatric symptoms in residents of young-onset dementia care units: a cluster randomised controlled trial. *Int. J. Nurs. Stud.* 96, 35–43. <https://doi.org/10.1016/j.ijnurstu.2019.03.006>.
- van Dulmen, S., McCormack, B., Eide, T., Skovdahl, K., Eide, H., 2017. *Future directions for person-centred healthcare research*. *Person-Centred Healthcare Research*. John Wiley & Sons Ltd., pp. 209–218.
- von Werthern, A., 2020. *Theoriebasierte Evaluation. Entwicklung und Anwendung eines Verfahrensmodells zur Programmtheoriekonstruktion*. Springer VS, Wiesbaden.
- Walsh, D., Evans, K., 2014. Critical realism: an important theoretical perspective for midwifery research. *Midwifery* 30 (1), e1–e6. <https://doi.org/10.1016/j.midw.2013.09.002>.
- Williams, L., Rycroft-Malone, J., Burton, C.R., 2017. Bringing critical realism to nursing practice: Roy Bhaskar's contribution. *Nurs. Philos.* 18 (2). <https://doi.org/10.1111/nup.12130>.
- Xiao, J., Chow, K.M., Choi, K.C., Ng, S.N.M., Huang, C., Ding, J., Chan, W.H.C., 2022. Effects of family-oriented dignity therapy on dignity, depression and spiritual well-being of patients with lung cancer undergoing chemotherapy: a randomised controlled trial. *Int. J. Nurs. Stud.* 129, 104217. <https://doi.org/10.1016/j.ijnurstu.2022.104217>.