

Nielsen, K., & Miraglia, M. (2016, August 23). Critical essay: What works for whom in which circumstances? On the need to move beyond the “what works?” question in organizational intervention research. *Human Relations* (accepted for publication).

'This article may not exactly replicate the final version published in Human Relations. It is not the copy of record.'

Critical essay: What works for whom in which circumstances? On the need to move beyond the “what works?” question in organizational intervention research

Abstract:

Organizational interventions can be defined as planned, behavioural, theory-based actions that aim to improve employee health and well-being through changing the way work is designed, organized and managed. The gold standard for evaluating the effectiveness of such interventions is the use of randomized, controlled trials and meta-analyses. The key question according to this paradigm is whether this type of interventions work or not, however, answering this questions offers little practical insights into how organizations and policy makers may plan and implement future organizational interventions. In the present critical essay, we discuss the limitations of this paradigm. We argue that realist evaluation and realist syntheses may offer a more informative alternative. The key questions of realist evaluation are what works for whom in which circumstances. These questions are explored through the identification of the mechanisms that bring about desired outcomes and in which contextual circumstances such mechanisms may be triggered (CMO-configurations). In the present essay, we suggest that organizational intervention content and process mechanisms may help bring about the desired outcomes of improved employee health and well-being and that the omnibus and discrete contextual factors determine whether mechanisms are facilitated.

Keywords: Meta-analysis, randomized controlled trial, realist evaluation, realist synthesis, CMO-configurations, organizational interventions, critical essay

Critical essay: What works for whom in which circumstances? On the need to move beyond the “what works?” question in organizational intervention research

Organizational interventions can be defined as planned, behavioural, theory-based actions that aim to improve employee health and well-being through changing the way work is designed, organized and managed (e.g. Nielsen, 2013a; Richardson and Rothstein, 2008). Such interventions most often employ a problem solving cycle approach consisting of five phases: Setting up the intervention, screening of existing employees’ working conditions and employee health and well-being and feedback of screening results to participants, developing action plans to improve working conditions, implementation of action plans, and evaluation of the intervention’s outcomes. Organizational interventions most often employ a participatory approach where employees and (line) managers collaboratively decide on the process and the content of the intervention (Nielsen et al., 2010). This type of intervention is recommended by the World Health Organization (WHO, 1995), the International Labor Organization (ILO, 2001), and the European Agency for Occupational Safety and Health (EU-OSHA, 2010) because they target the causes of poor well-being, i.e., a poor working environment. Furthermore, the framework European Directive (European Communities, 1989; ETUC, 2004) requires that employers ensure the health and well-being of employees are based on prevention principles, i.e. avoidance of the risks, evaluation of the risks which cannot be avoided, fighting the risks at source, and adaption of the work to employees (Malchaire, 2004). Yet reviews and meta-analyses conclude that organizational interventions are by and large ineffective (Bhui et al., 2012; Richardson and Rothstein, 2008; van der Klink et al., 2001).

In the present essay, we discuss the limitations of classification systems that assume that the randomized, controlled, trial (RCT) and meta-analyses offer the highest level of

scientific evidence and propose that realistic evaluation and realist syntheses (Pawson and Tilley, 1997; Pawson, 2013) with its questions of what works for whom in which circumstances may successfully be applied to the area of organizational intervention evaluation. Realist evaluation seeks to answer the questions of what works for whom in which circumstances through studying what the Mechanisms of an intervention are (what makes an intervention work?) and the Context in which these may be triggered (what are the conditions in which an intervention is effective?) and bring about certain Outcomes (which improvements in working conditions and in employee health and well-being can be observed?), in what is also known as CMO-configurations (Pawson and Tilley, 1997).

In realist evaluation, an intervention's programme theory, i.e. the theory of what causes change, is the unit of analysis, not whether an intervention worked or not (Greenhalgh et al., 2015). A programme theory can be seen as an implicit set of assumptions that steers the choice and design of an intervention (Pawson, 2013) and spells out the coordinated sequence of activities (e.g., preparation, screening, action planning, implementation and evaluation) that are thought to be necessary to achieve the intervention's intended outcomes (Manzano and Pawson, 2014).

RCTs and meta-analyses - what are they good for?

Central tenets of a study's quality concern its internal and external validity (generalizability). Internal validity is an indicator of the extent to which a cause-and-effect relationship between working conditions and its outcome is well-founded (Hoyle et al., 2001), and it is crucial to evaluate whether an intervention produced any observed changes and to estimate the magnitude of these changes (Newcomer et al., 2010). To determine which designs are the most robust in terms of internal validity, the so-called 'levels of evidence' are used (Guyatt et al., 1995). These levels describe the hierarchical ordering of research design quality. A study has high internal validity when it fulfils the three conditions required for

causal inference: covariation, time–order relationship, and elimination of plausible alternative causes (Shaugnessy et al., 2006). Drawing on the medical field, randomized controlled trials (RCTs) randomly allocate participants into either an intervention or a control group. The basic assumption is that any differences in effects in the two groups are attributable to the results of the intervention rather than to any individual differences at baseline, e.g. gender, age or existing levels of resources. The goal of the RCT design is to assess whether the intervention had the desired effect or not, to answer the question of “what works?”. RCTs are supposed to deal with the threats to internal validity, by addressing the three aforementioned conditions for causal inference (Shaugnessy et al., 2006).

However, the complexity of organizational interventions must be considered when applying the RCT design to this type of interventions.

First, the RCT design only allows researchers to determine whether there was an effect or not; it does not demonstrate whether this effect can be ascribed to the intervention itself or to other factors. For example, it is not possible to determine whether any changes in working conditions and employee health and well-being have been caused by the intervention or other factors such as concurrent events in the form of downsizing or additional resource allocation.

Second, it is well-established that the intervention process influences the intervention’s outcomes. For example, it has been argued that the behaviours and involvement of key players in the intervention process actively craft the intervention (Nielsen, 2013a). Research has established that line managers play an important role in shaping and supporting the intervention process and that these behaviours are related to observed improvements in targeted working conditions and employee well-being (Nielsen and Randall, 2009). Also the participatory process, i.e. the extent to which employees are involved in determining the intervention’s processes and content, plays a role in ensuring a

successful intervention outcome (Nielsen and Randall, 2012). These outcomes are also strongly influenced by the multiple perceptions and interpretations of the intervention by the participants. In fact, due to the participatory, problem solving cycle, the content of the intervention, i.e. the action plans, are not determined prior to the intervention but jointly developed by (line) managers and employees. RCTs assume that the intervention is the causal effect, but at the same time they exclude the many components that may explain an intervention's outcomes, possibly disregarding the real causes of the intervention's success or failure (Pawson, 2013). It is essential to recognize that workplace interventions are "active" rather than passive programmes (Pawson, 2006) that continuously respond to contextual factors and emerging processes.

Third, organizational interventions target an entire work unit; however, not all of participants in an organizational unit, e.g. a department, will suffer from the same issues. For instance, although a majority of employees may think lack of managerial support may be a problem, not all employees may agree and they may also have different perceptions on the appropriate activities to be implemented to improve such support. These various interpretations may differently affect the intervention functioning and realization.

Third, there are several challenges associated with the use of control groups. It is notoriously difficult to finding equivalent control groups that enable reliable comparisons between the intervention and the control group (Pawson, 2013). If an entire department is participating, there may not be another department in the organization that provides an adequate match. For example, in hospitals there is only one Accidents and Emergency, and an Accident and Emergency at another unit may have different procedures and policies. Different production units in manufacturing plants may differ in terms of size and production methods within the same organization. The basic tenet of the RCT design is that randomization controls for baseline levels but because of contamination risks, organizational

interventions employ cluster randomization and thus N is too small to rule out differences between groups. Also organizational interventions may be a response to national legislation (e.g. in Denmark it is a legal requirement to assess the working environment at least every third year and take appropriate measures to address any issues, Nielsen, 2013b) and thus the intervention must be implemented in the entire organization. In addition, it may be considered unethical to withhold the intervention from a group that has been identified as “at risk” and wait-list control design may result in aggravation of problems in the wait-list group (Nielsen, 2013a). Furthermore, research indicates that the sheer act of randomization increases readiness for change in the intervention group due to feelings of “having won the lottery” (Nabe-Nielsen et al., 2015).

A further challenge of RCTs is related to their external validity or generalizability of the findings. External validity refers to the extent to which an intervention will have the same effects in other settings or with other groups (Newcomer et al., 2010). One way of achieving generalizability is to conduct the intervention across multiple sites. Other ways include making explicit the settings in which an intervention is expected to work. For example, is an intervention that relies on high levels of technology expected to be generalizable beyond a group of “tech savvy” engineers? It is also important to evaluate the generalizability from the sample size. Does an intervention with ten postal service workers represent the entire population across an entire country regardless of whether they work in urban or rural areas? RCTs are implemented in a specific context and discrete period of time, making the generalization of the findings questionable across contexts (Karlan et al., 2009) and offering limited information on the transferability of complex interventions.

Addressing these challenges related to internal and external validity, Newcomer et al. (2010) recommended that when conducting evaluations attention should be paid to measuring the extent to which an intervention was implemented (a manipulation check), examining

whether any other events may account for the observed changes, evaluating whether sufficient time has lapsed for any changes to be attributed to the intervention (e.g. an intervention aimed at improving trust between management and employees may not work overnight but may take time to take effect), and assessing whether similar interventions have produced similar results.

In organizational intervention evaluation, meta-analyses are considered the gold standard when aiming to estimate the overall effectiveness of interventions and to draw evidence-based conclusions and recommendations (Rousseau et al., 2008). By quantitatively combining existing studies, meta-analyses statistically synthesize the results of several RCTs across contexts, to obtain a single weighted average measure of the effect of a certain intervention (i.e., its effect size). This strategy could give the illusion that RCTs may overcome the contextual specificity of RCTs. Meta-analyses have the merit to attain statistical precision, by weighting the estimated effect size on sample size and study artifacts (Schmidt and Hunter, 2015), and by assessing across-study variability around this effect size (Borenstein et al., 2009). Moreover, they strive to achieve methodological rigor and replicability, by transparently reporting the numerous procedural judgements and choices a meta-analyst faces.

Overall, meta-analyses of organizational interventions have concluded that organizational interventions are ineffective (Rothstein and Richardson, 2008; van der Klink et al., 2001). In a synthesis of systematic reviews and meta-analyses, Bhui et al. (2012) concluded that organizational interventions show limited and mixed effects. We question whether such meta-analyses and reviews are appropriate when synthesizing organizational intervention studies, and we argue that they suffer from a number of pitfalls that may invalidate their findings and any related evidence-based recommendations. Some of these

drawbacks are idiosyncratic to meta-analyses in general, others are even more critical when trying to draw conclusions on the effectiveness of organizational interventions.

Pawson (2006) challenged the overall logic of the meta-analytic approach for intervention evaluation, by critically illustrating how the multiple stages of meta-analysis (from formulating the review question to disseminating the findings) fail to realistically review interventions in the social science domain. The main point of his critique is related to the over-simplistic approach underlying meta-analytic evaluation. Indeed, in order to achieve rigour and net effect-size estimates, simplifications are used at multiple stages of the meta-analytic process (Pawson, 2006). For example, when first sifting the literature, clear criteria are set to decide which studies should be included or excluded from the analyses, eliminating most of the contextual information and details of the interventions. This over-simplification can produce statistically accurate information about the interventions' effectiveness (e.g., organizational interventions only show little effect), but may actually fail to detect the underlying mechanisms of an organizational intervention that would disclose how any effects are brought about. Organizational interventions most often operate through a participatory, dynamic approach, where employees and managers in collaboration decide on the process and the content of intervention activities and through on-going negotiations shape the intervention process and the content of action plans (Nielsen et al., 2010; Nielsen et al., 2014). It has been argued that the intervention's component, i.e. the development of action plans, trigger certain behaviours of participants, i.e. their engagement in developing and implementing these action plans and it is these behaviours that produce outcomes rather than the component "developing action plans" (Nielsen and Abildgaard, 2013).

A common criticism against meta-analysis is that it mixes apples and oranges. This is particularly critical for intervention studies, which are highly dependable on contextual factors. It could be argued that meta-analyses can perform moderator (and sensitivity)

analyses in order to investigate those differences among primary studies attributable to such contextual factors or to intervention types. Moderator analyses are considered one of the strengths of the meta-analytic approach, because they recognize variation in the distribution of the estimated effect size and they aim to explain it (Borenstein et al., 2009). However, this variability may not easily be disentangled, since its source may not be disclosed from primary studies. In fact, primary studies may have not measured the variables that account for any differences among studies, such as contextual variables (or they may not report net statistical indicators of these). Moreover, moderator analyses are often performed through meta-regressions, which require to select an even more strict number of studies (i.e., all that studies which performed the same intervention in relation to the same outcomes, and reported useful statistics for the same moderators), further limiting the number of studies included in the subgroup analysis, and thus the power of the analyses (Borenstein et al., 2009).

A further challenge associated with meta-analyses is that the effect size does not provide useful insights into how organizations may target resources or maximize the impact of their efforts to improve employee health and well-being. Due to their high complexity, organizational interventions are likely to produce a complex configuration of outcomes (Pawson, 2006) and to provide wide-spread effects through organizational learning, which cannot be quantitatively expressed through a meta-analytic estimation. Indeed, meta-analyses need to precisely operationalize the expected outcomes to estimate the relations of interest, making difficult to assess any possible emerging “products” of it.

Finally, it is worth briefly mentioning some more general problems of meta-analyses, which further show the limits of the meta-analytic approach to evaluate organizational interventions. First, qualitative studies are excluded from meta-analyses for obvious reasons, but they can be useful to understand the processes, the contextual factors and mechanisms underlying the effectiveness of organizational interventions. The second issue pertains to the

quality of the primary studies. Since one poorly conducted study could erroneously alter the meta-analytic effect size, strict criteria are set to sift the literature in order to include in the meta-analysis only those studies which meet certain standards. In this way, the accurateness of the estimation is assured, but at the same time a whole set of valuable information about processes and mechanisms is excluded. Indeed, poorly conducted or unsuccessful studies can reveal a lot about how and why organizational interventions worked or did not, offering useful insights for planning future interventions (e.g. Nielsen et al., 2006). Third, directly linked to this issue, meta-analyses are vulnerable to publication bias, which occurs when the retrieved primary studies are systematically unrepresentative of the study population (Rothstein et al., 2005). Studies may be excluded from the literature (i.e., not published) because they did not confirm the expected hypotheses or did not achieve statistically significant results. Although recently more advanced methodologies have been used to assess the effect of publication bias (e.g., Duval and Tweedie, 2000; Sterne and Egger, 2005; Sterne et al., 2005), the failure to include those unpublished studies – that report unsuccessful results from an intervention point of view – not only threatens the validity of the meta-analysis (Dickersin, 2005), but also limits our understanding of the key factors underlying intervention effectiveness. An example of the limited usefulness of meta-analyses is the Nielsen et al. (2006) study, which compared two intervention studies and two non-randomized control groups. Improvements in working conditions and employee health and well-being were found in one intervention and one control group whereas deteriorations or no effects could be observed in the other intervention and control groups. A meta-analytic comparison of these groups would conclude that the intervention was ineffective. However, the study's process evaluation revealed that contextual differences and the fact that the "successful" control group implemented similar activities as the "successful" intervention group would suggest that the proposed mechanism of the intervention (e.g. team building activities) could help

explain outcomes. These results indicate that the intervention had an effective program theory (e.g. improving employee health and well-being through activities that improve social relations) and that the mechanism team building had a positive effect in the right circumstances (a context where baseline levels of social support were low). This example illustrates the difficulty for meta-analyses of dealing with inconsistent findings, which can result in misevaluating intervention studies and in losing important information about their functioning.

In summary, meta-analyses may fail to address the complexity of organizational interventions, thus underestimating their effectiveness. Meta-analyses aim to quantitatively summarize the specific relationships between certain causes and a set of expected outcomes. In so doing, they fail to consider the complex patterns made of contextual elements, emerging mechanisms, participative and recursive processes, participants' expectations, perceptions and interpretations that ultimately determine the success (or failure) of organizational interventions. In other words, they may fail to identify the true relationships of interest and they may not (meta)analyze those elements that account for the intervention's outcomes. This is also reflected by another important stage of the meta-analytical process, namely the extraction of the relevant information from the primary studies. In fact, key component of the intervention process and context are difficult to translate into precise, numerical variables, which is exactly what is needed to run a meta-analysis. Thus complex and precious information is lost.

In the attempt to address the limitations of RCTs and meta-analyses, there has been an increasing interest in understanding which specific elements of an organizational intervention may work (Nielsen et al. 2010) and how and why such interventions work or not (Nielsen et al., 2006; Nielsen and Randall, 2012). Some models have been developed that consider evaluating the process of interventions (Nielsen and Abildgaard, 2013; Nielsen and

Randall, 2013), however, existing thinking on the evaluation of organizational interventions suffer from not having an overall framework to guide such assessment. Although arguments for integrating process and effect evaluation have been put forward (Nielsen and Abildgaard, 2013), such models are not theoretically founded. Current frameworks are based on existing research and are therefore limited in their inclusion of the elements they recommend researchers to consider: It is possible that process factors may influence outcomes that have yet to be covered in the current literature.

In the present essay, we argue that realist evaluation (Pawson and Tilley, 1997, Pawson, 2013) may present a fruitful avenue to go down when evaluating organizational interventions. Realist evaluation offers a way to conduct rigorous, theory-based analyses of what works for whom in which circumstances. Such analyses provide researchers and evaluators to with insights into how to improve existing interventions and inform future interventions, and ensure both internal validity and external validity. In the following sections of the present essay, we will introduce realist evaluation and discuss how realist evaluation and its CMO-configurations may be applied to organizational intervention evaluation, and we discuss the realist synthesis as an alternative to meta-analyses. Realist evaluation may offer an opportunity to develop an integrated process and effect evaluation framework that may advance our theoretical understanding on which elements of organizational interventions may be effective and in which conditions we can expect positive outcomes.

The basic principles of realist evaluation

According to realist evaluation there are patterns that may explain why an intervention succeeds or fails to bring about the intended outcomes and we can build and test models to explain these patterns (Pawson, 2013). The focus of realist evaluation is to answer the questions of “what works for whom in which circumstances?” in an attempt to open the black box of how and why interventions may or may not work. The central tenet of realist

evaluation is to answer these questions through theoretically developing and testing Context+Mechanism = Outcome (CMO)-configurations (Pawson and Tilley, 1997; Pawson, 2013). A CMO-configuration can pertain to an entire intervention or to parts of it and one CMO-configuration can be embedded in another (Pawson and Tilley, 1997). The realist evaluation strategy thus focuses on three themes: Understanding the Mechanisms through which an intervention achieves its Outcomes, understanding the Contextual conditions necessary for triggering these mechanisms, and understanding outcome patterns (Pawson and Tilley, 1997). Some factors in the context may enable certain mechanisms to trigger intended outcomes (Greenhalgh et al., 2015) and therefore interventions cannot simply be transferred from one context to the other; there is always an interaction between context and mechanisms and it is this interaction that creates the intervention's outcomes (Greenhalgh et al., 2015). It is the interplay between participants in the intervention and the structures in which the intervention is embedded that determine the outcomes of the intervention and research should thus focus on how these agent-structure interactions produce outcomes (Greenhalgh et al., 2015).

Realist evaluators argue that it is crucial to understand that because interventions work differently in different contexts and through different mechanisms they cannot easily be transferred from one setting to another, however, the in-depth understanding of what works for whom in which circumstances is portable (Goodridge et al., 2015). Realist evaluation has primarily been applied in the evaluation of healthcare (Marchal et al., 2012) and policy research (Pawson, 2013). In the following we offer an operationalization of mechanisms and context in organizational intervention research.

Mechanisms: What makes organizational interventions work?

Interventions work by giving participants the opportunity to make different choices about their agency (Greenhalgh et al., 2015). Making and sustaining such changes in behaviour requires a change in people's reasoning and/or the resources they have available to them (Greenhalgh et al., 2015). Mechanisms can be expressed through interpretations, considerations, decisions and behaviours of participants and outcomes are the result of these actions of participants and their interpretations of the intervention (Pawson, 2013). Despite the definition of mechanisms put forward by Pawson (2013), there is confusion as to what the term includes. In a review of the mechanism concept in realist evaluation, Lacouture et al. (2015) identified 49 studies based on realist evaluation. They concluded in their analyses of the characteristics of mechanisms that 1) mechanisms can be defined as the reasoning and reactions of participants, 2) mechanisms may change over the duration of the intervention and that they are latent and only reveal themselves during the implementation of the intervention, 3) although mechanisms cannot be observed they are real and exist regardless of whether they are activated or not, and 4) mechanisms interact with each other and may be linked in negative or positive feedback loops. In a review, Marchal et al. (2012) identified mechanisms at the individual (resistant behaviours), group (social capital building) and organizational (management behaviours) level. For example, if a group of employees do not believe a particular action plan will address a problem they are unlikely to drive the implementation of this action plan and thus the action plan will not achieve its intended outcome(s) because it was not implemented.

Although there has been little explicit use of realist evaluation in the organizational intervention domain, a few studies on organizational practices and policies have employed a realist evaluation approach and these identify mechanisms that may also be applicable to organizational interventions. Higgins et al. (2015) conducted an intervention aimed to facilitate return to work among employees on long-term sickness absence. One of the

mechanisms that brought about a fast return to work was identified as line managers' communication with employees on sick leave. Line managers' regular contact to employees at an early stage was also recognized as a mechanism as was senior managers providing clear guidance on how to manage employees on sick leave.

Goodridge et al. (2015) explored the implementation of Lean in Canadian healthcare. The aim of Lean was to “develop an infrastructure to support and coordinate continuous quality improvement” (Goodridge et al., 2015, p 2). The focus of the study was to develop hypotheses for CMO-configurations that could be tested in future studies of Lean implementation. Based on interviews with employees in the healthcare, the following mechanisms were proposed would trigger the intended outcomes of the ability of the healthcare to “do more with less” and provide better quality of care. First, authorization of resources to be spent on Lean activities will result in employees participating in such activities. Second, the use of a common set of Lean tools will ensure consistent improvement process and practices to be implemented thus generating improved decision making which will lead to better quality of care. Third, leaders at the local level will play a key role in promoting the use of Lean tools and engagement in quality improvement processes. Fourth, increased levels of staff and patient participation in quality improvement processes and visibility of leaders will contribute to increased transparency which will in turn ensure leader accountability for implementing Lean (Goodridge et al., 2015).

Although rarely explicitly examined in organizational intervention research, a number of studies have implicitly examined the mechanisms of such interventions. We argue that two main categories of mechanisms may be at play in organizational interventions. First, the content of the intervention, i.e. the content of action plans, may realize the intended outcomes. Second, the participatory process of the intervention may also be an important mechanism that brings about improved employee health and well-being.

Holman and Axtell (2015) provided two examples of the content of the intervention as a mechanism that brought about the intended improvements in employee well-being in UK call centre study. The authors suggested two content mechanisms by which the intervention may have an effect. First, administrative tasks being completed by employees rather than line managers and the discretion over how minor customer queries and complains were dealt with would result in employees experiencing higher levels of job control and in turn well-being. Second, clarifying the performance criteria and regular team briefings would lead to improved feedback and well-being among staff.

Bond et al. (2008) provided an example of how the intervention participatory process may be a mechanism. Bond et al. (2008) argued that as employees jointly decide on how to make changes to the way work is organized, designed and managed this lead to improvements in job control and employee well-being. A limitation of both these studies is that although both studies formulated the mechanisms these were not tested empirically, for example the extent to which employees took over administrative roles was not measured and not linked to outcomes. Interestingly, the two studies both included job control as an outcome and argued that increased control could be either the result of the process or the content of the intervention. This lack of clarity calls for studies that empirically examine and measure both types of mechanism to determine whether job control is triggered by the intervention's content or process.

In a qualitative study among Danish postal service workers. Nielsen et al. (2014) explored a tailored questionnaire as an intervention tool. Based on interviews with managers and employees, a context-specific questionnaire was developed. To capture the cognitive appraisal of employees (Lazarus and Folkman, 1984), response categories were anchored around the appraisal of an aspect of the work environment being a problem or positive. Frequencies were fed back to participants and odds ratios were used to link working

conditions to employee health and well-being outcomes. The mechanisms underlying the questionnaire was that 1) it enabled collective sensemaking of the results of the tailored questionnaire and the problems experienced by the group, 2) it facilitated prioritization because participants could see how many employees reported an aspect of the work environment to be a problem and how strongly this aspect was linked to outcomes, 3) it created ownership among participants because they felt management were interested in their problems and their perspectives, and 4) it facilitated action planning because it pointed directly to issues specific to postal service workers, e.g. employees' opportunities to be heard in connection with changes in the postal routes.

Understanding context in organizational interventions

Interventions work in different ways for different people depending on the context within which they find themselves, and thus whether and how the mechanisms trigger outcomes depend on the context (Greenhalgh et al., 2015). The two aforementioned studies by Higgins et al. (2015) and Goodridge et al. (2015) offer some useful examples on context that may also apply to organizational interventions.

The contextual factors identified by Higgins et al. (2015) included an organizational climate where senior management had good relations with line managers and the trade unions, communication between line managers and employees was respectful, line managers had adequate training to deal with employees on sick leave and where financial pressures motivated line managers to develop flexible return to work practices, such a flexitime working.

In the Goodridge et al. (2015) study, the contextual factors that enabled the mechanisms that triggered outcomes were also identified in interviews. Goodridge et al. (2015) suggested that in regions 1) where Lean was poorly integrated and aligned with other initiatives, 2) where leaders who were responsible for Lean implementation perceived a lack

of centralized support to implement Lean, 3) where central authorities did not support adjustment to local context, 4) where measurement data on which to base quality improvement processes are not readily available or of poor quality, and 5) where leaders feared they could not make mistakes without repercussions, Lean would not be implemented.

Exploring context in organizational interventions

Pawson and Tilley (1997) defined context as the conditions that interventions find themselves in. Macfarlane et al. (2011) analysed the context at four levels: The individual (e.g. values, roles, and knowledge), the interpersonal (e.g. communication, collaboration and networks), the institutional (e.g. informal rules, organizational culture, leadership and regulations), and the infrastructural (e.g. political support). The model put forward by Nielsen and Randall (2015) only considered the individual level and the organizational level, incorporating the group, the institutional and the infrastructural level into the organizational context.

In recent models of organizational intervention evaluation, a distinction between omnibus and discrete context has been suggested (Nielsen and Abildgaard, 2013; Nielsen and Randall, 2013). Essentially, it could be argued that the omnibus context concerns the maturity of the organization in terms of readiness for change, existing working conditions, and health and well-being of employees, while the discrete context concerns the concurrent changes taking place during the intervention period. This distinction is useful in realist evaluation. Traditionally realist evaluation has understood the context as the setting in which the intervention takes place, e.g. a postal service where existing levels of employee health and well-being point to the need for intervention. This approach has been criticized for being too limited in that contexts are dynamic and thus the focus on a pre-existing and stable context does not capture the complexity of context (Dahler-Larsen, 2001). It could be argued that the omnibus context captures the stable, cultural aspects whereas the discrete context captures the

dynamic context. Examples of the omnibus context are found in the Goodridge et al. (2015) and Higgins et al. (2015) studies.

Examples of omnibus context at the individual level have been reported in organizational interventions. Bond et al. (2008) found that an organizational intervention had positive effects on mental health and absence rates for employees high in psychological flexibility. Readiness for change or openness to the intervention has also been found to be important: Participants need to acknowledge problems in the current situation, see the need for change, believe the intervention will have the desired effect and be motivated to follow the requirements for behavioural change made by the intervention (Nytrø et al., 2000). For example, Cunningham et al. (2002) found that the degree to which employees welcomed change, and felt they could effectively handle the change, predicted the degree to which they participated in change activities and felt that they themselves had made a significant contribution to the change. Employees' resources are also an important contextual factor; Nielsen et al. (2006) found employees with little formal education found it challenging to engage in participatory processes. In their review, Lyubomirsky et al. (2005) concluded that individuals high in affective well-being were more proactive and creative, and engaged in more activities and problem solving than their counterparts with lower levels of affective well-being and thus high well-being levels may facilitate the participatory process.

At the group level, in groups where employees were not used to communicate or solve problems at work, participation in the intervention process was reported as difficult and it found that insight about problems did not lead to willingness or ability to make things happen (Mikkelsen et al., 2000). At the institutional level, poor pre-intervention working conditions and well-being have been shown to limit implementation processes (Taris et al., 2003).

Another body of research suggests a ceiling effect. Nielsen et al. (2006) found that in an intervention group that had, prior to the intervention, worked with organizational development, few improvements were observed as a result of an intervention, and Nielsen and Randall (2012) found that high levels of pre-intervention social support were related to few changes in existing working procedures in an intervention aiming to improve social support. These studies suggest it is important to evaluate the “maturity” of the organization in order to understand how the organizational omnibus context has influenced the intervention (Taris et al., 2003).

In organizational intervention research, concurrent changes have often been described as “disturbing” the intervention and as “unnecessary noise” in organizational intervention studies (Nielsen et al., 2006; Mikkelsen and Saksvik, 1999). Discrete events both at the intra-organizational level, for example the introduction of conflicting initiatives (Nielsen et al., 2006) or of new technology (Saksvik, et al., 2002), or at the extra-organizational level, for example economic recession leading to layoffs within the organization (Landsbergis and Vivona-Vaughan, 1995; Mikkelsen and Saksvik, 1999) have been found to impact intervention outcomes negatively. In today’s globalized economy, organizations constantly reorganize and restructure in order to adapt to the demands of their environment (Grant, 2007) and transferring the realist evaluation thinking to organizational intervention research would enable the analysis of the impact of changing organizations on the outcomes of organizational interventions. We need to understand how the dynamic context may act as a barrier or a facilitator that may facilitate certain mechanisms (Noblet and LaMontagne, 2009). For example, if a new line manager is appointed, employees may not be familiar with him or her and therefore may not feel confident sharing their views and ideas openly during the participatory process and therefore the mechanisms of the participatory process may not be facilitated.

Analyzing organizational interventions using CMO-configurations

As mentioned, realist evaluation has rarely been used in organizational intervention research and many studies of organizational interventions study either only the context or the mechanisms without explicitly using realist evaluation as their framework. A central tenet of realist evaluation is that context and mechanisms must be linked to outcomes and not explored separately (Pawson, 2013). In this section, we present a few cases of existing research where CMO-configurations can be applied to understand the intervention's outcomes.

Nielsen and Daniels (2012) examined the effects of training newly appointed team leaders in how to implement and manage teams during a time of team implementation. The underlying mechanism was that through training leaders would come to see team implementation as a positive challenge to develop in their jobs rather than a threat to their status based on cognitive appraisal theory (Lazarus and Folkman, 1984). Nielsen and Randall (2012) found that training leaders (mechanism) lead to improved leader well-being (outcome), but only in circumstances where employees reported being ready to work in teams (omnibus context). Where employees did not report being ready to work in teams, trained leaders experienced poorer well-being compared to leaders who had not received training. This multi-level intervention (team implementation and leader training) provides important insights in what works for whom in which circumstances. Nielsen and Daniels (2012) thus operationalized mechanisms as mediators (a component of the programme) that intervene between baseline and follow-up to influence the outcome, and context as moderators (which represent a relationship between intervention components that is enabled or conditioned by a third factor).

In another study, Nielsen and Randall (2012) explored the outcomes of a teamwork intervention aimed at improving employee job well-being. Nielsen and Randall (2012)

suggested that two mechanisms would trigger improvements in employee well-being. First, the participatory process, in which employees and managers went through a joint problem solving approach would improve autonomy and social support and thereby employee well-being (defined as job satisfaction and affective well-being). Second, changes in work procedures to reflect teamwork would increase employees' opportunities to make independent decisions on how to do their job in a collaborative fashion would increase autonomy and social support and therefore also well-being. Nielsen and Randall (2012) also explored the extent to which the organizational context, i.e. pre-existing working conditions (social support, autonomy and well-being) facilitated the two mechanisms. In a structural equation model, Nielsen and Randall (2012) found that pre-existing levels of affective well-being and social support (omnibus context) triggered the changes in procedures (mechanism) which led to improvements autonomy, affective well-being and job satisfaction. They also found that pre-existing levels of autonomy and job satisfaction (omnibus context) enabled the participatory process (mechanism) which led to improvements in social support and job satisfaction.

Although not discussed by Nielsen and Randall (2012), we can identify two separate CMO-configurations: First, C1 (among employees who have high levels of affective well-being and social support) facilitated M1 (changes in procedures to implement teamwork are introduced) which leads to O1 (leading to increased autonomy and in turn affective well-being and job satisfaction). Second, C2 (pre-existing levels of autonomy and job satisfaction) triggers M2 (participatory intervention processes) which lead to the outcome of job satisfaction. Transferring the understanding obtained through these two CMO-configurations to the design of future organizational interventions it would be suggested that the mechanism of a participatory process would be facilitated when employees experienced high levels of pre-intervention autonomy, i.e. have previous experience engaging in independent decision

making and who are satisfied with their jobs. Employees who are less satisfied are unlikely to engage in a participatory process with leaders and colleagues. The participatory intervention process could be expected to bring about increased levels of social support and job satisfaction, but not affective well-being. The mechanism of changes in work procedures would only be triggered if groups of employees had high levels of affective well-being, i.e. the energy to change working procedures and where they felt supported by their colleagues prior to the intervention (omnibus context). In these circumstances the mechanism of changes in procedures could be expected to result in both affective well-being and job satisfaction. It could be argued that this particular study examined both content and process mechanisms. The changes in procedures relate to the teamwork intervention, i.e. the content of the intervention as it was measured whether teamwork practices were introduced. The participation mechanism on the other hand relate to the process as it concerned how employees engaged in the participatory process employed to implement teamwork.

Synthesizing the CMOs of organizational interventions: Realist synthesis

In the beginning of this critical essay, we argued that meta-analyses and considerations of RCTs as the gold standard offer limited value in providing knowledge on internal and external validity but that such types of validity may be better captured studying what specific elements of the intervention work, how and why and in which circumstances. Synthesizing the literature reviews exploring what works for whom in which circumstances in health care, Straus et al. (2016) found that alternative review methods have gained popularity: In 2000, ten alternative reviews had been published, but in 2013 the number has risen to 300 alternative reviews. In organizational intervention research, however, the systematic reviews and meta-analyses still dominate (Bhui et al., 2012).

The question remains: How can we synthesize learning and offer recommendations on how to design and implement future organizational interventions? We argue that realist

synthesis which builds on the principles of realist evaluation can be successfully transferred to organizational intervention research. We thus suggest that reviews should focus on synthesizing CMO-configurations in realist syntheses.

In realist synthesis the unit of enquiry is the programme theory and the aim is to develop a series of research questions about what works for whom in which circumstances that can be supported by the existing research – or lead to amendments of the programme theory. The programme theory can be used to develop and evaluate future interventions and provide recommendations on how to conduct organizational interventions in particular settings, for example, if change agents aim to improve employee health and well-being among employee groups A and B, then they should implement x and be aware that contextual factors 1 and 2 need to be present. If you are targeting employee groups C and D, you are better off implementing z and y and be aware of the pitfalls of 2 and 3.

Realist syntheses in organizational intervention research would thus synthesize the mechanisms of the content and the process of the intervention and identify how participants' reactions (both cognitive and behavioural) shape the interventions' outcomes and in which contexts these mechanisms may be triggered. For example, is the intervention only likely to be successful in large organizations where existing health and safety management structures are well established or can elements of the intervention be transferred to interventions in Small and Medium-sized Enterprises?

Potential research questions in a realist synthesis of organizational interventions could include: How and why do organizational interventions work? That is: What are the programme theories for making changes to the way work is organized, designed and managed and the processes by which change is brought about? More specific research questions could include: Which theories can be used to explain process mechanisms? Which theories can be used to explain content mechanisms? Which discrete context factors trigger a) process and b)

content mechanisms? Which omnibus context factors trigger a) process and b) content mechanisms? What are the most important Outcomes brought about by these CM configurations? What areas for further research can be identified to inform and develop sustainable organizational interventions?

Conclusion

In the present critical essay, we have argued that realist evaluation offers great opportunities to unpack the black box of organizational interventions that are generally recommended by official bodies such as EU-OSHA, ILO and WHO. After illustrating the limitations of RCTs and meta-analyses when evaluating the effectiveness of organizational interventions, we have put forward realistic evaluation and realistic synthesis as more suitable approaches to understand the outcomes of an intervention and the reasons (i.e., context and mechanisms) behind them. Acknowledging the challenges of current realist evaluation in determining which factors can be defined as context and mechanisms, we propose an operationalization of context into omnibus and discrete context and mechanisms concerning the content and the process of interventions. The knowledge and learning obtained from realist evaluation can be used by a wide range of stakeholders. First, this approach can be used by researchers to develop and test theories of how, why organizational interventions work and for whom and in which circumstances. It can thus be used to develop organizational intervention theory and test the extent to which important mechanisms related to the content or the process of the intervention may trigger certain outcomes and in which circumstances. Second, realist evaluation can be used by organizations, HR and occupational health professionals who wish to improve employee health and well-being through changing the way work is organized, designed and managed. Developing programme theories of how and why specific intervention components may help organizational members to plan interventions better and have realistic expectations of which outcomes can be achieved. It also facilitates

the transfer of one intervention to another. For example, in a large multi-national company the analysis of what works for whom can shed light on the extent to which an intervention needs tailoring to a particular national context and how it should be tailored. Finally, realist evaluation can help inform policy at the national and intervention level and can lead to the development of methods, tools and guidelines that policy bodies can recommend organizations use in their efforts to improve employee health and well-being.

References

- Bhui K, Dinos S, Stansfeld S, and White P. (2012) A synthesis of the evidence for managing stress at work: a review of the reviews reporting on anxiety, depression, and absenteeism. *Journal of Environmental and Public Health* 2012: 515-874.
- Bond FW, Flaxman PE and Bunce D (2008) The influence of psychological flexibility on work redesign: Mediated moderation of a work reorganization intervention. *Journal of Applied Psychology* 93(3): 645-54.
- Borenstein M, Hedges LV, Higgins JPT and Rothstein HR (2009). *Introduction to meta-analysis*. Chichester: Wiley.
- Dahler-Larsen P (2001) From Programme Theory to Constructivism On Tragic, Magic and Competing Programmes. *Evaluation*, 7(3), 331-49.
- Cunningham CE, Woodward CA, Shannon HS, MacIntosh J, Lendrum B, Rosenbloom D et al. (2002) Readiness for organizational change: A longitudinal study of workplace, psychological and behavioural correlates. *Journal of Occupational and Organizational Psychology*, 75, 377-392.
- Dickersin K. (2005). Publication bias: Recognizing the problem, understanding its origins and scope, and preventing harm. In: Rothstein HR, Sutton AJ, Borenstein M (Eds.), *Publication bias in meta-analysis: Prevention, assessment, and adjustments*. West Sussex, England: John Wiley, 11-33.
- Duval SJ and Tweedie RL (2000) Trim and fill: A simple funnel plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* 56: 455-463.
- ETUC (2004). *Framework agreement on work-related stress*. Brussels: European Trade Union Confederation.
- EU-OSHA (2010). European Agency for Safety and Health at Work, *European Survey of Enterprises on New and Emerging Risks*, 2010. Available at: www.esener.eu
- ILO (2001). *Guidelines on occupational safety and health management systems*. Geneva: International Labor Office.
- Goodridge D, Westhorp G, Rotter T, Dobson R, Bath B (2015) Lean and leadership practices: development of an initial realist program theory. *BMC Health Services Research* 15(362), 1-15.
- Grant AM (2007) Relational job design and the motivation to make a prosocial difference. *Academy of Management Review* 32(2): 393-417.
- Greenhalgh T, Wong G, Jagosh J, Greenhalgh J, Manzano A, Westhorp G, Pawson R (2015) Protocol-the RAMESES II study: Developing guidance and reporting standards for realist evaluation, *BMJ Open*, 5.8
- Guyatt G, Sackett D, Sinclair J, Hayward R, Cook D, and Cook R (1995) User's guides to the medical literature. IX. A method for grading health care recommendations. *Journal of the American Medical Association* 274: 1800-1804.
- Higgins A, O'Halloran P, Porter S (2012) Management of long term sickness absence: a systematic realist review. *Journal of Occupational Rehabilitation* 22(3):322-332.
- Holman D and Axtell C (2015) Can job redesign interventions influence a broad range of employee outcomes by changing multiple job characteristics? A quasi-experimental study. *Journal of Occupational Health Psychology*, doi.org/10.1037/a0039962.
- Hoyle, R. H., Harris, M. J., & Judd, C. M. (2001). *Research methods in social relations* (7th ed.). Belmont, CA: Wadsworth.

- Karlan D., Goldberg N. and Copestake J. (2009) Randomized control trials are the best way to measure impact of microfinance programmes and improve microfinance product designs. *Enterprise Development and Microfinance* 20(3): 167-176.
- Lacouture, A, Breton, E, Guichard, A and Ridde, V (2015). The concept of mechanism from a realist approach: a scoping review to facilitate its operationalization in public health program evaluation. *Implementation Science* 10: 153-163.
- Landsbergis P and Vivona-Vaughan E (1995) Evaluation of an occupational stress intervention in a public agency. *Journal of Organizational Behavior* 16: 29-48.
- Lazarus R and Folkman S (1992) *Stress, Appraisal and Coping*. New York: Springer Publications.
- Lyubomirsky S, King L and Diener E (2005) The benefits of frequent positive affect: Does happiness lead to success? *Psychological Bulletin* 131(6): 803-855.
- Macfarlane F, Greenhalgh T, Humphrey C, Hughes J, Butler C, Pawson R. (2011) A new workforce in the making? A case study of strategic human resource management in a whole-system change effort in healthcare. *Journal of Health Organization Management* 25(1):55-72.
- Malchaire, JB (2004) The SOBANE risk management strategy and the De'paris method for the participatory screening of the risks. *International Archives of Occupational Environment and Health* 77: 443-50.
- Manzano A and Pawson R (2014) Evaluating consent to deceased organ donation: The need for programme theory. *Journal of Healthcare Organization & Management* 28(3): 366-385.
- Marchal B, van Belle S, van Olmen J, Hoérée T, Kegels G (2012) Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research. *Evaluation* 18(2):192-212.
- Mikkelsen A, Saksvik PØ and Landsbergis P (2000) The impact of a participatory organizational intervention on job stress in community health care institutions. *Work & Stress* 14(2): 156-70.
- Mikkelsen A and Saksvik PØ (1999) Impact of a participatory organizational intervention on job characteristics and job stress. *International Journal of Health Services* 29(4): 871-93.
- Nabe-Nielsen K, Persson R, Nielsen K, Olsen O, Carneiro IG and Garde AH (2015) Perspective on randomization and readiness for change in a workplace intervention study. In Karanika-Murray, M., and Biron, C. *Derailed organizational stress and well-being interventions: Confessions of failure and solutions for success* (pp.201-208). Springer. Dordrecht.
- Newcomer KE, Hatry HP and Wholey JS (2010) Planning and Designing Useful Evaluations. In JS Wholey, HP Hatry, and KE Newcomer. (3rd ed.), *Handbook of practical program evaluation* (pp. 5-29). San Francisco: Jossey-Bass.
- Nielsen K (2013a) How can we make organizational interventions work? Employees and line managers as actively crafting interventions. *Human Relations*. 66, 1029-1050.
- Nielsen K (2013b) Managing psychosocial risks in Denmark: research and policy initiatives. In J.M. Peiro and C. Molina. *International yearbook of psychosocial risk prevention and quality of life at work. Evaluation development of psychosocial risks in Europe. The state of scientific research and institutional experiences*. Spain, Jaen: Secretary of Labour Health and Environment UGT-CEC.

- Nielsen K and Abildgaard JS (2013) Organizational interventions: A research-based framework for the evaluation of both process and effects. *Work & Stress* 27: 278-297.
- Nielsen K, Abildgaard JS and Daniels K (2014) Putting context into organizational intervention design: Using tailored questionnaires to measure initiatives for worker well-being. *Human Relations* 67(2): 1537-1560.
- Nielsen K, Fredslund H, Christensen KB and Albertsen K (2006) Success or failure? Interpreting and understanding the impact of interventions in four similar worksites. *Work & Stress* 20(3): 272-87.
- Nielsen K and Randall R (2009) Managers' active support when implementing teams: The impact on employee well-being. *Applied Psychology: Health and Well-being* 1(3): 374-390.
- Nielsen K and Randall R (2012) The Importance of Employee Participation and Perception of Changes in Procedures in a Teamworking Intervention. *Work & Stress* 29(1): 111.
- Nielsen K and Randall R (2013) Opening the black box: A framework for evaluating organizational-level occupational health interventions. *European Journal of Work and Organizational Psychology* 22(5): 601-617.
- Nielsen K and Randall R (2015) Addressing the fit of planned interventions to the organizational context. In Karanika-Murray, M., and Biron, C. *Derailed organizational stress and well-being interventions: Confessions of failure and solutions for success* (pp. 107-118).Springer. Dordrecht.
- Nielsen K, Randall R, Holten AL and Rial González E (2010) Conducting Organizational-level Occupational Health Interventions: What Works? *Work & Stress* 24(3): 234-59.
- Noblet A and LaMontagne A (2009) The challenges of developing, implementing, and evaluating interventions. In: Cartwright S, Cooper CL (eds) *The Oxford Handbook of Organizational Well-being*. Oxford: Oxford University Press, 466-496.
- Nytrø K, Saksvik PØ, Mikkelsen A, Bohle P and Quinlan M (2000) An appraisal of key factors in the implementation of occupational stress interventions. *Work & Stress* 14(3): 213-225.
- Pawson R (2006) *Evidence-based policy: A realistic perspective*. Los Angeles, CA and London: Sage.
- Pawson R (2013) *The science of evaluation: a realist manifesto*. Los Angeles, CA and London: Sage.
- Pawson R and Tilley N (1997) *Realistic evaluation*: Los Angeles, CA and London: Sage.
- Richardson KM and Rothstein HR (2008) Effects of occupational stress management intervention programs: a meta-analysis. *Journal of Occupational Health Psychology* 13: 69-93.
- Rothstein HR, Sutton AJ and Borenstein M (2005) Publication bias in meta-analysis. In: Rothstein HR, Sutton AJ, Borenstein M (Eds.) *Publication bias in meta-analysis: Prevention, assessment, and adjustments*. West Sussex, England: John Wiley, 1-7.
- Rousseau DM, Manning J and Denyer D (2008) Evidence in management and organizational science: Assembling the full weight of scientific knowledge through synthesis. *The Academy of Management Annals* 2: 475-515.
- Saksvik PØ, Nytrø K, Dahl-Jørgensen C and Mikkelsen A (2002) A process evaluation of individual and organizational occupational stress and health interventions. *Work & Stress* 16(1): 37-57.
- Shaugnessy JJ, Zechmeister EB and Zechmeister JS (2006) *Research methods in psychology* (7th ed.) New York, NY: McGraw-Hill.

- Schmidt FL and Hunter JE (2015) *Methods of meta-analysis: Correcting error and bias in research synthesis* (3rd Ed.). Thousand Oaks, CA: Sage.
- Sterne JA and Egger M (2005) Regression methods to detect publication bias and other bias in meta-analysis. In: Rothstein HR, Sutton AJ, Borenstein M (Eds.), *Publication bias in meta-analysis: Prevention, assessment, and adjustments*. West Sussex, England: John Wiley, 99-110.
- Sterne JA, Gavaghan D and Egger M (2005) The funnel plot. In: Rothstein HR, Sutton AJ, Borenstein M (Eds.) *Publication bias in meta-analysis: Prevention, assessment, and adjustments*. West Sussex, England: John Wiley, 75-98.
- Straus SE, Kastner M, Soobiah C, Antony J and Tricco AC (2016). Engaging researchers on developing, using, and improving knowledge synthesis methods: introduction to a series of articles describing the results of a scoping review on emerging knowledge synthesis methods. *Journal of Clinical Epidemiology*.
[doi:10.1016/j.jclinepi.2016.01.031](https://doi.org/10.1016/j.jclinepi.2016.01.031)
- Taris TW, Kompier M, Geurts S, Schreurs P, Shaufeli P, de Boer E, Sepmeijer K and Watez C (2003) Stress management interventions in the Dutch domiciliary care sector: Findings from 81 organisations. *International Journal of Stress Management* 10(4): 297-325.
- Van der Klink, J. Blonk,R., Schene A and van Dijk F (2001) The benefits of interventions for work-related stress. *American Journal of Public Health* 91: 270–276.
- WHO (1995). *Global strategy on occupational health for all: the way to health at work*. Geneva: World Health Organization; 1995.