BMJ Open

Responsive evaluation: an innovative evaluation methodology for workplace health promotion interventions

Hanneke van Heijster ,[1] Jantien van Berkel,[1,2] Cécile RL Boot,[3] Tineke Abma,[4] Emely de Vet[1]

To cite: van Heijster H, van Berkel J, Boot CRL, *et al*. Responsive evaluation: an innovative evaluation methodology for workplace health promotion interventions. *BMJ Open* 2022;**12**:e062320. doi:10.1136/bmjopen-2022-062320

► Prepublication history for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2022-062320).

Received 23 February 2022
Accepted 27 November 2022

[1]Department of Social Sciences, Chair Group Consumption and Healthy Lifestyles, Wageningen University & Research, Wageningen, The Netherlands
[2]Department of Interdisciplinary Social Science: Public Health, Utrecht University, Utrecht, The Netherlands
[3]Department of Public and Occupational Health, Amsterdam Public Health research institute, Amsterdam UMC, VU University, Amsterdam, The Netherlands
[4]Department of Public Health, Leiden University Medical Center, Leiden, The Netherlands

Correspondence to
Hanneke van Heijster;
hanneke.vanheijster@wur.nl

ABSTRACT

Background: Workplace health promotion (WHP) interventions have limited effects on the health of employees with low socioeconomic position (SEP). This paper argues that this limited effectiveness can be partly explained by the methodology applied to evaluate the intervention, often a randomised controlled trial (RCT). Frequently, the desired outcomes of traditional evaluations may not match employees'—and in particular employees with low SEP—needs and lifeworld. Furthermore, traditional evaluation methodologies do not function well in work settings characterised by change resulting from internal and external developments.

Objective: In this communication, responsive evaluation is proposed as an alternative approach to evaluating WHP interventions. Responsive evaluation's potential added value for WHP interventions for employees with low SEP in particular is described, as well as how the methodology differs from RCTs. The paper also elaborates on the different scientific philosophies underpinning the two methodologies as this allows researchers to judge the suitability and quality of responsive evaluation in light of the corresponding criteria for good science.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ This study presents an innovative evaluation method in the field of workplace health promotion, with empirical examples and explanations of underlying concepts.
⇒ This study informs readers about various methodologies and about the situations in which each methodology might be more or less suitable.
⇒ This study focuses on one innovative methodology, although others exist.
⇒ It was decided to focus on one methodology to be able to provide sufficient examples and explain underlying concepts.

INTRODUCTION

Although the workplace is a promising setting for workplace health promotion (WHP), the effects of WHP interventions have been limited, as shown in several individual participant data meta-analyses performed in the last decade.[1–5] Theoretically, the workplace is a promising setting for health promotion because existing physical and social structures at work, such as the physical environment and group norms, provide a basis on which to build.[6 7] WHP is of particular relevance to employees with low socioeconomic position (SEP) because these employees generally have poorer health than those with high SEP.[8] However, employees with low SEP participate less frequently in WHP interventions,[9 10] and if they do the effectiveness on their health is limited.[1–3] Therefore, the question arises as to why WHP interventions are not living up to their potential.

Various explanations have been offered for the limited effectiveness of WHP interventions. First, the limited effectiveness could result from inappropriate interventions, for example based on the wrong theories (theory failure). Second, poor implementation could explain the intervention's limited effectiveness, for example when participation in the intervention is low or when participants drop out before the end of the intervention (programme failure).[11 12] Although these explanations are legitimate, we argue that part of the explanation can be found in the methodologies used to evaluate WHP interventions. WHP intervention evaluation has already been a topic of discussion. Several articles have questioned the suitability of the randomised controlled trial (RCT), one of the methodologies most used in the occupational health field to evaluate WHP interventions.[13–15] One reason for questioning an RCT's suitability for evaluation in the work setting is that the setting is subject to change.[16] Changes can be large scale on the socioeconomic, political, technological and demographic level (eg, legislation, ageing population), macro at the industry or company level (eg, downsizing, outsourcing), or micro at the organisation level (eg, workload, participation, support).[16] These

inevitable changes make it impossible to fully control the work setting, which is a prerequisite to an RCT's internal validity.[16 17] Other challenges for RCTs relate to the desires of organisations, which may not agree with the randomisation of employees and a control group because this means that some employees will not receive the intervention. Moreover, the organisation may want to make changes to the intervention protocol, for example due to developments in the context (eg, the intervention has to be postponed due to the COVID-19 pandemic).[13] These challenges hamper RCTs' desired execution, thereby masking the intervention's impact.

In addition, traditional evaluation methodologies may not pay sufficient attention to employees' subjective experiences, which are relevant to ascertaining what they find important in WHP. Traditional evaluation methodologies often focus on outcomes such as body mass index (BMI), lifestyle behaviours or organisation-related outcomes such as productivity,[18] as these are outcomes that fit the typical design of an RCT in which changes in certain measurable outcomes are evaluated. Outcomes such as BMI and lifestyle behaviours are relevant outcomes from an epidemiological perspective, but often are not the top priority in employees with low SEP daily lives.[10 19] In other words, the outcomes of traditional evaluations may not match employees with low SEP lifeworld. The lifeworld is a person's or a group's background of shared assumptions, meanings and understandings about the world,[20] which differ between people and groups.[21] The lifeworld of employees with low SEP probably differs from high SEP employees' lifeworld because these employees work under different working conditions,[7] and their norms and values, for example with regard to health, may differ.[22] Each person's lifeworld influences what that person considers important. Hence, it presumably also encompasses what that person considers a relevant outcome of a WHP intervention. Therefore, the outcomes in WHP intervention evaluations should be defined in terms of employees with low SEP lifeworld.

Alternative evaluation methodologies for WHP interventions have been proposed, such as the cluster RCT and stepped-wedge design[18] and observational (non-randomised) design.[14] Although these alternatives tackle some of the challenges regarding randomisation, control group and intervention, they do not address the possible mismatch between what is measured and employees with low SEP lifeworld. Furthermore, these alternative evaluation methods still face challenges when changes occur in and outside the work setting. Therefore, other evaluation methodologies should also be considered.

This paper proposes responsive evaluation as an approach to evaluating WHP interventions. This approach was introduced by Stake[23] for evaluating educational programmes[23 24] and extended by Guba and Lincoln.[25] The methodology was further developed and introduced to the public health field by Abma,[26] who added more interactive and participatory elements. This paper aims to provide an extensive description of responsive evaluation and its potential added value for WHP evaluation in light of the changeability of the work setting and the need for WHP evaluation to take employees with low SEP lifeworld into account. Consequently, the aims, methods and type of evidence used in responsive evaluation are described and compared with these elements in an RCT, thereby aiming to inform researchers about the differences between both methodologies, including diverging underlying scientific philosophies. These philosophies are described, highlighting some characteristics of the scientific philosophy underpinning responsive evaluation that might be of particular interest for WHP evaluation. In addition, the role played by dialogue and values in responsive evaluation is explained, as these are two typical elements of responsive evaluation.

Throughout this commentary, examples are given from two recent WHP responsive evaluations performed over 2 years in two Dutch organisations with employees with low SEP: a harbour service provider (2018–2021) and a sheltered workplace (2019–2021). The evaluation papers have been/will be published elsewhere.[27]

AIMS, METHODS AND TYPE OF EVIDENCE

In the following sections, the aims, methods and types of evidence in RCTs and responsive evaluation are described. A summary of important differences is provided in table 1.

Aims

The RCT's aim is to examine a causal relationship between an intervention or manipulation and an observable change in a predefined outcome of interest.[17] The RCT has become the gold standard in evaluative medical research because it allows the effect of certain medical therapies to be examined.[13] If an RCT is used in the WHP field, the aim is to examine a causal relationship between a WHP intervention and a particular health outcome or organisational outcomes such as productivity or job satisfaction. If a causal relationship between intervention and effect is found, it can be concluded that the intervention is effective under the circumstances in which the RCT was performed.

Responsive evaluation, on the other hand, starts from the notion that social reality is too complex to detect clear cause–effect relations between an intervention and an outcome.[28] However, this evaluation methodology has other aims that are of interest. Responsive evaluation aims to improve interventions by aligning them more closely with practice, for example the work setting. To achieve this, stakeholders (eg, employees, Human Resources Management (HRM), management) are involved in defining relevant themes and outcomes of the interventions.[26] Their concerns and suggestions are the priority and the starting point of the evaluation. In the harbour service provider and the sheltered workplace, operational employees, supervisors and management were asked what they consider important outcomes of an intervention. By bringing together stakeholders' perspectives, responsive

Table 1 Important differences between traditional methodologies and responsive evaluation

Differences	Traditional methodologies (e.g., randomised controlled trial)	Responsive evaluation
Aims	Examine causality between the intervention and the outcome	Better match interventions to a target group's lifeworld and evaluate change
Design and methods	Rigid design Measurement	Emergent design Qualitative methods or mixed methods
Type of evidence	Statistical evidence	Argumentative evidence
Philosophy underlying the evaluation	Positivism, postpositivism	Social constructivism, hermeneutics, interpretative approaches
View on reality	Reality is external and can be observed from outside	Reality is socially constructed and can be understood by participating in it
Researcher's attitude	Objective (observes from outside)	Participating observer (observes from inside and through interaction)
Role of values	Values play a role in deciding the direction of research and judging methodology on ethics	Specific values* are the main drivers of research and underlie the approach
Favourable circumstances for the different evaluation designs	Controllable settings Target population easy to reach and include	Settings susceptible to change Target population difficult to reach and include

*Empowerment, social inclusion, emancipation and epistemic justice.

evaluation facilitates mutual learning and greater understanding and acting on these.[24 25 29] Special attention is paid to people who are generally not involved in knowledge creation, as is often the case for employees with low SEP in WHP intervention evaluations.[10] Involving employees and taking their experiences as a starting point for evaluation can shape the evaluation from the perspective of their lifeworld.

Responsive evaluation aims to evaluate the impact of an intervention. This can take the form of experiences with the intervention or mapping concrete actions or changes that take place during or after the intervention. These experiences and changes can then be compared with participants' desired outcomes in order to monitor the relevance of the changes for the stakeholders. With the aim of improving an intervention, changes in the organisational context can be taken into account to adapt the intervention and evaluation, if this will improve the intervention's relevance for employees. For example, after an early COVID-19 outbreak, an inventory was made in the sheltered workplace, which was located in a heavily affected region in the Netherlands, of whether the impact of the virus on employees' (mental) health should be addressed. This was not the case, but responsive evaluation would have allowed a slight shift in focus if this had been necessary to maintain the intervention's relevance.

Methods

Methods used in an RCT are mostly quantitative, such as validated questionnaires or bodily measurements of biomedical risk factors of health such as BMI. Measurements are usually performed before and after the intervention. Through randomisation, other confounding factors that may influence these outcomes are assumed to be equally distributed over the groups. This allows conclusive statements to be made about an intervention's effects on the outcome of interest and consequently about the internal validity of the outcomes.[17]

In responsive evaluation, both qualitative and quantitative methods can be used. However, qualitative methods are important because they are appropriate for gaining insights into the experiences and complexity of the social world,[26] thereby making qualitative methods suitable for gaining insights into employees with low SEP lifeworld. Participant observations (ie, participating at the workplace for several days) and interviews led by employees' stories are examples of possible qualitative methods that can be used in responsive evaluation. If possible, qualitative methods are combined with quantitative methods to enhance the validity of findings (data triangulation).[30] For example, in the responsive evaluation in the harbour service provider, qualitative methods (interviews and participant observations) revealed that high workloads, mental health and burnout were important issues for employees. These issues were confirmed by quantitative data consisting of periodical medical evaluations.

Different methods result in different types of outcomes. Whereas RCTs provide information about measurable outcomes, the impact of responsive evaluation is often described as a learning process. In this learning process, understanding is gained on issues that are important to stakeholders, leading to changes in understanding, attitude, and sometimes organisational and/or individual behaviour.[26 31 32] In the responsive evaluation of the harbour service provider, changes were observed on four

levels: case, individual, team and organisational.[27] These changes included the implementation of a programme to enhance mutual understanding between departments, learning from one another and management's improved understanding of employees' issues.

Responsive evaluation takes place continuously, rather than at fixed junctures. This means that the intervention is evaluated throughout the evaluation period, and changes can be found at any stage of the project.[33 34] This allows an understanding to be obtained of the experiences with the intervention early in the evaluation. In addition, even when major changes in the organisational context take place (eg, the COVID-19 outbreak), the findings in the phases before the major change can be taken into account. Furthermore, due to responsive evaluation's emergent design, the methods can be adapted if this is considered necessary consequent to changes like the COVID-19 outbreak. For example, the final stages of the harbour service provider's responsive evaluation took place during the COVID-19 outbreak in March–April 2020. Due to the high level of sick leave during that period, it was decided to replace the planned final employee interviews with one interview with management (ie, decision-makers) to avoid overwhelming the employees. This was the best option in practice and still provided relevant insights for evaluation. If an RCT evaluation methodology had been applied, the decline in the usefulness of the results would have been greater because RCTs rely on performing evaluations as planned in advance.

Type of evidence

Measurements in a controlled environment can provide statistical evidence in an RCT about whether or not a causal relationship exists between intervention and outcome. In responsive evaluation, changes are substantiated not by statistical but by argumentative evidence. This evidence is qualitative and sometimes also quantitative, together making a plausible argument that certain changes have taken place as a result of the intervention. For example, in the harbour service provider, the evidence consisted of a collection of stakeholders' stories about the perceived changes and communications with the organisation's decision-makers about the changes implemented.

UNDERLYING PHILOSOPHIES

The differences in aims, methods and evidence do not stand alone, but fit in the underlying philosophies in which traditional methodologies such as the RCT and responsive evaluation are embedded (table 1). Although the philosophical embedding of evaluation methodologies can be nuanced, RCT and responsive evaluation are associated with two diverging philosophies frequently criticised by users of the one and the other,[35] namely positivism and social constructivism. These scientific traditions have different views on what good science is and how it should be performed. It may be helpful for researchers considering responsive evaluation, but who are accustomed to working from a positivist tradition, to understand responsive evaluation's underlying philosophy and how it has scientific value in its own right. Moreover, some characteristics, such as the view on reality and the researcher's role in social constructivism, may be especially relevant for WHPs for groups whose lifeworld is little known, such as employees with low SEP. The following section provides a reflection on those characteristics of social constructivism, after first describing how each of these characteristics is interpreted in positivism.

View on reality and how to understand it

The positivist tradition underlying most traditional WHP evaluation methodologies, including the RCT, is originally situated in the natural sciences model. In this model, reality is considered to be external, with properties that can and should be measured through objective methods.[36] In pure positivism, knowledge about this reality is significant only if it is based on objective, value-free observations. However, postpositivism has rejected the idea of objective, merely sensory, observation. From a postpositivist perspective, the world should be studied through measurement and objective methods that are value-neutral and have operationalised indicators.[36 37] In WHP evaluations, operationalised indicators could include, for example, BMI, physical fitness or productivity. Self-reported data can also be used, although these raise questions about bias in the positivist tradition.[38] These standardised measures allow reality to be described objectively, or, in WHP evaluations, health.

Whereas positivism stems from the natural sciences model, social constructivism starts from the belief that this model is inadequate for studying social phenomena.[39] In the social constructivist tradition, which underpins responsive evaluation, reality is considered to be socially constructed. This means that reality is constructed by people who ascribe different meanings to their world.[36] Translated into WHP evaluation, this means that, although health can be measured through operationalised indicators, it is also socially constructed by people with different definitions of what it means to be healthy. From a social constructivist position, meanings are considered to be credible as long as they are understood from the perspective of the people under study.[39] Researchers try to understand the subjective meanings that people give to a certain phenomenon[36 37 39] such as health. Subjective does not mean biased or opinionated, but rather the meaning that something has for the observed human. These subjective meanings allow the people and their behaviour under study to be understood.[39] For example, in the responsive evaluation of the social enterprise, a single woman with a disability explained that she was well aware that quitting smoking would be much better for her health. She had heart problems and was treated for this in the hospital, and her doctor had already urged her to quit several times. However, she explained that it was hard to quit because she enjoyed smoking, especially when she arrived

home alone from work or other activities. Although this may seem irrational from a medical perspective, it shows that knowing is not enough to quit smoking. This helps us to better understand—although not necessarily agree with—this woman's decision with regard to smoking. Understanding the subjective meaning of health for this woman is relevant for WHP because it could lead to the conclusion that education about the disadvantages of smoking is not sufficient to support her health.

The researcher's role

In positivism, the researcher is independent of what is being researched. The researcher observes what is or is not the case, in the third-person attitude.[37] In this attitude, the researcher observes the study object from the outside (rather than from the inside, as is the case in social constructivist research). Evaluation of outcomes through standardised questionnaires and measurements facilitates comparisons with data collected in other settings.[36 40] In the positivist tradition, these comparisons are important for interpreting results and making generalisations.

In social constructivist approaches on the other hand, the researcher adopts the performative attitude, in which the researcher participates in communicative action.[37 39] Unlike in traditional approaches, the researcher has to be in proximity to the people under study. Standardised evaluation methods are less suitable when one is working from the performative attitude because they do not allow space for experiences outside the questionnaire and measurement method. The latter is of particular relevance when researchers are less familiar with the target group under study. Qualitative methods such as interviews and participant observations are more suitable for this purpose, provided they are used in such a way that the participants' experiences are the starting point of the methods. In communicative action, the researcher will be confronted with the so-called non-cognitive claims to reality: the speakers will refer not only to something in the objective world, but also to something in their social world (eg, norms) and personal world (subjective experiences), for example the employee's subjective experience with smoking from the abovementioned example.[21] In this research role, changes in the work context that affect those subjective experiences are not a disruption of the research process, but rather a development that offers the possibility of a better understanding of employees' lifeworld.

THE ROLE OF DIALOGUE AND VALUES IN RESPONSIVE EVALUATION

Dialogue

In the most recent version of responsive evaluation,[25] dialogue plays an important role. In addition to elucidating people's subjective meanings and perspectives, these experiences should be related to one another through dialogue. This argument is based on Gadamer's theory of hermeneutics, in which gaining insight through the interpretation of experiences and opinions, through dialogue, is central.[41] Dialogue facilitates (1) acknowledging the other; (2) being open about one's perspective on reality; and (3) striving for mutual understanding, learning and insight.[42] These outcomes of dialogues are not only helpful in gaining more insight into employees' lifeworld in WHP evaluation because employees share their perspectives on health, but also relevant for WHP evaluation given the many stakeholders involved in WHP.[43 44] In the role of 'Socratic guide' in responsive evaluation, the researcher facilitates dialogue and, through this, learning among the stakeholders.[45] For example, in the sheltered workplace, exchange of perspectives between employees led to the management understanding that the WHP activities were sometimes too challenging for some employees, as a result of which these employees often did not participate.

Values

Responsive evaluation starts from a normative position, working from intrinsic values such as emancipation, social justice and empowerment.[34 46] These values are reflected in some of the characteristics of responsive evaluation. For example, they are reflected in the inclusion in the evaluation of those with the least heard voice, by creating a safe communication climate in which they can speak freely.[47] In the sheltered workplace, a safe communication climate was created through organising dialogues where employees shared their perspectives and ideas on how to improve the existing WHP activities. The sheltered workplace planned to continue these dialogues after the responsive evaluation to ensure that employees' input would be given a permanent place in the design of WHP activities. Including different perspectives and respecting knowledge diversity also reflect the underlying value of epistemic justice.[48] Epistemic justice is another guiding value in responsive evaluation, as it is in other participatory approaches such as participatory health research.[49] To achieve epistemic justice, the researcher has the moral responsibility to create room for all voices and various forms of knowledge, and also to consider the various stakeholders' perspectives as equally relevant.

DISCUSSION

The aim of this paper was to provide an extensive image of responsive evaluation and its potential added value as a different approach to evaluating WHP interventions, in light of the fact that it is suitable for changeable settings such as the workplace and for gaining more insight into employees with low SEP lifeworld. Responsive evaluation's emergent design means that change can be anticipated by making adjustments during the evaluation, if this increases the quality of the evaluation (eg, by enhancing relevance for stakeholders). Moreover, given the principles of social constructivism in which people give their own meanings to their world, the ongoing influence of change on various stakeholders in the work setting,

rather than impeding the evaluation, can be addressed and included in the evaluation. In dialogue—an indispensable element of responsive evaluation—the various meanings that people give to health and WHP can be connected, thereby enabling various perspectives on WHP to be heard, including of those who are generally not involved in WHP design and evaluation. Thus, it can contribute to finding ways to better match interventions to employees with low SEP lifeworld.

However, there are also challenges that should be acknowledged when using responsive evaluation as a methodology for WHP evaluation. The first challenge is related to responsive evaluation's aim to understand how programmes work in a particular context rather than to make generalisations about how programmes work in general.[24] This may be seen as problematic if organisations want to base their decisions about WHP interventions on proof of effectiveness in other organisations.[10 17] However, traditional approaches such as RCTs face similar issues with external validity[35] when programmes work in a particular setting but not in another.[13] The findings of responsive evaluation are transferable rather than generalisable. The rich data mined by responsive evaluation provide a thick description of the work context, stakeholders, circumstances and outcomes of the evaluation. Other workplaces may recognise some or more elements of this thick description and can extrapolate some of the findings to their setting,[25] that is, form an idea of how an intervention could work in their setting. Thus, the findings of a responsive evaluation in one workplace can be translated to other, similar workplaces. The second challenge is related to confirmability, that is, the possibility for other researchers to verify the interpretations of the data.[30] All researchers have biases, for instance due to their (cultural, educational or social) background that can trickle down to the interpretation of data.[50] Researchers should be aware of their biases and reflect on this especially in responsive evaluation, as often a large part of the data is qualitative, of which a part consists of data collected through informal conversations during participant observations. This form of data collection is less controlled like in an interview which is audio-recorded, which requires researchers to keep an audit trail of all data collected and reflect on their biases.[30] All in all, researchers who use responsive evaluation must face the challenge of demonstrating the relevance of findings to other settings in a different way than in traditional evaluation methodologies, as well as putting a lot of efforts in systematically keeping track of all the different forms of data collection.

Strengths and limitations of this communication

A strength of this communication is that it provides examples of two responsive evaluations in the work setting to illustrate responsive evaluation's potential added value. Other approaches for WHP evaluation have also been proposed, but these generally do not include empirical examples. Yet examples can help researchers get a better idea of how a new evaluation methodology in the WHP field works. The second strength is that, in addition to describing the various characteristics of responsive evaluation such as its aims and frequently used methods, we have compared it with a traditional evaluation methodology, the RCT. The reason for doing this was not so much the comparison as to inform the reader of the different scientific grounds on which both are based. These grounds impact what is considered good science and what is not. Information on the underlying scientific philosophies may enable researchers who are considering responsive evaluation to better evaluate its added value.

A limitation of this communication is that it describes only one methodology for evaluating WHP interventions. There certainly are other approaches that are of interest for WHP evaluation as well, for instance realist evaluation[12] and citizen science,[51] which have similarities and differences with responsive evaluations. However, the added value of this paper lies in the comprehensive, empirically illustrated and focused description of responsive evaluation. This allows other researchers to understand, consider and use this approach for their own research. A comparison between approaches certainly also has its value as this provides an overview of various methodologies that may be an alternative to traditional approaches for WHP evaluation. However, this would most likely reduce the extensiveness of the insight in the various methodologies, compared with focusing on one approach.

CONCLUSION

This communication presents responsive evaluation as an innovative methodology for evaluating WHP interventions. It describes why responsive evaluation is suitable for addressing workplace evaluation challenges such as the changeability of the work setting and the need to better match evaluations to the lifeworld of employees, especially those with low SEP. In addition, it provides insights into the scientific philosophy underlying responsive evaluation, how it has different expectations of what constitutes good science and why some elements might be relevant for WHP evaluation. Responsive evaluation also faces challenges, such as the translation of findings from one setting to another, although suggestions on how to do so are provided. Other methodologies for WHP intervention evaluation should be explored in future research to further contribute to finding ways to evaluate WHP interventions.

Acknowledgements We would like to thank Leon Pijnenburg for fruitful discussions on the topic in this communication.

Contributors HvH and JvB worked on the general outline. CB, TA and EdV provided intellectual input on the general outline and contributed their knowledge on research methodologies. HvH drafted the paper; all other authors contributed to further writing of the manuscript and read and approved the final version.

Funding The study was funded by the Netherlands Organisation for Health Research and Development (ZonMw; grant number 53-1001-411).

Competing interests None declared.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iD

Hanneke van Heijster http://orcid.org/0000-0001-9064-7788

REFERENCES

1 Robroek SJW, Oude Hengel KM, van der Beek AJ, *et al*. Socio-economic inequalities in the effectiveness of workplace health promotion programmes on body mass index: an individual participant data meta-analysis. *Obes Rev* 2020;21:e13101.
2 Coenen P, Robroek SJW, van der Beek AJ, *et al*. Socioeconomic inequalities in effectiveness of and compliance to workplace health promotion programs: an individual participant data (IPD) meta-analysis. *Int J Behav Nutr Phys Act* 2020;17:112.
3 van Heijster H, Boot CRL, Robroek SJW, *et al*. The effectiveness of workplace health promotion programs on self-perceived health of employees with a low socioeconomic position: an individual participant data meta-analysis. *SSM Popul Health* 2021;13:100743.
4 Cairns J-M, Bambra C, Hillier-Brown FC, *et al*. Weighing up the evidence: a systematic review of the effectiveness of workplace interventions to tackle socio-economic inequalities in obesity. *J Public Health* 2015;37:659–70.
5 Rongen A, Robroek SJW, van Lenthe FJ, *et al*. Workplace health promotion: a meta-analysis of effectiveness. *Am J Prev Med* 2013;44:406–15.
6 Stiehl E, Shivaprakash N, Thatcher E, *et al*. Worksite health promotion for low-wage workers: a scoping literature review. *Am J Health Promot* 2018;32:359–73.
7 Baron SL, Beard S, Davis LK, *et al*. Promoting integrated approaches to reducing health inequities among low-income workers: applying a social ecological framework. *Am J Ind Med* 2014;57:539–56.
8 Robroek SJW, Rongen A, Arts CH, *et al*. Educational inequalities in exit from paid employment among Dutch workers: the influence of health, lifestyle and work. *PLoS One* 2015;10:e0134867.
9 van der Put AC, Mandemakers JJ, de Wit JBF, *et al*. Worksite health promotion and social inequalities in health. *SSM Popul Health* 2020;10:100543.
10 Meershoek A, Horstman K. Creating a market in workplace health promotion: the performative role of public health sciences and technologies. *Crit Public Health* 2016;26:269–80.
11 Kristensen TS. Intervention studies in occupational epidemiology. *Occup Environ Med* 2005;62:205–10.
12 Robroek SJ, Coenen P, Oude Hengel KM. Decades of workplace health promotion research: marginal gains or a bright future ahead. *Scand J Work Environ Health* 2021;47:561–4.
13 Schelvis RMC, Oude Hengel KM, Burdorf A, *et al*. Evaluation of occupational health interventions using a randomized controlled trial: challenges and alternative research designs. *Scand J Work Environ Health* 2015;41:491–503.
14 Craig P, Dieppe P, Macintyre S, *et al*. Developing and evaluating complex interventions: the new medical Research Council guidance. *BMJ* 2008;337:a1655.
15 Larouche A, Potvin L. Stimulating innovative research in health promotion. *Glob Health Promot* 2013;20:64–9.
16 Griffiths A. Organizational interventions: facing the limits of the natural science paradigm. *Scand J Work Environ Health* 1999;25:589–96.
17 Cartwright N. Causal Inference. In: *Philosophy of social science - A new introduction*. United Kingdom: Oxford University Press, 2014: 308–25.
18 Schelvis RMC, Wiezer NM, van der Beek AJ, *et al*. The effect of an organizational level participatory intervention in secondary vocational education on work-related health outcomes: results of a controlled trial. *BMC Public Health* 2017;17:141.
19 van Heijster H, van Berkel J, Abma T, *et al*. Responsive evaluation of stakeholder dialogue as a worksite health promotion intervention to contribute to the reduction of Sep related health inequalities: a study protocol. *BMC Health Serv Res* 2020;20:196.
20 Baxter H. System and life-world in Habermas'stheory of communicative action. *Theory Soc* 1987;16:39–86.
21 Habermas J. Intermediate reflections: system and lifeworld. In: *The theory of communicative action*. London: Heinemann, 1984: 113–52.
22 De Neve J-W, Kawachi I. Spillovers between siblings and from offspring to parents are understudied: a review and future directions for research. *Soc Sci Med* 2017;183:56–61.
23 Stake RE. *Evaluating the arts in education: a responsive approach*. Columbus, Ohio: Merill, 1975.
24 Stake RE. Evaluation roots. In: *Development*, 2006: 1–5.
25 Guba EG, Lincoln YS. *Fourth generation evaluation*. Beverly Hills, CA: Sage, 1989.
26 Abma TA. Responsive evaluation: its meaning and special contribution to health promotion. *Eval Program Plann* 2005;28:279–89.
27 van Heijster H, van Berkel J, Boot CRL, *et al*. Stakeholder dialogue on dilemmas at work as a workplace health promotion intervention including employees with a low SEP: a responsive evaluation. *BMC Public Health* 2022;22:407.
28 Abma TA, Noordegraaf M. Public managers amidst ambiguity. *Evaluation* 2003;9:285–306.
29 Abma TA, Leyerzapf H, Landeweer E. Responsive evaluation in the interference zone between system and lifeworld. *Am J Eval* 2017;38:507–20.
30 Frambach JM, van der Vleuten CPM, Durning SJ. Am last page. quality criteria in qualitative and quantitative research. *Acad Med* 2013;88:552.
31 van der Borg WE, Verdonk P, Dauwerse L, *et al*. Work-related change in residential elderly care: trust, space and connectedness. *Hum Relat* 2017;70:805–35.
32 Baur V, Abma T. 'The Taste Buddies': participation and empowerment in a residential home for older people. *Ageing Soc* 2012;32:1055–78.
33 Abma TA. Book review. *Eval Program Plann* 2019;74:18–19.
34 Abma TA, Widdershoven GAM. Responsieve methodologie. In: *Interactief Onderzoek in de praktijk*. 1st ed. Den Haag: LEMMA, 2006.
35 Bonell C, Moore G, Warren E, *et al*. Are randomised controlled trials positivist? reviewing the social science and philosophy literature to assess positivist tendencies of trials of social interventions in public health and health services. *Trials* 2018;19:15–17.
36 Easterby-Smith M, Thorpe R, Jackson PR. The philosophy of management research. In: *Management research*. London: SAGE Publications Ltd, 2008: 55–77.
37 Habermas J. Reconstruction and Interpretation in the Social Sciences. In: *Moral consciousness and communicative action*. Cambridge, MA: MIT Press, 1990: 21–42.
38 Madsen IE, Rugulies R. Understanding the impact of psychosocial working conditions on workers' health: we have come a long way, but are we there yet? *Scand J Work Environ Health* 2021;47:483–7.
39 Lee AS. Integrating positivist and interpretive approaches to organizational research. *Organ Sci* 1991;2:342–65.
40 Block G, Sternfeld B, Block CH, *et al*. Development of alive! (a lifestyle intervention via Email), and its effect on health-related quality of life, presenteeism, and other behavioral outcomes: randomized controlled trial. *J Med Internet Res* 2008;10:e43.
41 Gadamer HG. *Wahrheit und Methode*. Tübingen: JCB Mohr, 1960.
42 Molewijk AC, Abma T, Stolper M, *et al*. Teaching ethics in the clinic. The theory and practice of moral case deliberation. *J Med Ethics* 2008;34:120–4.
43 Allender S, Colquhoun D, Kelly P. Competing discourses of workplace health. *Health* 2006;10:75–93.
44 van Berkel J, Meershoek A, Janssens RMJPA, *et al*. Ethical considerations of worksite health promotion: an exploration of stakeholders' views. *BMC Public Health* 2014;14:458.
45 Abma TA, Widdershoven GAM. *Responsieve methodologie*. LEMMA, 2006.
46 Abma T, Banks S, Cook T, *et al*. *Participatory research for health and social well-being*. 1st ed. Springer Nature Switzerland AG, 2019.
47 Abma T. Ethics work for good participatory action research. *Beleidsonderzoek Online* 2020.
48 Fricker M. *Epistemic injustice: power and the ethics of knowing*. Oxford: Oxford University Press, 2009.
49 Groot B, Haveman A, Abma T. Relational, ethically sound co-production in mental health care research: epistemic injustice and the need for an ethics of care. *Crit Public Health* 2022;32:230–40.
50 Green J, Thorogood N. *Qualitative methods for health research*. 3rd ed. London: SAGE Publications Ltd, 2014.
51 Lelie L, van der Molen HF, van den Berge M, *et al*. The process evaluation of a citizen science approach to design and implement workplace health promotion programs. *BMC Public Health* 2022;22:1610.