Effectiveness of health-oriented leadership interventions for improving health and wellbeing of employees: a systematic review

Indra Dannheim1 · Helena Ludwig-Walz2 · Anette E. Buyken3 · Valerie Grimm1 · Anja Kroke1,2

Received: 28 June 2021 / Accepted: 3 October 2021 / Published online: 23 November 2021
© The Author(s) 2021

Abstract
Aim To identify and summarize the evidence for the effect of health-oriented leadership interventions on health and well-being outcomes at the employee level following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement (Moher et al. 2009).

Subject and Methods A systematic search of relevant studies was conducted in multiple databases. Randomized controlled trials (RCTs), cluster-randomized controlled trials (cRCTs) and controlled before–after studies (CBAs) were included based on the following criteria: interventions that addressed supervisors, to raise awareness for the importance of health issues, teach mindfulness practices for conscious awareness, reduce stress and promote resources at the level of individual behavior, and evaluated the effect on at least one outcome of psychomental stress, absenteeism and well-being on the employee level.

Results Of 6126 publications retrieved, ten studies were identified for analysis. Significant effects of leadership training were reported on exhaustion tendency, self-reported sickness absence, work-related sickness absence and job satisfaction in studies comparing health-oriented training programs to no intervention. Studies comparing health-oriented leadership training to other training did not report significant effects. Risk of bias was judged to be high in seven studies and unclear in three studies.

Conclusion Evidence for the effectiveness of health-oriented leadership interventions on employees’ stress, absenteeism or well-being is judged to be low, clearly indicating the need for more and higher-quality research.

Keywords Leadership · Intervention · Occupational health · Employees · Health-oriented leadership · Systematic review

Introduction
Employees’ health and well-being is important not only for individuals themselves, but also for companies and public organizations, as it leads to higher productivity levels, less absenteeism, less employee turnover and lower health care costs (Wright and Bonett 2007; Ford et al. 2011; Faragher et al. 2013; Kramer and Son 2016; Beehr 2019). Creating and developing a healthy workplace is therefore an important challenge for companies, organizations and political stakeholders (World Health Organization 2005).

Supervisors function as connecting link between individual health and organizational health promotion and are promoters of occupational health and safety in worksite settings (Kelloway and Barling 2010; Rump et al. 2016; Straub et al. 2017; Rudolph et al. 2020). However, health promotion as a leadership task is often not recognized as such or is not included in leadership training. Therefore, such leadership interventions have grown in popularity (Struhs-Wehr 2017; Spitzenverband Bund der Krankenkassen 2018; Rudolph et al. 2020). Consequently, questions about the effectiveness of these interventions have arisen. To that end, a recent systematic review (Kuehnl et al. 2019) addressed the effectiveness of leadership interventions on outcomes at the employee level. Kuehnl et al. (2019) examined the effects of human resource management training of supervisors for improving the health and
well-being of employees. The review included interventions aiming to improve supervisor-employee interaction and the supervisors’ capability of designing the work environment. Interventions addressing health-oriented behaviors were excluded. Twenty-five studies were included in that review. Overall, inconsistent evidence was found for a positive impact of human resource management training programs on employees’ stress, well-being or absenteeism when compared to no intervention, and no evidence of a considerable effect was found when compared to other interventions. These results are surprising, since the influence of leadership behavior on employee health has repeatedly been postulated (Kuoppala et al. 2008; Nyberg et al. 2009; Skakon et al. 2010; Gregersen et al. 2011; Schyns and Schilling 2013; Montano et al. 2017). This may be partly attributable to the inclusion and exclusion criteria for studies proposed by Kuehnl et al. (2019).

According to the concept of health-oriented leadership described by Franke et al. (2014), leaders’ management of their own health (described as i.e., leaders’ self-care), composed of the three dimensions’ value, awareness and behavior, serves as a relevant precondition for health-oriented leadership behavior (Franke et al. 2015). In line with this concept, leaders must perceive their own health as important (dimension value), must be aware of their health status (dimension awareness) and take care of their health (dimension behavior) in order to be able to lead in a health-oriented way. Franke et al. (2015) suggested that leadership training should be built on these dimensions and as such should (i) raise awareness for the importance of health issues (dimension value), (ii) teach mindfulness practices for conscious awareness (dimensions awareness) and (iii) reduce stress and promote resources at the level of individual behavior (dimension behavior). However, these aspects of leadership training were not taken into account by Kuehnl et al. (2019). Therefore, the aim of this systematic review was to identify and summarize the evidence for the effects of health-oriented training programs targeted at supervisors—which explicitly address leaders’ management of their own health based on the concept of health-oriented leadership by Franke et al. (2014)—on the psychosomatic stress, absenteeism or well-being of employees.

**Methods**

For this systematic review, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al. 2009; checklist contained in Appendix 1) were adhered to, and it was registered in the International Prospective Register of Systematic Reviews (PROSPERO) (CRD42020205502).

**Data sources**

A literature search was performed on September 3, 2020, using four databases: Cochrane Library, MEDLINE, PsycINFO and Web of Science. No search restrictions were made regarding language or year of publication. Searches were performed using database-specific index terms (e.g. Medical Subject Headings) and relevant free text words in titles and abstracts. The specific search strategies, including applied search terms for each database, are outlined in Appendix 2. To identify unpublished and ongoing trials, the following trial registers were searched: ClinicalTrials.gov, WHO International Clinical Trials Registry Platform, and Trials Register of Promoting Health Interventions. A manual search was performed in Google Scholar and the Federal Institute for Occupational Safety and Health (BAuA). Additionally, systematic reviews and reference lists of included studies were checked for eligible references.

**Inclusion and exclusion criteria**

The population-intervention-comparison-outcome-study design schema (Higgins et al. 2020) was utilized to determine the eligibility criteria:

1. **Population (P):** Studies that enrolled supervisors as the target group, independently of gender, management level and number of employees.
2. **Intervention (I):** Studies that conducted interventions to (i) raise awareness of the importance of health issues (dimension value), (ii) teach mindfulness practices for conscious awareness (dimensions awareness) and (iii) reduce stress and promote resources at the level of individual behavior (dimension behavior). Interventions that focused on improving supervisor–employee interaction (e.g. providing social support, communication skills) and were also open to employees as participants were excluded.
3. **Comparison (C):** Studies comparing health-oriented interventions targeted at supervisors with a passive control group or with an active control group receiving an alternative intervention.
4. **Outcome (O):** Studies with measurements of psychosomatic stress, such as the Maslach Burnout Inventory (Maslach et al. 1996), any estimate of absenteeism and measures of well-being, such as the Utrecht Work Engagement Scale (Schaufeli et al. 2002). Outcome measurements must have been performed in employees, not supervisors.
5. **Study design (S):** Randomized controlled trials (RCTs), cluster-randomized controlled trials (cRCTs) and con-
trolled before–after studies (CBAs) measuring outcomes both before and after the intervention.

Following the Cochrane Handbook of Systematic Reviews of Interventions (Higgins et al. 2020) suggestions, two reviewers (I.D. and H.L.-W.) independently searched the different databases for potential studies. In a first step, titles and abstracts were screened and irrelevant studies were excluded. In a second step, the full text of the remaining studies was retrieved and screened according to the inclusion and exclusion criteria. Any disagreements were resolved by discussion or, if necessary, by third-party consultation (A.K.). Reasons for study exclusion are outlined in Appendix 3.

**Data extraction**

Two researchers (I.D. and V.G.) independently extracted data from each study into a predefined data extraction form. Information was collected on study framework, intervention, sample size and relevant outcomes including description of measurement, statistical analyses and description of main results. Any disagreement was resolved by discussion or, if necessary, by third-party consultation (H.L.-W.).

**Risk of bias assessment**

RCTs and cRCTs were assessed using the Revised Cochrane Tool for randomized trials, described in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins et al. 2020). The following potential sources of bias were added for cluster-randomized trials: recruitment bias, baseline imbalance, loss of cluster, incorrect statistical analysis, comparability with randomized trials. CBAs were evaluated following the Cochrane Effective Practice and Organisation of Care (EPOC) criteria (EPOC 2017). Based on the overall risk of bias judgment, each domain of bias of included studies was rated as high, low or unclear. At study level, risk of bias was considered to be high when one or more key domains were judged to be at high risk of bias (Higgins et al. 2020). In RCTs and cRCTs, key domains were random sequence generation, allocation concealment, incomplete outcome data and selective outcome reporting. In CBAs, key domains were similarity of baseline outcome measurements, similarity of baseline participant characteristics, adequately addressing incomplete outcome data, adequately preventing knowledge of the allocated interventions during the study, adequately protecting against contamination, and freedom from selective outcome reporting. Rating criteria for the risk of bias assessment that was independently performed by two reviewers (I.D. and V.G.), are presented in Appendix 4 and Appendix 5. Any disagreement was resolved by consensus or, where needed, with consultation of a third reviewer (H.L.-W.).

**Results**

**Study selection**

In total, 6126 results were generated from the initial database search. Sixteen articles were identified through manual search. After removal of 861 duplicates, title-abstract screening of 5281 publications was performed. Of these, 31 full-text papers were selected for detailed eligibility screening. Finally, ten studies remained for qualitative analyses. Figure 1 illustrates the selection process. Reasons for exclusion of the full-text screened studies are described in Appendix 3.

**Study characteristics**

Among the ten identified studies, six were cRCTs and four were controlled before–after studies. Seven studies were performed between 2013 and 2019 (Angelo and Chambel 2013; Elo et al. 2014; Stansfeld et al. 2015; Milligan-Saville et al. 2017; Barrech et al. 2018; Lange and Rowold 2019; Veloso-Besio et al. 2019), the others in 2006 and 2005 (Kawakami et al. 2005; Kawakami et al. 2006; Takao et al. 2006). Five studies were conducted in Europe (Angelo and Chambel 2013; Elo et al. 2014; Stansfeld et al. 2015; Barrech et al. 2018; Lange and Rowold 2019), three in Japan (Kawakami et al. 2005; Kawakami et al. 2006; Takao et al. 2006), one in Chile (Veloso-Besio et al. 2019) and one in Australia (Milligan-Saville et al. 2017). Interventions were heterogeneous, ranging from a single 4-hour training session (Milligan-Saville et al. 2017) to comprehensive training programs lasting several days, including short lectures on well-being, role playing and group discussions on leadership (Elo et al. 2014). In addition to face-to-face interventions, three studies (Kawakami et al. 2005; Kawakami et al. 2006; Stansfeld et al. 2015) applied a web-based training program for supervisors. Interventions were carried out in different organizational settings, including fire departments (Angelo and Chambel 2013; Milligan-Saville et al. 2017), a computer software engineering company (Kawakami et al. 2006) and a public hospital (Veloso-Besio et al. 2019). The number of employees included in the statistical analyses ranged from 70 (40 in intervention group and 30 in control group) (Veloso-Besio et al. 2019) to 1966 (1233 employees in intervention group, 733 employees in control group) (Milligan-Saville et al. 2017). All studies were screened for the assessment of stress, absenteeism and well-being of employees. Eight studies measured outcomes of stress using seven different questionnaires (Kawakami et al. 2005; Kawakami et al. 2006; Takao et al. 2006; Angelo
and Chambel 2013; Elo et al. 2014; Stansfeld et al. 2015; Barrech et al. 2018; Lange and Rowold 2019), two studies measured outcomes of absenteeism differentiating between self-reported, work-related and standard sick leave (Stansfeld et al. 2015; Milligan-Saville et al. 2017) and three studies measured well-being outcomes with three different scales (Angelo and Chambel 2013; Stansfeld et al. 2015; Veloso-Besio et al. 2019). Only one study addressed all three outcomes simultaneously (Stansfeld et al. 2015).

Table 1 presents a summary of the characteristics of the ten studies.

**Risk of bias**

A summary of the risk of bias assessment is presented in Fig. 2. Overall, seven studies were judged to be at an overall high risk of bias, and three studies to be at an overall unclear risk of bias. For detailed justifications for risk of bias judgment, see Appendix 6. Due to the low methodological quality of most studies and the insufficient number of outcomes within each comparison, meta-analysis and a quality of evidence assessment were not performed.

**Effects of interventions**

Table 2 presents the reported results of the ten included intervention studies regarding the investigated outcomes of stress, absenteeism and well-being. Results are presented separately according to control type (intervention versus no intervention; intervention versus other intervention). Studies were grouped according to study design, follow-up times and outcome. Follow-up times were categorized into short-term (less than 3 months), mid-term (3 months to 1 year) and long-term (more than 1 year).

**Type of control group**

Eight studies evaluated the effectiveness of their intervention compared to no intervention (Takao et al. 2006; Angelo and Chambel 2013; Stansfeld et al. 2015; Elo et al. 2014; Milligan-Saville et al. 2017; Barrech et al. 2018; Lange and Rowold 2019; Veloso-Besio et al. 2019), and three studies compared health-oriented training programs of supervisors to another intervention (Kawakami et al. 2005; Kawakami et al. 2006; Lange and Rowold 2019).
| Author; Year Country | Study Design | Occupational Setting | Intervention(s) | Sample size | Outcome measurements and applied assessment instruments |
|----------------------|--------------|----------------------|-----------------|-------------|--------------------------------------------------------|
| Angelo and Chambel 2013; Portugal | CBA | Fire department | Stress management training; 21 hours distributed over 3 days | Supervisors: IG & CG: 33; Employees: IG: 67; CG: 37 | Psychomental stress: 1. Emotional exhaustion. 5 items, Maslach Burnout Inventory (Maslach et al. 1996) 2. Cynicism. 3 items, Maslach Burnout Inventory (Maslach et al. 1996) Absenteeism: none Well-being: 1. Vigor. 5 items, Utrecht Work Engagement Scale (Schaufeli et al. 2002) 2. Dedication. 5 items, Utrecht Work Engagement Scale (Schaufeli et al. 2002) |
| Barrech et al. 2018; Switzerland | cRCT | Multinational healthcare provider | Intervention to reduce job insecurity; six training sessions (3 seminars; 3 peer-counselling sessions); 2-4 h each over three months | Supervisors: IG: 52 (ITT), 20 (PP); CG: 47 (ITT), 4 (PP) Employees: IG: 391 (ITT), 51 (PP); CG: 424 (ITT), 28 (PP) | Psychomental stress: 1. Anxiety. 7 items, Hospital Anxiety and Depression Scale (HADS) (Herrmann-Lingen et al. 2011) 2. Depression. 7 items, Hospital Anxiety and Depression Scale (HADS) (Herrmann-Lingen et al. 2011) 3. Exhaustion tendency. 3 items, Giessen Subjective Complaints List (GBB) (Brähler and Scheer 1995) Absenteeism: none Well-being: none |
| Elo et al. 2014; Finland | CBA | Public sector organization (maintains and constructs streets, green areas and public buildings) | A personal growth-orientated leadership intervention; 7.5 days carried out in one to three-day sessions over a six-month period | Supervisors: IG: 8; CG: 32 Employees: IG: 49; CG: 96 | Psychomental stress: 1. Emotional exhaustion. 5 items, Maslach Burnout Inventory (Maslach et al. 1996) 2. Perceived stress symptoms. 1 item, single-item measure of stress symptoms (Elo et al. 2003) Absenteeism: none Well-being: none |
| Kawakami et al. 2006; Japan | cRCT | Computer software engineering company | Web-based supervisor training focusing on worksite mental health; nine chapters; 3-5h over a four-week training period; self-administered | Supervisors: IG: 22 (ITT), 21 (PP); CG: 23 Employees: IG: 85 (ITT), 81 (PP); CG: 114 (ITT), 108 (PP) | Psychomental stress: 1. Psychological distress. 18 items, Brief Job Stress Questionnaire (BJSQ) (Shimomitsu et al. 2000) Absenteeism: none Well-being: none |
| Author; Year Country | Study Design | Occupational Setting | Intervention(s) | Sample size | Outcome measurements and applied assessment instruments |
|----------------------|--------------|----------------------|-----------------|-------------|----------------------------------------------------------|
| Kawakami et al. 2005; Japan | cRCT | Office machine sales and service company | Web-based supervisor training focusing on worksite mental health; nine chapters; 3-5h over a four-week training period; self-administered | Supervisors: IG: 9 (ITT), 8 (PP); CG: 7 Employees: IG: 82 (ITT), 75 (PP); CG: 84 (ITT), 84 (PP) | Psychomental stress: 1. Psychological distress. 18 items, Brief Job Stress Questionnaire (BJSQ) (Shimomitsu et al. 2000) Absenteeism: none Well-being: none |
| Lange and Rowold 2019; Germany | CBA | Different sectors (industry, finance & insurance, health, attendance, consulting & auditing, public services and others) | Mindful Leadership Intervention covering the topics stress, stress management, mindfulness, leadership and communication; one-day training (7hr), two follow-up sessions (one-on-one coaching (30min); group session (90min) and digital based mindfulness/relaxation instructional videos (voluntary) over a three-month period | Supervisors: IG: 19; CG1: 21; CG2: 18 Employees: IG: 127 (ITT), 97 (PP); CG1: 70 (ITT), 65 (PP); CG2: 63 (ITT), 57 (PP) | Psychomental stress: 1. Irritation. 8 items, Irritation Scale (Mohr et al. 2006) Absenteeism: none Well-being: none |
| Milligan-Saville et al. 2017; Australia | cRCT | Fire department | Mental health training; single four-hour training session | Supervisors: IG: 46 (ITT), 45 (PP); CG: 42 (ITT), 40 (PP) Employees: IG: 1233; CG: 733 | Psychomental stress: none Absenteeism: 1. Work-related sick leave. Administrative sickness absence records (human resource department of the Fire and Rescue New South Wales (FRNSW)) 2. Standard sick leave. Administrative sickness absence records (human resource department of the Fire and Rescue New South Wales (FRNSW)) Well-being: none |
| Author; Year | Country | Study Design | Occupational Setting | Intervention(s) | Sample size | Outcome measurements and applied assessment instruments |
|-------------|---------|--------------|---------------------|-----------------|-------------|-------------------------------------------------------|
| Stansfeld et al. 2015; UK | UK | cRCT | National Health Service (NHS) Mental Health Trust | Guided e-learning health programme focusing on work-related stress; six separate modules; weekly to bi-weekly modules over a three-month period | Supervisors: IG: 49 (ITT), 21 (PP); CG: 11 Employees: IG: 341 (ITT), 225 (PP – wellbeing), 294 (PP – sickness absence HR Data), 198 (PP – sickness absence self-report); CG: 83 (ITT), 59 (PP – wellbeing), 66 (PP – sickness absence HR Data), 51 (PP – sickness absence self-report) | Psychomental stress: 1. Psychological distress. 12 items, General Health Questionnaire (GHQ) (Goldberg and Williams 1988) Absenteeism: 1. Sickness absence. Measured in days excluding absences greater than 21 days (reporting system of the NHS Trust and local Services) 2. Self-report sickness absence. Medium-term (7-21 days) Well-being: 1. Employee well-being. 14 items, Warwick Edinburgh Mental Wellbeing Scale (WEMWBS) (Tennant et al. 2007) |
| Takao et al. 2006; Japan | Japan | cRCT | Sake brewery | Supervisory education with focus on mental health promotion; single session (1hr), active listening training (3hr); three-month education program | Supervisors: IG: 24 (ITT), 23 (PP); CG: 22 Employees: IG: 154 (ITT), 134 (PP); CG: 101 (ITT), 92 (PP) | Psychomental stress: 1. Psychological distress. 18 items, Brief Job Stress Questionnaire (BJSQ) (Shimomitsu et al. 2000) Absenteeism: none Well-being: none |
| Veloso-Besio et al. 2019; Chile | Chile | CBA | Public hospital | Intervention program based on positive psychology and social skills; twelve sessions (60-90 min); twice a week during 1 month and three weeks | Supervisors: IG: 7; CG: 8 Employees: IG: 41 (ITT), 40 (PP); CG: 36 (ITT), 30 (PP) | Psychomental stress: none Absenteeism: none Well-being: 1. Job satisfaction. 15 items, Overall Job Satisfaction Scale (Warr et al. 1979) |
Outcome variable Stress

Three cRCTs examined the effectiveness of health-oriented training programs on outcomes of stress compared to no intervention using a mid-term follow-up (Takao et al. 2006; Stansfeld et al. 2015; Barrech et al. 2018). Statistically significant intervention effects were only reported on exhaustion tendency as measured by the Giessen Subjective Complaints List (Brähler and Scheer 1995). Regarding mid- and long-term follow-up of CBAs in comparison to no intervention, no statistically significant intervention effect on employees’ stress was found for supervisor training (Angelo and Chambel 2013; Elo et al. 2014; Lange and Rowold 2019). In comparison to another training, no intervention effect was found on irritation, a stress outcome, at mid-term follow-up of one CBA (Lange and Rowold 2019) and on psychological distress stress in two cRCTs (Kawakami et al. 2005; Kawakami et al. 2006).

Outcome variable Absenteeism

Two cRCTs analyzed whether training of supervisors had an effect on employees’ absenteeism (compared to no intervention) using mid-term follow-up (Stansfeld et al. 2015; Milligan-Saville et al. 2017). Marginally significant intervention effects ($p < 0.10$) were seen on self-reported sickness absence (Stansfeld et al. 2015). Statistically significant intervention effects ($p < 0.05$) were reported on work-related sick leave (Milligan-Saville et al. 2017).

Outcome variable Well-Being

Well-being as outcome was examined in three studies comparing the intervention to no intervention, one cRCT using a mid-term follow-up (Stansfeld et al. 2015), one CBA using a short-term follow-up (Veloso-Besio et al. 2019) and one CBA using a mid-term follow-up (Angelo and Chambel 2013). Intervention effects of supervisor training on well-being outcomes were only reported in two CBAs. Marginally significant intervention effects ($p < 0.10$) were seen on vigor at mid-term follow-up (Angelo and Chambel 2013), and statistically significant changes ($p < 0.05$) were seen on job satisfaction at short-term follow-up (Veloso-Besio et al. 2019).

Discussion

Summary and evaluation of available evidence is a fundamental prerequisite for developing effective evidence-based health promotion interventions and preventive activities.
| Study                                                                 | Statistical analysis | Intervention                                                                 |
|---------------------------------------------------------------------|----------------------|-----------------------------------------------------------------------------|
| **Training vs. no intervention/outcome stress/mid-term**             |                      |                                                                             |
| Angelo and Chambel 2013; CBA                                        | Analysis of Variance | Emotional exhaustion, mean (SD): IG: 1.90 (1.38), CG: 1.45 (1.18)          |
|                                                                     |                      | Cynicism, mean (SD): IG: 0.91 (1.26), CG: 0.96 (1.46)                       |
| Barrech et al. 2018; cRCT                                            | Analysis of covariance | Anxiety: IG: NR, CG: NR                                                     |
|                                                                     |                      | Depression: IG: NR, CG: NR                                                  |
|                                                                     |                      | Exhaustion tendency: IG: NR, CG: NR                                         |
| Lange and Rowold 2019; CBA                                          | Analysis of covariance | Irritation, mean (SD): IG: 2.81 (2.63), CG: 2.75 (0.82)                    |
| Stansfeld et al. 2015, cRCT                                         | Random effects model | Psychological distress, mean (SD): IG: 2.8 (3.5), CG: 3.2 (3.4)            |
| Takao et al. 2006; cRCT                                             | Analysis of Variance | Psychological distress, mean (SE): IG: 26.8 (0.81), CG: 27.4 (0.88)        |
| **Outcome Stress/long-term**                                        |                      |                                                                             |
| Elo et al. 2014; CBA                                                | Analysis of Variance | Emotional exhaustion, mean (SD): IG: 2.29 (1.49), CG: 1.63 (1.16)          |
|                                                                     |                      | Stress symptoms, mean (SD): IG: 2.69 (1.07), CG: 2.14 (0.92)               |
| **Training vs. no intervention/outcome absenteeism/mid-term**        |                      |                                                                             |
| Milligan-Saville et al. 2017; cRCT                                  | Linear regression with generalized estimating equations and robust SEs   | Work-related sick leave, mean (SE): IG: 1.56% (0.23), CG: 0.95% (0.20)    |
|                                                                     |                      | Standard sick leave, mean (SE): IG: 4.97% (0.22), CG: 5.27% (0.21)         |
| Stansfeld et al. 2015, cRCT                                         | Random effects model  | Days off sick from HR data, mean (SD): IG: 1.2 (3.2), CG: 0.9 (2.0)       |
|                                                                     |                      | Days off sick self-report, mean (SD): IG: 1.0 (3.0), CG: 1.2 (3.5)         |
| **Training vs. no intervention/outcome well-being/short-term**       |                      |                                                                             |
| Veloso-Besio et al. 2019; CBA                                       | Analysis of Variance | Job satisfaction, mean (SD): IG: 64.55 (16.17), CG: 66.90 (18.39)          |
| **Outcome Well-Being/mid-term**                                     |                      |                                                                             |
| Angelo and Chambel 2013; CBA                                        | Analysis of Variance | Vigor, mean (SD): IG: 5.22 (0.67), CG: 5.34 (0.54)                         |
|                                                                     |                      | Dedication, mean (SD): IG: 5.24 (0.85), CG: 5.40 (0.76)                    |
| Stansfeld et al. 2015, cRCT                                         | Random effects model  | Employee well-being, mean (SD): IG: 51.0 (8.3), CG: 50.4 (8.0)            |
| **Training vs. other training/outcome stress/mid-term**             |                      |                                                                             |
| Kawakami 2006; cRCT                                                 | Analysis of Variance | Psychological distress, mean (SD): IG: 52.5 (9.3), CG: 50.2 (10.0)         |
|                                                                     |                      | Psychological distress, mean (SD): IG: 50.6 (9.9), CG: 49.2 (11.5)         |
| **Outcome Well-Being/mid-term**                                     |                      |                                                                             |
| Angel and Chambel 2013; CBA                                         | Analysis of Variance | Vigor, mean (SD): IG: 5.22 (0.67), CG: 5.34 (0.54)                         |
|                                                                     |                      | Dedication, mean (SD): IG: 5.24 (0.85), CG: 5.40 (0.76)                    |
| Stansfeld et al. 2015, cRCT                                         | Random effects model  | Employee well-being, mean (SD): IG: 49.9 (8.3), CG: 49.0 (8.5)            |
|                                                                     |                      | Psychological distress, mean (SD): IG: 50.6 (9.9), CG: 49.2 (11.5)         |

| Effect | Stress/mid-term | Emotional exhaustion, p value: 0.317 |
|        | Anxiety, p value: 0.779 |
|        | Depression, p value (95% CI): 0.099 (-0.57 to 6.54) |
|        | Exhaustion tendency, p value (95% CI): 0.013 (-1.64 to -0.20) |
|        | Stress symptoms, p value: 0.70 |
|        | Psychological well-being, p value (95% CI): 0.2 (-2.0 to 2.5) |
|        | Psychological distress, p value (95% CI): 0.715 |
|        | Work-related sick leave, p value: 0.049 |
|        | Standard sick leave, p value: 0.73 |
|        | Days off sick from HR data, p value (95% CI): 0.6 (-1.4 to 2.6) |
|        | Days off sick self-report, p value (95% CI): 0.1 (-2.2 to 2.4) |
|        | Job satisfaction, p value: 0.01 |
|        | Work-related sick leave, p value: 0.049 |
|        | Standard sick leave, p value: 0.73 |
|        | Days off sick from HR data, p value (95% CI): 0.6 (-1.4 to 2.6) |
|        | Days off sick self-report, p value (95% CI): 0.1 (-2.2 to 2.4) |
|        | Job satisfaction, p value: 0.01 |
|        | Work-related sick leave, p value: 0.049 |
|        | Standard sick leave, p value: 0.73 |
|        | Days off sick from HR data, p value (95% CI): 0.6 (-1.4 to 2.6) |
|        | Days off sick self-report, p value (95% CI): 0.1 (-2.2 to 2.4) |
|        | Job satisfaction, p value: 0.01 |
|        | Work-related sick leave, p value: 0.049 |
|        | Standard sick leave, p value: 0.73 |
|        | Days off sick from HR data, p value (95% CI): 0.6 (-1.4 to 2.6) |
|        | Days off sick self-report, p value (95% CI): 0.1 (-2.2 to 2.4) |
|        | Job satisfaction, p value: 0.01 |
|        | Work-related sick leave, p value: 0.049 |
|        | Standard sick leave, p value: 0.73 |
|        | Days off sick from HR data, p value (95% CI): 0.6 (-1.4 to 2.6) |
|        | Days off sick self-report, p value (95% CI): 0.1 (-2.2 to 2.4) |
|        | Job satisfaction, p value: 0.01 |
Therefore, this systematic review identified and summarized the evidence for the effectiveness of health-oriented training programs targeted at supervisors that addressed leaders’ management of their own health, based on the concept of health-oriented leadership by Franke et al. (2014), on psychomental stress, absenteeism or well-being of employees. Overall, ten intervention studies could be identified assessing at least one employee-related outcome. Significant positive effects of leadership training were reported on exhaustion tendency (Barrech et al. 2018), self-reported sickness absence (Stansfeld et al. 2015), work-related sickness absence (Milligan-Saville et al. 2017) and job satisfaction (Veloso-Besio et al. 2019) in studies comparing health-oriented training programs to no intervention. However, due to the high overall risk of bias assessed for three studies and the overall unclear risk of bias judgment for one study, these effects should be interpreted with caution. Nonsignificant effects, on the other hand, were reported from studies comparing health-oriented leadership training to other training. Also, these results should be interpreted with caution due to the overall low study quality. In addition to the aforementioned high risk of bias in most studies, small sample sizes, nonrandomized study designs and nonrandom allocation, as well as incomplete and selective reporting, were identified. Based on these findings, the evidence for the effectiveness of the evaluated health-oriented trainings programs targeted at supervisors on employees’ psychomental stress, absenteeism or wellbeing should be judged as low. This finding is similar to that of Kuehnl et al. (2019) and Stuber et al. (2020), who also found no clear evidence for the effectiveness of the evaluated training programs targeted at supervisors. Hence, similar to their conclusion, a strong need for further well-designed studies is to be stated.

Beyond study quality, other explanations for the findings should also be explored, given the strong notion of a relevant role of leaders in promoting employee’s health (Gregersen et al. 2011; Montano et al. 2017; Kaluza et al. 2020). Firstly, (the majority) of leadership training programs may not be effective; i.e. the underlying theories, applied methods and training content focusing on improving leaders’ behavior and capabilities might be not sufficient to yield comprehensive improvements in employees’ health and well-being.

Second, studies may be unable to measure the effectiveness of health-oriented leadership interventions on outcomes at the employee level, as these are complex interventions with a comprehensive, multistep pathway from leadership training to improved employee health and well-being: training programs must be perceived as helpful by supervisors, must induce changes in attitudes and must result in successful acquisition of knowledge and skills. Eventually, these new attitudes, knowledge and skills need to be transferred into practice. The modified practice then has to effectively exert its influence on employees in terms of improved health.
or well-being. Finally, these changes must be assessed. Given this complex multistep pathway, future studies might better focus on single steps.

A third explanation relates to the workplace settings in which the studies are conducted. Carrying out intervention studies in this dynamic and complex setting has inherent complications, which may influence the effectiveness of the interventions. These are mostly “sideline” activities which are not directly relevant to core task completion (Kristensen 2005). The resulting incomplete study compliance and rapidly changing organizational contexts might hamper long-term follow-up measurements (Nielsen et al. 2010). In addition, competing high job demands or available resources also shape leadership behavior (Zimber et al. 2015; Knudsen et al. 2009; Mohr and Wolfram 2010; Arnold and Rigotti 2020). Therefore, mutual commitment between researchers and organizations/managers needs to be built up first in order to improve the evidence base in occupational health research (Kristensen 2005; Schelvis et al. 2016).

Limitations of the included studies

The included studies are subject to various limitations. First, due to the natural settings, randomization and blinding was difficult to realize, thereby increasing the likelihood of bias. Second, outcomes were mainly assessed using measurement tools relying on self-reports, which might have distorted study results. Third, a wide variety of different measurement tools were applied to assess study outcomes. Fourth, high dropout rates increased the likelihood of an underestimated intervention effect. Fifth, it remained unclear which moderating or mediating factors influenced the observed intervention effects. Sixth, the time between intervention and follow-up might have been too short to produce intervention effects on employees’ health or well-being. Finally, the heterogeneity in study design (e.g. training content, timing and mode of delivery of interventions) reduced study comparability.

Limitations of the systematic review

Although this systematic review was conducted according to the standards of the PRISMA statement (Moher et al. 2009), several limitations have to be considered. It is possible that not all relevant intervention studies were identified, although an expanded search in various databases was conducted. Other or additional search terms might have led to more potentially relevant publications. The inclusion and exclusion criteria were based on a selected health-oriented leadership concept, that of Franke et al. (2014). However, previous authors have criticized the construct proliferation (accumulation of seemingly different healthy leadership concepts but potentially identical constructs) of different health-oriented leadership approaches (Rudolph et al. 2020). Therefore, the consideration of different leadership concepts considering different behaviors, e.g. relationship-oriented, task-oriented, change-oriented (Wegge et al. 2014; Inceoglu et al. 2018) or health-beneficial leadership styles (such as transformational leadership (Bass and Riggio 2006), may add further insights.

Conclusion

In summary, evidence for the effectiveness of the evaluated health-oriented leadership interventions on employees’ stress, absenteeism or well-being is judged to be low. Instead, the results of this systematic review call for more and higher-quality research. Future results might then allow recommendations to be made for the conception of effective leadership interventions, that is, to exert significant positive effects on employees’ health. Thus, research on health-oriented leadership remains a central task in the field of occupational health research (Rudolph et al. 2020).

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10389-021-01664-1.

Acknowledgement The authors acknowledge the financial support provided within the framework of the „DEAL“ project (FKZ 01ZZ1520). The authors would like to thank all colleagues and partners of the DEAL project for their important contribution to the present systematic review. In addition, we want to thank all publications that were part of this review. The quality of the presented data depends on the quality of the original research. We encourage the authors of future health-oriented leadership interventions to improve the evidence base in occupational health research (Rudolph et al. 2020).

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not
References

Angelo R-P, Chambel M-J (2013) An intervention with firefighters to promote psychological occupational health according to the Job Demands-Resources Model. Rev Psicol Soc 28(2):197–210. https://doi.org/10.1174/021347413806196753

Arnold M, Rigotti T (2020) The leader in the spotlight: Health-oriented leadership and its antecedents and outcomes. Proceedings 2020(1):16724. https://doi.org/10.5465/AMBPP.2020.16724abstr

Barrech A, Seubert C, Glaser J, Gündel H (2018) Can a workplace leadership intervention reduce job insecurity and improve health? Results from a field study. Int Arch Occup Environ Health 91(5):547–557. https://doi.org/10.1007/s00420-018-1302-y

Bass BM, Riggio RE (2006) Transformational leadership. L. Erlbaum Associates, Mahwah

Beehr TA (2019) Interventions in occupational health psychology. J

Bass BM, Riggio RE (2006) Transformational leadership. L. Erlbaum Associates, Mahwah

Beex TA (2019) Interventions in occupational health psychology. J Occup Health Psychol 24(1):1–3. https://doi.org/10.1037/ocpp000140

Brähler E, Scheer JW (1995) Der Gießener Beschwerdebogen. [The Giessen Subjective Complaints List], 2nd edn. Huber, Bern

Cochrane Effective Practice and Organisation of Care (EPOC) (2017) Suggested risk of bias criteria for EPOC reviews. http://www.epoc.cochrane.org/epoce-resources-review-authors. Accessed 02 October 2020

Elo A-L, Leppänen A, Jäähkola A (2003) Validity of a single-item measure of stress symptoms. Scand J Work Environ Health 29(6):444–451. https://doi.org/10.5271/sjweh.752

Elo A-L, Ervasti J, Kuosma E, Mattila-Holappa P (2014) Effect of a leadership intervention on subordinate well-being. J Mgmt Dev 33(3):182–195. https://doi.org/10.1108/JMD-11-2012-0146

Faragher EB, Cass M, Cooper CL (2013) The Relationship between Job Satisfaction and Health: A Meta-Analysis. In: Cooper CL (ed) From Stress to Wellbeing. The Theory and Research on Occupational Stress and Wellbeing, 1st edn, vol 1. Palgrave Macmillan, UK, London, pp 254–271

Ford MT, Cerasoli CP, Higgens JA, Deccesare AL (2011) Relationships between psychological, physical, and behavioural health and work performance: A review and meta-analysis. Work Stress 25(3):185–204. https://doi.org/10.1080/02678373.2011.609035

Franke F, Felle J, Punti A (2014) The Impact of Health-Oriented Leadership on Follower Health: Development and Test of a New Instrument Measuring Health-Promoting Leadership. Ger J Hum Resour Manag: Zeitschrift f&#252;r Personalforschung 28(1-2):139–161. https://doi.org/10.1079/239700221402800108

Franke F, Ducki A, Felle J (2015) Gesundheitsförderliche Führung. In: Felle J (ed) Trends der psychologischen Führungsforschung. Hogrefe, Göttingen, pp 253–263

Goldberg DP, Williams P (1988) A User’s guide to the GHQ. NFER Nelson, Windsor

Gregersen S, Kuhnert S, Zimber A, Nienhaus A (2011) Führungsverhalten und Gesundheit - Zum Stand der Forschung (Leadership behaviour and health - current research state). Gesundheitswesen (Bundesverband der Ärzte des Öffentlichen Gesundheitsdienstes (Germany)) 73(1):3–12. https://doi.org/10.1055/s-0029-1246180

Herrmann-Lingen C, Buss U, Snaith R (2011) Hospital anxiety and depression scale—Deutsche Version (HADS-D) [Hospital Anxiety and Depression Scale—German version (HADS-D)]. 3. aktual. Hans Huber, Bern

Higgins JPT, Thomas J, Chandler J, Cumpton M, Li T, Page MJ, Welch VA (eds) (2021) Cochrane handbook for systematic reviews of interventions version 6.2 (updated February 2021). Cochrane. Available from www.training.cochrane.org/handbook

Inceoglu I, Thomas G, Chu C, Plans D, Gerbasi A (2018) Leadership behavior and employee well-being: An integrated review and a future research agenda. Leadersh Q 29(1):179–202. https://doi.org/10.1016/j.leaqua.2017.12.006

Kaluza AJ, Boer D, Buengeler C, van Dick R (2020) Leadership behaviour and leader self-reported well-being: A review, integration and meta-analytic examination. Work Stress 34(1):34–56. https://doi.org/10.1080/02678373.2019.1617369

Kawakami N, Kobayashi Y, Takao S, Tsutsuji A (2005) Effects of web-based supervisor training on supervisor support and psychological distress among workers: A randomized controlled trial. Prev Med 41(2):471–478. https://doi.org/10.1016/j.ypmed.2005.01.001

Kawakami N, Takao S, Kobayashi Y, Tsutsuji A (2006) Effects of web-based supervisor training on job stressors and psychological distress among workers: a workplace-based randomized controlled trial. J Occup Health 48(1):28–34. https://doi.org/10.1539/joh.48.28

Kelloway EK, Barling J (2010) Leadership development as an intervention in occupational health psychology. Work Stress 24(3):260–279. https://doi.org/10.1080/02678373.2010.518441

Knorr L, Kroke A (2012) Evidence-based public health nutrition: what constitutes good evidence? CAB Reviews 7(045). https://doi.org/10.1079/PAVSNR20127045

Knudsen HK, Ducharme LJ, Roman PM (2009) Turnover intention and emotional exhaustion “at the top”: adapting the job demands-resources model to leaders of addiction treatment organizations. J Occup Health Psychol 14(1):84–95. https://doi.org/10.1037/a0013822

Kramer A, Son J (2016) Who Cares about the Health of Health Care Professionals? An 18-Year Longitudinal Study of Working Time, Health, and Occupational Turnover. ILR Rev 69(4):939–960. https://doi.org/10.1177/0019793916640492

Kristensen TS (2005) Intervention studies in occupational epidemiology. Occup Environ Med 62(3):205–210. https://doi.org/10.1136/oem.2004.016097

Kuechler A, Seubert C, Rehfues E, Elm E von, Nowak D, Glaser J (2019) Human resource management training of supervisors for improving health and well-being of employees. Cochrane Database Syst Rev 9:CD010905

Kuoppala J, Lamminpää A, Liira J, Vainio H (2008) Leadership, job well-being, and health effects—a systematic review and a meta-analysis. J Occup Environ Med 50(8):904–915. https://doi.org/10.1097/JOM.0b013e31817e918d

S, Rowold J (2019) Mindful leadership: Evaluation of a mindfulness-based leader intervention. Gruppodynamik 50(3):319–335. https://doi.org/10.1007/s11612-019-00482-0

Maslach C, Jackson SE, Leiter MP (1996) Maslach burnout inventory manual, 2nd edn. Consulting Psychologists Press, Palo Alto

Milligan-Saville JS, Tan L, Gayed A, Barnes C, Madan I, Dobson M, Bryant RA, Christensen H, Mykletun A, Harvey SB (2017) Workplace mental health training for managers and its effect on sick leave in employees: a cluster randomised controlled trial. Lancet Psychiatry 4(11):850–858. https://doi.org/10.1016/S2215-0366(17)30372-3

Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 6(7):e1000097. https://doi.org/10.1371/journal.pmed.1000097
Mohr G, Müller A, Rigotti T, Aycan Z, Tschan F (2006) The assessment of psychological strain in work contexts. Eur J Psychol Assess 22(3):198–206. https://doi.org/10.1027/1015-5759.22.3.198

Mohr G, Wolfram H-J (2010) Stress among managers: the importance of dynamic tasks, predictability, and social support in unpredictable times. J Occup Health Psychol 15(2):167–179. https://doi.org/10.1037/a0018892

Montano D, Reeske A, Franke F, Hüffmeier J (2017) Leadership, followers' mental health and job performance in organizations: A comprehensive meta-analysis from an occupational health perspective. J Organ Behav 38(3):327–350. https://doi.org/10.1002/job.2124

Nielsen K, Taris TW, Cox T (2010) The future of organizational interventions: Addressing the challenges of today’s organizations. Work Stress 24(3):219–233. https://doi.org/10.1080/02678373.2010.519176

Nyberg A, Alfredsson L, Theorell T, Westerlund H, Vahtera J, Kivimäki M (2009) Managerial leadership and ischaemic heart disease among employees: the Swedish WOLF study. Occup Environ Med 66(1):51–55. https://doi.org/10.1136/oem.2008.039362

Rudolph CW, Murphy LD, Zacher H (2020) A systematic review and critique of research on “healthy leadership”. Leadersh Q 31(1):101335. https://doi.org/10.1016/j.leaqua.2019.101335

Rump J, Schiedhelm M, Eilers S (2016) Gesundheit anordnen? Die Rolle der Führungskultur im Rahmen des Betrieblichen Gesundheitsmanagements. In: Badura B, Dücki A, Schröder H, Klose J, Meyer M (eds) Fehlzeiten-Report 2016: Unternehmenskultur und Gesundheit - Herausforderungen und Chancen, Fehlzeiten-Report, vol 2016. Springer, Berlin, Heidelberg, pp 95–103

Schaufler WB, Salanova M, González-romá V, Bakker AB (2002) The Measurement of Engagement and Burnout: A Two Sample Confirmatory Factor Analytic Approach. J Happiness Stud 3(1):71–92. https://doi.org/10.1023/A:1015630930326

Schelvis RMC, Wiezer NM, Blatter BM, van Genabeek JAGM, Oude Hengel KM, Bohlmeijer ET, van der Beek AJ (2016) Evaluating the implementation process of a participatory organizational level occupational health intervention in schools. BMC Public Health 16(1):1212. https://doi.org/10.1186/s12889-016-3869-0

Schyns B, Schilling J (2013) How bad are the effects of bad leaders? A meta-analysis of destructive leadership and its outcomes. Leadersh Q 24(1):138–158. https://doi.org/10.1016/j.leaqua.2012.09.001

Shimomitsu T, Haratani T, Nakamura K, Kawakami N, Hayashi T, Hiro H et al (2000) The Final development of the Brief Job Stress Questionnaire mainly used for assessment of the individuals. In: Kato M (ed) Ministry of Labor sponsored grant for the prevention of work-related illness: The 1999 report. Tokyo Medical University, Tokyo, pp 126–164

Skakon J, Nielsen K, Borg V, Guzman J (2010) Are leaders’ wellbeing, behaviours and style associated with the affective wellbeing of their employees? A systematic review of three decades of research. Work Stress 24(2):107–139. https://doi.org/10.1080/02678373.2010.495262

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.