



DECREASING WORK STRESS IN TEACHERS

Roosmarijn M.C. Schelvis



DECREASING WORK STRESS IN TEACHERS

Roosmarijn M.C. Schelvis

The studies presented in this thesis were conducted at the Netherlands Organization of Applied Scientific Research (TNO), in collaboration with VU University Medical Center as part of the Body@Work cooperation between these institutions. The intervention study was funded by The Netherlands Organization for Health Research and Development (ZonMw) and the SOM. The grant was awarded in the round of subsidies on 'Sustainable Employability – effectiveness studies', and embedded in the research program 'Prevention' (dossier number 50-51400-98-019). The content of this thesis was not influenced by any kind of sponsorship or monetary contribution.

Financial support for the printing of this thesis was kindly provided by Body@Work, Research Center on Physical Activity, Work and Health.

VRIJE UNIVERSITEIT

Decreasing work stress in teachers

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. V. Subramaniam,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Geneeskunde
op dinsdag 7 november 2017 om 9.45 uur
in de aula van de universiteit,
De Boelelaan 1105

English title: Decreasing work stress in teachers
Dutch title: Werkstress verminderen bij docenten
ISBN-number: 978-94-028-0737-0

Design by: Lyanne Tonk (Persoonlijk Proefschrift)
Printed by: Ipskamp Printing, Amsterdam

© Copyright 2017, Roosmarijn M.C. Schelvis, The Netherlands
All rights reserved. No part of this thesis may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without prior written permission from the author, or when appropriate, from the publishers of the publications.

door
Roosmarijn Margaretha Cornelia Schelvis
geboren te Heemskerk

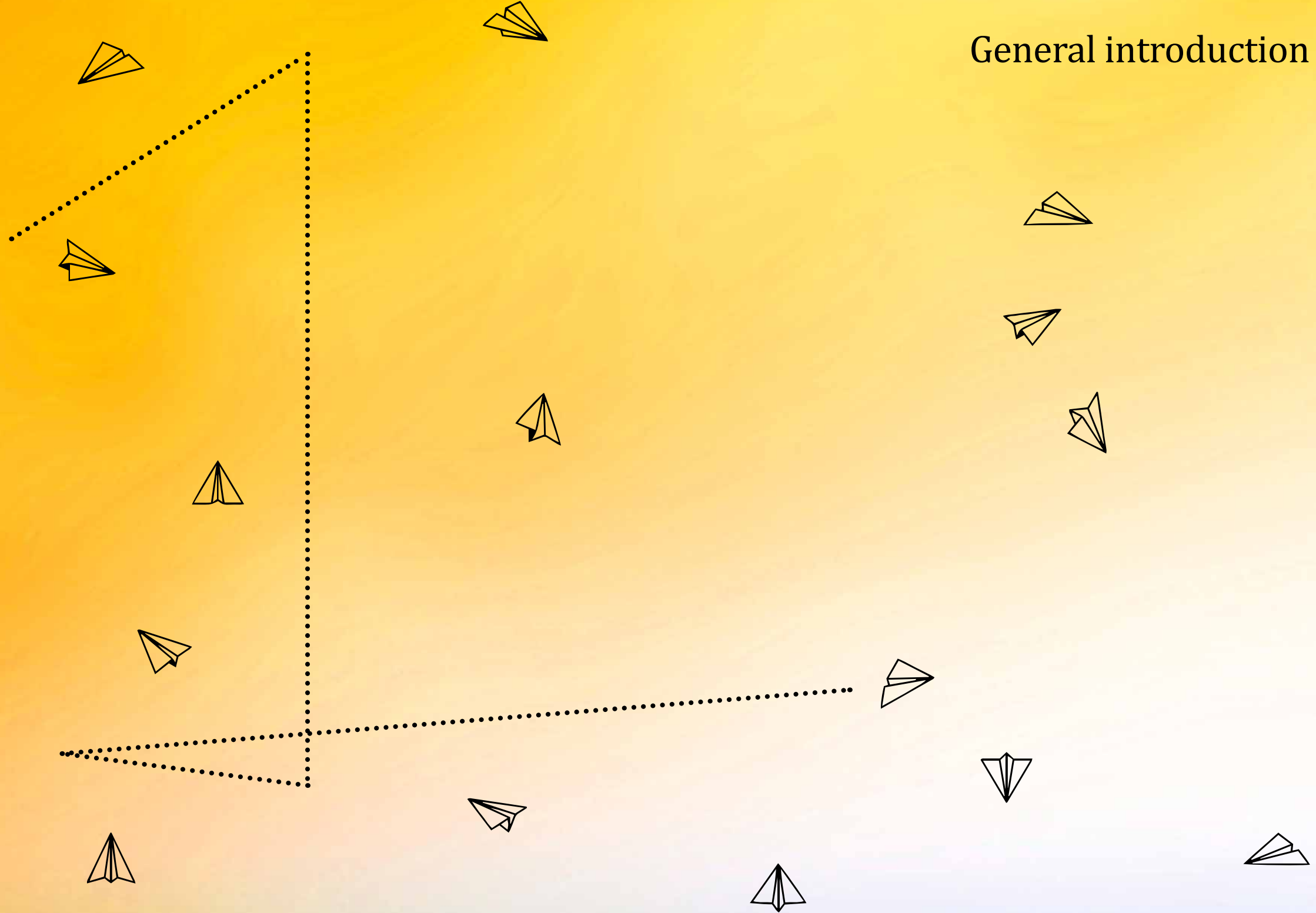
promotor: prof.dr. A.J. van der Beek
copromotoren: dr. N.M. Wiezer
dr. K.M. Oude Hengel

Voor Inge en Cees,
de beste docenten die ik ken

Contents

Chapter 1	General introduction	9
	AN INDIVIDUAL PERSPECTIVE	
Chapter 2	Interplay between mastery and work-related factors in influencing depression and work engagement: a three-wave longitudinal study among older teachers Submitted	27
	AN ORGANIZATIONAL PERSPECTIVE	
Chapter 3	Design of the Bottom-up Innovation project - a participatory, primary preventive, organizational level intervention on work-related stress and well-being for workers in Dutch vocational education. BMC Public Health. 2013;13(760):1-15	53
Chapter 4	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):1-14	83
Chapter 5	Evaluating the implementation process of a participatory organizational level occupational health intervention in schools BMC Public Health. 2016;11(1212):1-20	109
	AN EVALUATION PERSPECTIVE	
Chapter 6	Evaluation of occupational health interventions using a randomized controlled trial: challenges and alternative research designs Scandinavian Journal of Work, Environment & Health. 2015;41(5):491-503	147
Chapter 7	Process variables in organizational stress management intervention evaluation research: a systematic review Scandinavian Journal of Work, Environment & Health. 2016;42(5):371-381	171
Chapter 8	General discussion	191
	Summary	217
	Samenvatting	223
	About the author	233
	List of publications	235
	Dankwoord	238

General introduction



General introduction

“We want to rank among the top five of the world” ([1], p.17). This is not the Dutch national soccer coach speaking, it was the ambition as formulated by the Dutch government at the time of conducting this thesis, in the education paragraph of their coalition agreement. Good education supposedly ensures the competitive power of the economy. Education in the Netherlands is already of high standards, as evidenced by the top 10 notation for mathematics, science and reading skills in the influential OECD Programme for International Student Assessment [2]. However, to transform ‘good education’ into ‘excellent education’, the teacher and school organization are of quintessential importance [3]. If the general level of teaching is improved, the general level of education will improve [4, 5]. And the level of teaching will more easily improve if the school organization is functioning well [6]. However, a major threat to the improvement of teaching, is the high level of work stress among teachers [7]. This thesis aims to explore ways to decrease work stress in teachers, thereby contributing to the realization of the government’s ambition.

The topic is introduced in this chapter by a description of the prevalence and consequences of work stress (paragraph **‘Prevalence of work stress in education’**). In order to successfully find ways to decrease work stress, it is necessary to outline the several definitions of work stress, because the definition also determines the character and scope of interventions (paragraph **‘Theory and definition of work stress’**). Furthermore, existing interventions should be taken into account in the exploration of ways to decrease work stress in teachers (paragraph **‘Existing interventions for work stress’**). The relevance of approaching the work stress problem from an individual, organizational and (intervention) evaluation perspective is reasoned. This chapter concludes with the aim and outline of the current thesis.

Prevalence of work stress in education

One third of the workers in European Union countries reports a high work intensity, which is related to work stress [8]. Work stress is especially common among workers in education [7]. More specifically, stress levels of teachers more than doubled (42%) those found in other occupations (20% [9]). Also in the Netherlands the educational sector is front runner in work stress. According to a representative survey almost one in five employees in education suffers from work stress compared to one in eight in the general working population [10]. These employees feel emotionally drained and exhausted by their work, especially at the end of the work day. They also feel tired when they get up in the morning and are confronted with their work [11].

Across occupations work stress seems to result in several health problems, negative organizational outcomes, and economic costs. Firstly, work stress can result in a range of mental health problems including burnout [12], depression [13, 14], and anxiety [15]. With regard to physical health, research has

convincingly demonstrated that work stress can result in coronary heart disease years later [16]. In the shorter term, work stress might also lead to a range of health problems such as musculoskeletal problems [17], specifically, of the low back, neck, shoulder and forearm [18], gastrointestinal problems [19], and headache disorder (i.e. migraine, severe headaches [20]). Secondly, work stress might impact organizational outcomes negatively. For example, work stress has been related to decreased productivity (in terms of presenteeism; [21]) and work performance [22]. Also relations have been found with (long term) sickness absence from work [23, 24], and with turnover to another job within the same occupation as well as to other occupations [25]. Furthermore, longitudinal evidence from a study among older workers convincingly demonstrated that psychological health problems predict unemployment and early retirement [26]. Thirdly, the consequences of work stress can be regarded in terms of societal costs. Sickness absence due to work stress sum up to a reduced employability of the work force, which is costly. Work days lost due to presenteeism and sickness absence associated with mental health problems summed up to 2.7 billion Euros in 2008 in the Netherlands alone [27, 28].

These three types of work stress consequences all seem to be present in the educational sector as reflected in the finding that for 7% of the workers in Dutch education work stress results in being overworked or burned out, including long term sickness absence [10]. The high level of work stress among teachers is especially challenging because of the ageing working population. A little over half of the employees in education in the Netherlands are aged 45 years or older (51%; [10]), compared to 43% in the general working population. Employees in education are not only older in general, the outflow of the occupation is also more prominent than in the general working population. Many teachers retire before reaching the official retirement age [29], whereas approximately half of all novice teachers leave the sector within the first five years [6].

In sum, work stress has substantial consequences, which are especially alarming in the light of an already shrinking workforce. Some of these consequences might be prevented if adequate measures are taken. Formulating adequate measures starts with a clear conceptualization of work stress and its causes.

Theory and definition of work stress

Since the first introduction of the term ‘stress’ in science in 1936, the concept gradually found its way to the spoken language. The term is now often used by the general public to describe a range of symptoms, feelings, states, causes or consequences. This diverse use of the term resembles the scientific search for a definition and theorization of ‘stress’.

According to the first stress theory, stress was considered a psychological or physiological response to a threatening situation [30]. A threatening situation can be any external stimulus, for example a biological agent, an environmental

condition or event and is often referred to as 'stressor'. The bodily response to a stressor, also named *General Adaptation Syndrome* (GAS), is characterized by three states: the alarm state, the resistance state, and the exhaustion state. The alarm state lasts shortly, it is characterized by physiological changes that prepare the body to show a stress response, such as freeze, fight or flight. In the resistance state, the body tries to cope with the threat, thereby gradually depleting its resources. In the third state, either recovery or exhaustion takes place. If recovery occurs, the body's adaptation system was successful in adapting to the stressor, and the body returns to normal functioning, whereas the opposite is the case in the exhaustion state. In case the exhaustion state endures, the body can be damaged, possibly resulting in physical and mental illness.

In the following years, the biologically oriented definition of stress as a response to the environment, was extended by psychological insights based on theories such as the *Appraisal Theory* [31], the *Michigan model* [32], the *Person-Environment-Fit model* [33]. These theories have in common that the stress response is considered the result of an interaction between an individual and the environment. More specifically, the environment is the source of a stressor, the individual employee is the place where stress effects become visible. It was assumed that the individual appraisal or perception of a stressor determines the scope and duration of a stress response. Applied to the work context, this would mean that stress occurs if an employee cannot meet the demands posed on him/her by the environment and the employee perceives this as threatening.

This work paved the way for the most influential model as proposed by Karasek [34], the *Job Demand-Control model* (JD-C model). The model assumes that perceived job demands such as a high workload, a high work rate or emotionally demanding tasks are not stressors per se, but only if they coincide with a lack of control over the work, for instance due to poor decision latitude. Demands by the work and control over the work are considered dimensions that can be either low or high, resulting in four quadrants: active jobs (high demands, high control), passive jobs (low demands, low control), high strain jobs (high demands, low control), low strain jobs (low demands, high control). Each quadrant represents a different risk for stress and its consequences. The most desirable situation is the 'active job', because it is assumed to increase motivation and learning on the job. The least desirable situation is the 'high job strain', since it is likely to lead to psychological strain and physical illness.

In testing the JD-C-model, social support was discovered as an important additional work characteristic for the occurrence of work stress [35]. This dimension was added and the model was renamed *Job Demand Control-Support model* (JD-CS model). Social support of colleagues and supervisor is considered an accelerator of positive and negative effects: good social support stimulates the positive effects of high job demands and high control on the one hand, and if social support is absent the negative effects of too high job demands without control will be larger on the other. In the JD-C(S) model, individual factors are disregarded in order to avoid the inclination that an individual employee is held responsible for both his or her experience of work stress as well as for the solution of work stress.

Based on these earlier insights the *Job Demands-Resources-model* (JD-R model) [22, 36] was formulated. In line with earlier models, the JD-R model assumes that the balance between positive and negative work characteristics (i.e. job resources and job demands) determines whether positive or negative work-related outcomes occur (e.g. work engagement and burnout, respectively). The model differs from its predecessors in the assumption that any work characteristic can be a potential demand or resource, instead of proposing a set of predetermined, specific positive and negative characteristics [37]. Job demands are generally considered the physical, social or organizational aspects of the job that require sustained physical or psychological effort [36]. Job demands are related to the work content (e.g. workload, work schedule) or work context (e.g. organizational culture) [38]. Job resources on the other hand are the physical, social or organizational aspects of the job that may reduce job demands, help to achieve goals and stimulate learning and development [36]. Examples of job resources are social support by colleagues or supervisor, or decision latitude. Even though any work characteristic can be a demand or resource according to the JD-R model, it has been demonstrated that patterns of common job demands and job resources exist across and within occupational groups [39]. And therefore work stress interventions are probably most successful if they are tailored to a specific occupational sector or group. In sum, determinants of work stress differ as a function of occupation and should be taken into account in the design of workplace interventions.

The most recent versions of the JD-R model furthermore differ from the JD-C(S) model in the incorporation of individual factors. Since the stress response is considered the result of an interaction between the individual and the environment, incorporating individual factors might help explain the occurrence and course of work stress. These factors are mostly known as 'personal resources' and defined as: "the psychological characteristics or aspects of the self that are generally associated with resiliency and that refer to the ability to control and impact one's environment successfully" [37]. In line with the definition of job resources, personal resources are presumed to help in achieving goals and to stimulate learning and development. Examples are self-efficacy, self-esteem, or intrinsic work motivation. According to a critical review of the JD-R model by Schaufeli and Taris, personal resources have to date been integrated in the model in five ways [37]. First, as a direct influence on work stress (e.g. [40]). Second, as a moderator between job demands/job resources and work stress (e.g. [41]). Third, as a mediator in the relation between job demands/job resources and work stress (e.g. [42]). Fourth, as an antecedent of job demands/job resources [43]. And fifth, as a confounding variable (e.g. [44]). Nevertheless, it seems relevant to include personal resources in the exploration of ways to decrease work stress.

Somewhat parallel to the development of the JDC, the JD-CS, and the JD-R model, another line of stress research focused on recovery [45, 46]. Recovery is defined as "a process of psychophysiological unwinding after effort expenditure" (p. 482, [46]) and is considered a central element in the stress process. Recovery after work seems all the more relevant when recovery during work is insufficient. If

physiological activation continues after work and recovery is insufficient, evidence indicates that this will eventually result in chronic health impairment [46].

In this thesis on decreasing work stress in teachers, the JD-R model was the leading theoretical framework, because of its generic applicability and widespread use. The definition of work stress by the International Labour Organization (ILO) was used, which is in line with the JD-R model: *[work]* “stress is the harmful physical and emotional response caused by an imbalance between the perceived demands and the perceived resources and abilities of individuals to cope with those demands” *[i.e. personal resources]* ([47], italics by current author). In order to decrease work stress in organizations, job demands, job resources, and personal resources seem to be the starting point for interventions, conducted within a specific sector.

Existing interventions for work stress

Primary, secondary and tertiary prevention

Work stress interventions are often classified by the aim one has with the intervention, often labeled as primary prevention (prevent work stress before it ever occurs), secondary prevention (reduce impact of occurring work stress), or tertiary prevention (treat the consequences of occurring work stress, such as cardiovascular disease). Examples of primary preventive interventions are: job redesign in order to maintain a balance between demands and resources, or the enhancement of social support (e.g. [48]). Secondary preventive interventions on the other hand could be cognitive behavioral therapy sessions wherein problem solving skills are enhanced (e.g. [49]) or coaching sessions wherein coping skills are enhanced (e.g. [50]). An illustration of tertiary preventive interventions could be counseling or return-to-work programs after sickness absence due to work stress (e.g. [51]).

An occupational health principle with regard to interventions is the ‘hierarchy of (hazard) controls’ [52, 53]. The proposition of the principle is that methods that eliminate or substitute a stressor (i.e. prevention through design) are to be preferred over methods that protect workers from the stressor (e.g. personal protective equipment). Eliminating or substituting the stressor is believed to result in more sustainable effects than protecting workers from the stressor [54]. Eliminating or substituting stressors requires a change in the work environment and work organization, which can be done by conducting organizational-level interventions [54].

Organizational level and individual level interventions

Interventions aiming to change the work environment and work organization in order to decrease work stress are often labeled ‘organizational level interventions’ [55], as opposed to interventions targeting (personal resources of) individual employees, which have been termed ‘individual level interventions’. In the field of work stress the proposition of the ‘hierarchy of controls’ principle – translated as organizational level interventions producing more sustain-

able effects over individual level interventions - could thus far not be fully supported by empirical evidence. Two meta-analyses consistently demonstrated significant effects on health outcomes by individual level interventions, while the results of organizational level interventions yielded inconsistent findings [56, 57]. More specifically, Van der Klink and colleagues conducted a meta-analysis of 48 stress management interventions and found, in line with earlier qualitative reviews, that stress management interventions are effective in reducing stress. However, the strength of the effect differed by the type of intervention applied [56]. That is, the individual interventions (i.e. cognitive behavioral interventions, relaxation techniques, and multimodal programs to enhance passive and active coping) had a moderate to small effect size, whereas organizational level interventions were not effective. In an update of this meta-analysis seven years later, Richardson and Rothstein came largely to the same conclusion based on 36 studies, although the overall weighted effect size was somewhat larger [57]. Cognitive behavioral interventions again showed the largest effect on stress-related outcomes and organizational level interventions showed no significant effects.

On the other hand, there is some evidence of the relevance of organizational level interventions for work stress. Lamontagne and colleagues rated the 90 work stress interventions they included in their review [54]. A high rating was given if the intervention targeted both the organization and the individual, compared to moderate ratings (organization only) and low ratings (individual only). They found that the highly rated interventions were most effective in addressing the organizational and individual consequences of work stress. Another review of 39 organizational level work stress interventions found that the odds ratio of finding effects was higher in the more comprehensive interventions, addressing material, organizational and work-time issues at the same time [58]. Lastly, evidence suggests that some of the included interventions improved health outcomes [59] or business outcomes such as decreased absenteeism and staff turnover [60].

The inconclusiveness with regard to organizational level interventions for work stress could be due to several reasons. Firstly, the interventions might not be the right ones, because the theory on which the intervention is built is (partially) erroneous (i.e. theory failure)[61]. According to Kristensen, in an intervention process the intended occupational health intervention is presumed to lead to an intended reduction in exposure, which will lead to better health outcomes. In the case of theory failure, the intervention was implemented as intended, but the reduction in exposure and the improvement in health did not take place. In order to prevent theory failure in future intervention studies, more knowledge on the individual and organizational determinants of work stress for specific occupational groups is needed. Secondly, the organizational level interventions might also not be implemented as intended (program or implementation failure)[61]. For example because of lacking managerial support or because external events interfered with the intervention program. In order to prevent implementation failure, more knowledge on the implementation process of organizational level interventions is needed,

for example by conducting process evaluations of these type of interventions [62]. Even though the importance of the process evaluation as a relevant tool for assessing implementation is increasingly recognized over the last decade [63], there is no consensus on which process variables should be assessed in work stress interventions in order to grasp the often complex implementation process [64, 65]. A third explanation for the inconclusive evidence with regard to the effectiveness of organizational level interventions for work stress, could be that the study design was not suitable for detecting results [66]. In order to assess whether occupational health interventions decreased work stress effect evaluations are conducted. In intervention evaluation research, the randomized controlled trial (RCT) is the preferred research design (a 'gold standard') because of the possibility to draw causal inferences about the effects of the intervention under study. However, in the occupational setting practical and ethical challenges might exist that hamper the (correct) application of this design [67]. In intervention research in education this is evidenced by the results from a Cochrane review: the few studies that found effects of organizational interventions on well-being of teachers all were of low methodological quality [68]. There is thus a clear need within occupational health research for alternative research designs, which allow for (some degree of) causal inference. This underlines the relevance of the (intervention) evaluation perspective in this thesis wherein, amongst others, alternative research designs are explored, in order to ultimately decrease work stress in teachers.

Aim

As described in the previous paragraphs the evidence on the most effective ways to decrease work stress in teachers is inconclusive. More specifically, more theory-based knowledge on the individual and organizational determinants of work stress for specific occupational groups is needed. Furthermore, the research could be improved methodologically. The main objective of this thesis is therefore to explore ways to decrease work stress by combining the individual, organizational and intervention evaluation perspective, in the specific occupational group of teachers. The combination of evidence from these combined perspectives is believed to provide more insight into the main research question, i.e. how can we decrease work stress in teachers, than any of the three perspectives alone. Each perspective corresponds with a key question that will be addressed in this thesis:

1. How can we decrease work stress in teachers from an individual perspective?
2. How can we decrease work stress in teachers from an organizational perspective?
3. How can we gather the most relevant evidence in intervention studies in the occupational setting, for example to decrease work stress?

Outline of the thesis

How can we decrease work stress in teachers from an individual perspective?

In part one, the individual perspective is addressed by assessing the role of a personal resource in the decrease of work stress in a cohort of older teachers (**chapter 2**). More specifically, the interplay was explored between mastery and work-related factors (i.e. job demands and job resources) in influencing work stress related outcomes (i.e. depression and work engagement). Mastery was hypothesized to mediate the longitudinal effects of job demands and job resources on depression and engagement. For this chapter longitudinal data from the Study on Transitions in Employment, Ability and Motivation (STREAM) were used, which is a four-year longitudinal cohort study among older persons (aged 45-64 years) in the Netherlands [69].

How can we decrease work stress in teachers from an organizational perspective?

In part two, the organizational perspective is addressed by the evaluation of a practice-based, participatory prevention program for employees in schools. The evaluation was conducted within a controlled trial among 356 employees from two secondary vocational schools in the Netherlands. It was hypothesized that the prevention program would decrease need for recovery and increase vitality primarily. Several secondary outcomes relevant in relation to work stress were assessed as well (i.e. psychological job demands, decision authority, social support, work ability, job satisfaction, commitment, work engagement, occupational self-efficacy, and organizational efficacy). The prevention program and the study design are described in **chapter 3**. Whether implementation of the prevention program was successful was assessed using a comprehensive theoretical framework [70]. The framework included components with regard to the intervention design and implementation, the context, and mental models of the participants. A detailed evaluation of the implementation process is given in **chapter 4**. Whether the prevention program rendered the hypothesized effects is described in **chapter 5**, by comparing the effects in the intervention group to those in the control group on the primary and secondary outcomes.

How can we gather the most relevant evidence in intervention studies?

Part three of this thesis, the evaluation perspective on decreasing work stress by means of interventions, comprises a systematic review on process variables (**chapter 6**) and a narrative review on study designs for effect evaluations (**chapter 7**). The systematic literature review aimed to explore which process variables are used in evaluations of interventions to decrease work stress. The

narrative review on study designs for effect evaluations describes challenges in applying the RCT design in intervention studies in the occupational setting, and provides an overview of alternative observational and experimental study designs for the evaluation of occupational health interventions.

This thesis concludes with a summary of the main findings and a discussion of implications for practice and research for decreasing work stress (**chapter 8**).

References

- Rutte M, Samsom D: **Van goed naar excellent onderwijs Hoofdstuk VI**. In *Bruggen slaan, regeerakkoord VVD en PvdA*. Edited by Rutte M, Samsom D. 2012:16-17.
- PISA O: **PISA 2012 Results in Focus: What 15-year-olds know and what they can do with what they know**. 2014-12-03].<http://www.oecd.org/pisa/keyfindings/pisa-2012-results-overview.pdf> 2014.
- Hattie JAC: *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York: Routledge; 2009.
- Wright SP, Horn SP, Sanders WL: **Teacher and classroom context effects on student achievement: Implications for teacher evaluation**. *Journal of personnel evaluation in education* 1997, **11**(1):57-67.
- Wayne AJ, Youngs P: **Teacher characteristics and student achievement gains: A review**. *Review of Educational research* 2003, **73**(1):89-122.
- Pas ET, Bradshaw CP, Hershfeldt PA: **Teacher- and school-level predictors of teacher efficacy and burnout: Identifying potential areas for support**. *J School Psychol* 2012, **50**(1):129-145.
- Klassen R, Wilson E, Siu AFY, Hannok W, Wong MW, Wongsri N, Sonthisap P, Pibulchol C, Buranachaitavee Y, Jansem A: **Preservice teachers' work stress, self-efficacy, and occupational commitment in four countries**. *European Journal of Psychology of Education* 2013, **28**(4):1289-1309.
- Eurofound: **Sixth European Working Conditions Survey – Overview report**. 2016 :.
- Smith A, Brice C, Collins A, McNamara R, Matthews V: **Scale of occupational stress: a further analysis of the impact of demographic factors and type of job**. 2000, **Contract Research Report 311/2000**.
- van Zwieten MHJ, de Vroome EMM, Mol MEM, Mars GMJ, Koppes LLJ, van denBossche SNJ: *Nationale Enquête Arbeidsomstandigheden 2013: Methodologie en globale resultaten*. [Netherlands Working Conditions Survey 2013: Methodology and overall results]. Hoofddorp: TNO; 2014.
- Schaufeli WB, van Dieren-donck D: *Handleiding van de Utrechtse Burnout Schaal (ubos)*: Lisse: Swets & Zeitlingen; 2000.
- Maslach C, Schaufeli WB, Leiter MP: **Job burnout**. *Annual Review of Psychology* 2001, **52**:397-422.
- Netterstrom B, Conrad N, Bech P, Fink P, Olsen O, Rugulies R, Stansfeld S: **The relation between work-related psychosocial factors and the development of depression**. *Epidemiol Rev* 2008, **30**:118-132.
- Theorell T, Hammarstrom A, Aronsson G, Traskman Bendz L, Grape T, Hogstedt C, Marteinsdottir I, Skoog I, Hall C: **A systematic review including meta-analysis of work environment and depressive symptoms**. *BMC Public Health* 2015, **15**:738-015-1954-4.
- Jensen HK, Wieclaw J, Munch-Hansen T, Thulstrup AM, Bonde JP: **Does dissatisfaction with psychosocial work climate predict depressive, anxiety and substance abuse disorders? A prospective study of Danish public service employees**. *J Epidemiol Community Health* 2010, **64**(9):796-801.
- Kivimäki M, Nyberg ST, Batty GD, Fransson EI, Heikkilä K, Alfredsson L, Bjorner JB, Borritz M, Burr H, Casini A: **Job strain as a risk factor for coronary heart disease: a collaborative meta-analysis of individual participant data**. *The Lancet* 2012, **380**(9852):1491-1497.
- Bongers PM, de Winter CR, Kompier MA, Hildebrandt VH: **Psychosocial factors at work and musculoskeletal disease**. *Scand J Work Environ Health* 1993, :297-312.
- Eijkelhof B, Huysmans M, Garza JB, Blatter B, van Dieën J, Dennerlein J, van der Beek A: **The effects of workplace stressors on muscle activity in the neck-shoulder and forearm muscles during computer work: a systematic review and meta-analysis**. *Eur J Appl Physiol* 2013, **113**(12):2897-2912.
- Drossman DA, Creed FH, Olden KW, Svedlund J, Toner BB, Whitehead WE: **Psychosocial aspects of the functional gastrointestinal disorders**. *Gut* 1999, **45 Suppl 2**:II25-30.
- Santos I, Griep RH, Alves MGM, Goulart A, Lotufo P, Barreto S, Chor D, Benseñor I: **Job stress is associated with migraine in current workers: The Brazilian Longitudinal Study of Adult Health (ELSA-Brasil)**. *European Journal of Pain* 2014, **18**(9):1290-1297.
- Demerouti E, Le Blanc PM, Bakker AB, Schaufeli WB, Hox J: **Present but sick: a three-wave study on job demands, presenteeism and burnout**.

- Career Development International* 2009, **14**(1):50-68.
22. Schaufeli WB, Bakker AB: **Job demands, job resources, and their relationship with burnout and engagement: A multi-sample study.** *J Organ Behav* 2004, **25**(3):293-315.
 23. Bültmann U, Huibers MJ, van Amelsvoort LP, Kant I, Kasl S, Swaen GM: **Psychological distress, fatigue and long-term sickness absence: prospective results from the Maastricht Cohort Study.** *Journal of occupational and environmental medicine* 2005, **47**(9):941-947.
 24. Peterson U, Bergstrom G, Demerouti E, Gustavsson P, Asberg M, Nygren A: **Burnout levels and self-rated health prospectively predict future long-term sickness absence: a study among female health professionals.** *J Occup Environ Med* 2011, **53**(7):788-793.
 25. De Croon EM, Sluiter JK, Blonk RW, Broersen JP, Frings-Dresen MH: **Stressful work, psychological job strain, and turnover: a 2-year prospective cohort study of truck drivers.** *J Appl Psychol* 2004, **89**(3):442.
 26. Leijten FR, de Wind A, van den Heuvel SG, Ybema JF, van der Beek AJ, Robroek SJ, Burdorf A: **The influence of chronic health problems and work-related factors on loss of paid employment among older workers.** *J Epidemiol Community Health* 2015, **69**(11):1058-1065.
 27. De Graaf R, Tuithof M, Dorsselaer S: *Verzuim door psychische en somatische aandoeningen bij werkenden. resultaten van de 'Netherlands Mental Health Survey and Incidence Study-2' (NEMESIS-2)*: Utrecht: Trimbos Instituut; 2011.
 28. De Graaf R, Tuithof M, Van Dorsselaer S, Ten Have M: **Comparing the effects on work performance of mental and physical disorders.** *Soc Psychiatry Psychiatr Epidemiol* 2012, **47**(11):1873-1883.
 29. Seibt R, Lützkendorf L, Thinschmidt M: **Risk factors and resources of work ability in teachers and office workers.** *International Congress Series* 2005, **1280**:310-315.
 30. Selye H: **A syndrome produced by diverse noxious agents.** *Nature* 1936, **138**(3479):32.
 31. Lazarus RS, Folkman S: *Stress, appraisal, and coping*: Springer publishing company; 1984.
 32. Kahn RL, Wolfe DM, Quinn RP, Snoek JD, Rosenthal RA: **Organizational stress: Studies in role conflict and ambiguity.** 1964,.
 33. French JR, Rodgers W, Cobb S: **Adjustment as person-environment fit.** *Coping and adaptation* 1974, :316-333.
 34. Karasek Jr RA: **Job demands, job decision latitude, and mental strain: Implications for job redesign.** *Adm Sci Q* 1979, :285-308.
 35. Johnson JV, Hall EM: **Job strain, work place social support, and cardiovascular disease: a cross-sectional study of a random sample of the Swedish working population.** *Am J Public Health* 1988, **78**(10):1336-1342.
 36. Demerouti E, Nachreiner F, Bakker AB, Schaufeli WB: **The job demands-resources model of burnout.** *J Appl Psychol* 2001, **86**(3):499-512.
 37. Schaufeli WB, Taris TW: **A Critical Review of the Job Demands-Resources Model: Implications for Improving Work and Health.** In *Bridging Occupational, Organizational and Public Health: A Transdisciplinary Approach*. Edited by Bauer GF, Hämmig O. Dordrecht: Springer Science+Business Media; 2014:43.
 38. Cox T, Griffiths A, Barlow C, Randall R, Thompson T, Rial-Gonzalez E: *Organisational interventions for work stress: a risk management approach*: [Sudbury]: HSE Books; 2000.
 39. Johnson S, Cooper C, Cartwright S, Donald I, Taylor P, Millet C: **The experience of work-related stress across occupations.** *J Manage Psychol* 2005, **20**(2):178-187.
 40. Lorente Prieto L, Salanova Soria M, Martínez Martínez I, Schaufeli W: **Extension of the Job Demands-Resources model in the prediction of burnout and engagement among teachers over time.** 2008, **20**(3):354-360.
 41. Brenninkmeijer V, Demerouti E, le Blanc PM, Van Emmerik H: **Regulatory focus at work: The moderating role of regulatory focus in the job demands-resources model.** *Career Development International* 2010, **15**(7):708-728.
 42. Xanthopoulou D, Bakker AB, Demerouti E, Schaufeli WB: **The role of personal resources in the job demands-resources model.** *International Journal of Stress Management* 2007, **14**(2):121-141.
 43. Consiglio C, Borgogni L, Alessandri G, Schaufeli WB: **Does self-efficacy matter for burnout and sickness absenteeism? The mediating role of demands and resources at the individual and team levels.** *Work Stress* 2013, **27**(1):22-42.
 44. Bakker AB, Boyd CM, Dollard M, Gillespie N, Winefield AH, Stough C: **The role of personality in the job demands-resources model: A study of Australian academic staff.** *Career Development International* 2010, **15**(7):622-636.
 45. Meijman T, Mulder G: **Psychological aspects of workload.** In *Handbook of work and organizational psychology. Volume 2.* 2nd edition. Edited by Drenth P, Thierry H, de Wolff C. Hove (England): Psychology Press Ltd; 1998:5-33.
 46. Geurts SA, Sonnentag S: **Recovery as an explanatory mechanism in the relation between acute stress reactions and chronic health impairment.** *Scand J Work Environ Health* 2006, :482-492.
 47. International Labour Organization: **Workplace stress: A collective challenge.** 2016, **978-92-2-130642-9**.
 48. Bourbonnais R, Brisson C, Vezina M: **Long-term effects of an intervention on psychosocial work factors among healthcare professionals in a hospital setting.** *Occup Environ Med* 2011, **68**(7):479-486.
 49. Lexis MA, Jansen NW, Huibers MJ, van Amelsvoort LG, Berkouwer A, Tjin A Ton G, van den Brandt PA, Kant I: **Prevention of long-term sickness absence and major depression in high-risk employees: a randomised controlled trial.** *Occup Environ Med* 2011, **68**(6):400-407.
 50. Duijts SF, Kant I, van den Brandt PA, Swaen GM: **Effectiveness of a preventive coaching intervention for employees at risk for sickness absence due to psychosocial health complaints: results of a randomized controlled trial.** *J Occup Environ Med* 2008, **50**(7):765-776.
 51. Rebergen DS, Bruinvels DJ, Bos CM, van der Beek, Allard J, van Mechelen W: **Return to work and occupational physicians' management of common mental health problems-process evaluation of a randomized controlled trial.** *Scand J Work Environ Health* 2010:488-498.
 52. Peterson J: **Principles for controlling the occupational environment. The industrial environment—its evaluation and control.** 1973, 74-117.
 53. NIOSH: **The state of the national initiative on prevention through design.** 2013, **2014-123**.
 54. Lamontagne AD, Keegel T, Louie AM, Ostry A, Landsbergis PA: **A systematic review of the job-stress intervention evaluation literature, 1990-2005.** *Int J Occup Environ Health* 2007, **13**(3):268-280.
 55. Semmer NK: **Job stress interventions and organizational at work.** In *Handbook of occupational health psychology*. Edited by Quick JC, Tetrick LE. Washington: APA; 2003:.
 56. Van der Klink JJJ, Blonk RWB, Schene AH, Van Dijk FJH: **The benefits of interventions for work-related stress.** *Am J Public Health* 2001, **91**(2):270-276.
 57. Richardson KM, Rothstein HR: **Effects of Occupational Stress Management Intervention Programs: A Meta-Analysis.** *J Occup Health Psychol* 2008, **13**(1):69-93.
 58. Montano D, Hoven H, Siegrist J: **Effects of organizational-level interventions at work on employees' health: a systematic review.** *BMC Public Health* 2014, **14**:135-2458-14-135.
 59. Bamba C, Egan M, Thomas S, Petticrew M, Whitehead M: **The psychosocial and health effects of workplace reorganisation. 2. A systematic review of task restructuring interventions.** *J Epidemiol Community Health* 2007, **61**(12):1028-1037.
 60. Bond FW, Flaxman PE, Loivette S: *A business case for the management standards for stress*: Sudbury: Health and Safety Executive; 2006.
 61. Kristensen TS: **Intervention studies in occupational epidemiology.** *Occup Environ Med* 2005, **62**(3):205-210.
 62. Randall R, Griffiths A, Cox T: **Evaluating organizational stress-management interventions using adapted study designs.** *European Journal of Work and Organizational Psychology* 2005, **14**(1):23-41.
 63. Nielsen K, Randall R, Holten A-, González ER: **Conducting organizational-level occupational health interventions: What works?** *Work Stress* 2010, **24**(3):234-259.
 64. Nytrø K, Saksvik PØ, Mikkelsen A, Bohle P, Quinlan M: **An appraisal of key factors in the implementation of occupational stress interventions.** *Work & Stress* 2000, **14**(3):213-225.
 65. Egan M, Bamba C, Petticrew M, Whitehead M: **Reviewing evidence on complex social interventions: appraising implementation in systematic reviews of the health effects of organizational-level workplace interventions.** *J Epidemiol Community Health* 2009, **63**(1):4-11.
 66. Cox T, Karanika M, Griffiths A, Houdmont J: **Evaluating organizational-level work stress interventions:**

- Beyond traditional methods.** *Work & Stress* 2007, **21**(4):348-362.
67. Griffiths A: **Organizational interventions: facing the limits of the natural science paradigm.** *Scand J Work Environ Health* 1999, **25**(6):589-596.
68. Naghieh A, Montgomery P, Bonell CP, Thompson M, Aber JL: **Organisational interventions for improving wellbeing and reducing work-related stress in teachers.** *Cochrane Database Syst Rev* 2015, **April 8**(4):1-70.
69. Ybema JF, Geuskens GA, van den Heuvel, Swenne G, de Wind A, Leijten FR, Joling C, Blatter BM, Burdorf A, van der Beek, Allard J, Bongers PM: **Study on transitions in employment, ability and motivation (STREAM): The design of a four-year longitudinal cohort study among 15,118 persons aged 45 to 64 years.** *Br J Med Med Res* 2014, **4**:1383-99.
70. Nielsen K, Randall R: **Opening the black box: Presenting a model for evaluating organizational-level interventions.** *European Journal of Work and Organizational Psychology* 2013, **22**(5):601-617.

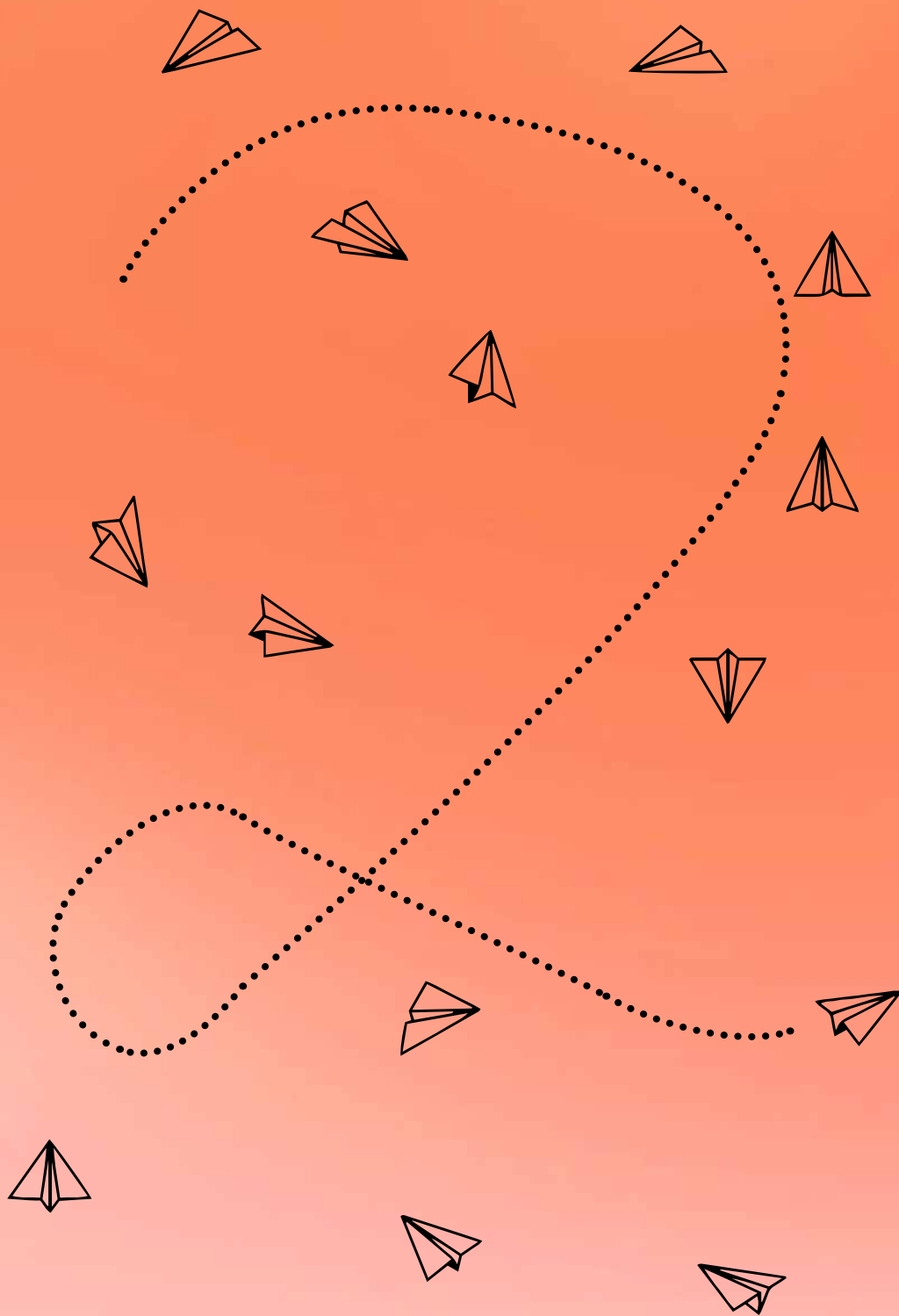
An individual perspective



Interplay between mastery and work-related factors in influencing depression and work engagement: a three-wave longitudinal study among older teachers

Roosmarijn M.C. Schelvis
Fenna R.M. Leijten
Noortje M. Wiezer
Karen M. Oude Hengel
Allard J. van der Beek
Jan Fekke Ybema

Submitted



Abstract

Purpose

The objective of the current study was to explore the interplay between mastery as a personal factor and work-related factors (i.e., job demands and job resources), in influencing changes in depression and work engagement in a longitudinal sample of ageing teachers. Mastery was hypothesized to mediate the longitudinal effects of (psychological and emotional) job demands and job resources (autonomy and social support) on depression and work engagement.

Methods

The study used data from a longitudinal cohort study among older persons (aged 45-64 years) in the Netherlands (STREAM). Data of 549 teachers from the first three waves (years 2010-2012) were included. A path model was built to test the longitudinal relations between all variables.

Results

Psychological job demands contributed to an increase in mastery; emotional job demands contributed to a decrease in mastery. Psychological demands at baseline (T1) were related to both an increase in depression at T3 (via T2 job demands) and a decrease in depression (via T2 mastery), but not to changes in work engagement. Emotional demands (T1) were related to an increase in depression (T3) and a decrease in work engagement (T3) (partially via T2 mastery). Contrary to expectations, job resources were not longitudinally related to mastery nor to depression or work engagement, and therefore no mediation by mastery was found. Finally, mastery was directly related to lower depression and higher work engagement one year later.

Conclusions

Increasing mastery in ageing teachers can be explored further as an intervention strategy to decrease depression and increase work engagement in this occupational group.

Introduction

Throughout the western world the teaching workforce is growing older [1]. In the Netherlands, for example, just over half of the employees in education are aged 45 years or older [2]. Furthermore, many teachers retire before reaching the official retirement age [3], whereas approximately half of all novice teachers leave the sector within the first five years [4]. These tendencies potentially threaten the quality of education and should urge society to keep teachers productive and in good health for longer, also referred to as 'sustainably employable' [5]. A major threat to the sustainable employability of teachers is mental health problems, such as work stress [6]. Stress levels are consistently higher in this occupational group compared to the general working population [7]. In a Scottish sample of teachers who retired early due to ill health, 37% indicated that mental health problems, such as work stress, were the most important reason [8]. An even higher number (45%) was found in a sample of German school principals [9].

In order to maintain or even increase sustainable employability in teachers, it is necessary to better understand the interplay of antecedents and outcomes of mental health. An often used theoretical model to describe this interplay is the Job Demands-Resources (JD-R) model [10]. The JD-R model distinguishes aggravating and alleviating work-related antecedents of mental health, job demands and job resources, respectively. Job demands, such as a high workload or job insecurity, are defined as "those physical, social or organizational aspects of the job that require sustained physical or mental effort and are therefore associated with certain physiological and psychological costs" ([10], p.501). Job resources, such as autonomy or social support from colleagues, are defined as: "(a) functional in achieving work goals; (b) reduce job demands at the associated physiological and psychological costs; (c) stimulate personal growth and development" ([10], p.501). Two review studies have shown that job demands, amongst others psychological job demands and bullying, may contribute to mental health problems, such as depression [11, 12]. Other studies have shown that job resources may contribute to flourishing mental health, such as work engagement [13]. Work engagement is defined as a positive, fulfilling, work-related state of mind that is characterized by vigor, dedication and absorption [14], and it has been shown to be strongly related to mental health [15].

Whereas prior research has often focused on work-related antecedents of mental health [16], the current study assessed associations of both work-related factors and a personal factor (i.e., mastery) related to mental health. Prior findings with regard to personal factors are mixed. A meta-analysis has demonstrated that a range of positive personal factors, such as optimism, locus of control or self-efficacy, were consistently and significantly negatively related to burnout [17]. In a large cross-sectional study among workers of an electrical engineering and electronics company, however, personal factors (i.e., self-efficacy and organization based self-esteem) were found to be a mediator in the

relation between job resources and engagement and exhaustion, but not in the relation between job demands and exhaustion [18]. Other studies have focused on the reciprocity of job resources, indicating the existence of cumulative, ‘positive gain spirals’ [19]. Schaufeli and Taris have argued for more systematically conducted research into the role of personal factors, clarifying whether these are antecedents, moderators, or mediators in the JD-R model [20].

Important personal factors are coping resources, such as mastery, because they are believed to be effective protective factors for stress [21]. Mastery is a generalized self-belief, and it has been conceptualized as “a conviction one is able to control the important circumstances that are currently impinging on one’s life” [22]. Xanthopoulou et al. [18] have described mastery in terms of the JD-R model, stating that employees with many personal resources (or personal factors) also have a higher level of mastery, which enables them to deal effectively with job demands and thus prevents them from experiencing negative mental health outcomes. However, Pearlin (2010) described in an overview article that the association between the belief and the actual control people can exert over their life course is still unclear [22].

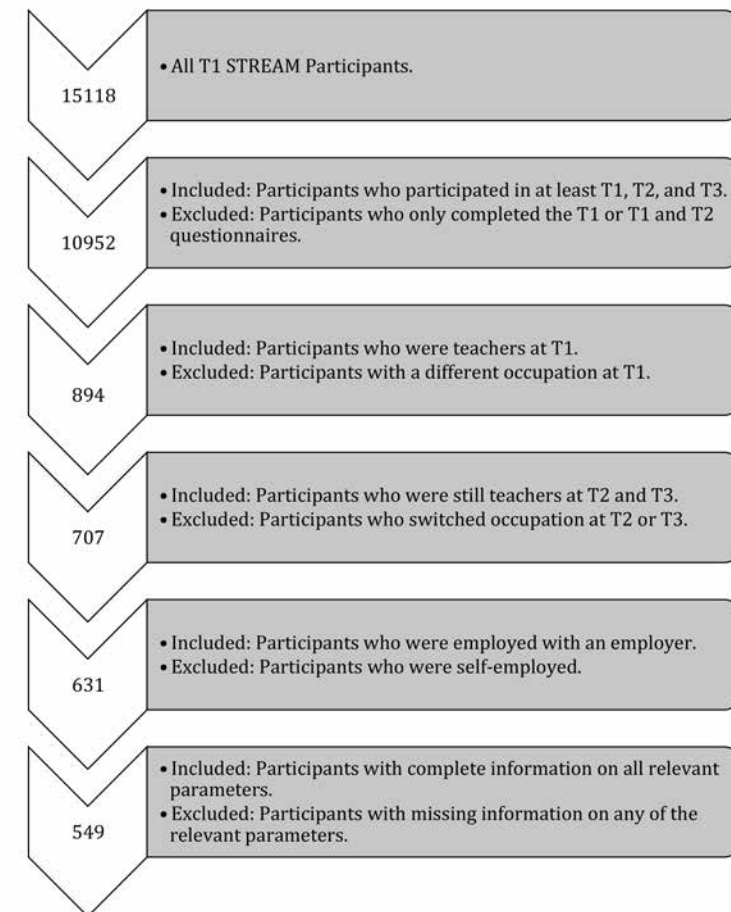
The objective of the current study is to explore the interplay between mastery and work-related factors in influencing positive (work engagement) and negative mental health (depression) in a longitudinal sample of ageing teachers (45 years and older; three time points). To this end, the first aim was to determine the relation between work-related factors, i.e., job demands and job resources, and mastery. It was hypothesized that higher job demands are related to decreased mastery (*hypothesis 1a*), whereas higher job resources are related to increased mastery (*hypothesis 1b*). The second aim was to determine the relation of mastery with depression and with work engagement. It was hypothesized that mastery was related to decreased depression (*hypothesis 2a*) and to increased work engagement (*hypothesis 2b*). The third aim was to determine the relation of work-related factors with depression and work engagement, and whether mastery mediated this relation. It was hypothesized that higher job demands were related to increased depression (*hypothesis 3a*) and to decreased work engagement (*hypothesis 3b*), whereas higher job resources were expected to relate to decreased depression (*hypothesis 3c*) and increased work engagement (*hypothesis 3d*). Mastery was hypothesized to mediate the longitudinal effects of job demands and job resources on depression and engagement (*hypothesis 3e*). In other words, mastery was expected to explain how demands and resources at work influence mental health, i.e., depression and work engagement.

Methods

Study design and sample

This study used data from the Study on Transitions in Employment, Ability and Motivation (STREAM), which is a longitudinal cohort study among older persons (aged 45-64 years) in the Netherlands [23]. The STREAM sample was drawn from an online panel and was stratified according to 5-year age groups and work status (i.e., employed, self-employed, non-employed). Participants annually filled out a questionnaire on a variety of topics including sociodemographic factors, work characteristics, health, and sustainable employability. Data from the first three waves of STREAM (2010-2012) were used for the current study. Participants were included if they had participated in the first three STREAM waves, were teachers and employed via an employer, and had complete information on all relevant parameters. Finally, 549 teachers were included; the participation flow is shown in Figure 1.

Figure 1 — Participation flow



Ethical considerations

The Medical Ethical Committee of the VU University Medical Centre Amsterdam declared that the Medical Research Involving Human Subjects Act (abbreviation in Dutch: WMO) does not apply to STREAM. The Medical Ethical Committee had no objection to the execution of this study. In the information for participants that accompanied the online questionnaire, it was emphasized that their privacy was guaranteed, that all answers to the questions were treated confidentially, and that all data were stored in secured computer systems.

Measures

All measures, except for covariates, were measured at all three waves (T1–T3). For all measures except the covariates, a mean score was calculated across items.

Work-related factors

Two types of work-related factors were distinguished: job demands and job resources. All work-related factors were measured on a five-point scale ranging from (almost) never to always. Two subscales of the Job Content Questionnaire [24] were used to measure *job demands*: psychological job demands (4 items; Cronbach's alpha 0.85) and emotional job demands (3 items; Cronbach's alpha 0.80). Sample items are: 'Do you have to work very fast?' and 'Is your work emotionally demanding?', respectively. Higher scores indicate higher job demands, which is unfavorable. Two types of *job resources* were measured: autonomy and social support from colleagues and supervisor. Higher scores reflect higher resources, which is considered favorable. Autonomy was measured with a 5-item scale based on the Job Content Questionnaire (Cronbach's alpha 0.70) [24]. The scale comprises items such as 'Can you decide for yourself how you conduct your work?'. Social support from colleagues and supervisor was measured with a 4-item subscale of the Copenhagen Psychosocial Questionnaire, e.g., 'How often do you get help and support from your colleagues?' (Cronbach's alpha 0.80) [25].

Mastery

Mastery was assessed using the Pearlin-Schooler Mastery Scale [26]. The scale consisted of 7 items (answer range 1=fully disagree to 5=fully agree), with higher scores indicating higher mastery, which is favorable (Cronbach's alpha 0.85). A sample item from this scale is: 'I have little control over the things that happen to me'.

Depression

Symptoms of depression were measured using the 10-item Center for Epidemiologic Studies Depression Scale (Cronbach's alpha 0.87) [27]. Higher scores indicate a higher frequency of depressive symptoms, which is unfavorable. Sample items are 'During the last week I was bothered by things that normally don't bother me' and 'During the last week I felt lonely'. Answers could range from 1=rarely, if ever (less than 1 day) to 4=most of the time, or always (5-7 days).

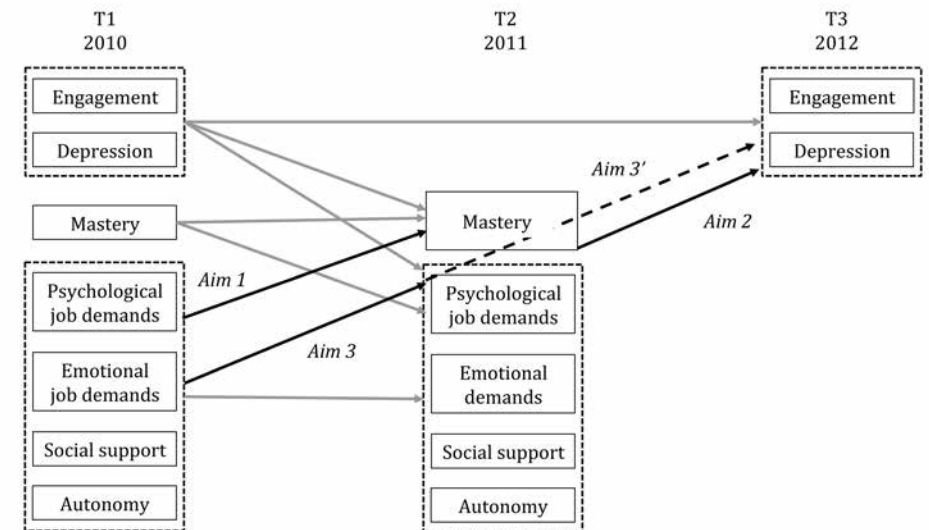
Work engagement

Work engagement was measured by the subscales Vigor (e.g. 'I feel strong and vigorous at work') and Dedication (e.g. 'I am proud of my work') of the Utrecht Work Engagement Scale-9 [14]. Absorption was not measured in the STREAM questionnaire. The scale comprises 6 items, and ranges from 0=never to 6=always or daily. Higher scores reflect higher work engagement, which is favorable (Cronbach's alpha 0.93).

Covariates

Information on age, gender and educational level was available and included in the analyses in order to correct for potential confounding.

Figure 2 — Tested relations in LISREL model, between job demands, job resources, mastery, work engagement and depression in a longitudinal observational cohort of teachers (aged 45-64 years at baseline) (N=549)



Note: Solid bold lines show the relations tested to address the research aims, grey lines show the additionally tested relations.

Statistical analyses

The data were analyzed using SPSS for Windows, and with the packages PRELIS 2 [28] and LISREL 8 [29].

Pearson's correlations (r) between all job demands and resources, mastery, depression and work engagement at baseline were calculated.

PRELIS was used for data screening and for the construction of covariance matrices that were used in LISREL. In the LISREL analyses, the model of the longitudinal relation between job demands, job resources, mastery, work engagement and depression as presented in Figure 2 was tested, using a Robust Maximum Likelihood (RML) method [30]. In RML, the standard errors of the parameter estimates are corrected for non-normal distributions by using the asymptotic covariance matrix. The PRELIS analysis showed that especially symptoms of depression were highly positively skewed and had a high kurtosis; to a lesser extent, work engagement and mastery were negatively skewed. The constructs were all analyzed as observed variables, leading to a path model rather than a structural equation model with latent variables. Figure 2 presents the longitudinal paths tested between the endogenous variables: psychological job demands (T1, T2), emotional job demands (T1, T2), social support (T1, T2), autonomy (T1, T2), mastery (T1, T2), work engagement (T1, T3), and symptoms of depression (T1, T3). Gender, age and education were included in the model as exogenous variables with free paths to all endogenous variables at the first measurement in order to explore how they related to variables of interest, and in order to control for potential confounding effects.

In the model, the contribution of work-related factors at T1 to mastery at T2 (aim 1) and of mastery at T2 to work engagement and depression at T3 (aim 2) were tested. The overall relation of work-related factors (T1) with work engagement and depression (T3) was determined, as well as the contribution of work-related factors (T1) to work engagement and depression (T3) through mastery at T2 (aim 3). In addition, as advocated by several authors (14,15), reversed causality was tested and taken into account by incorporating other longitudinal relations (marked in grey in Figure 2). All longitudinal relations were adjusted for work engagement and depression at T1. In addition, longitudinal relations of work engagement and depression at T1 to work-related factors and mastery at T2 were included in the model, as well as the longitudinal relations of mastery at T1 to the work-related factors at T2. Direct effects of work-related factors and mastery at T1 to work engagement and depression at T3 were constrained to be zero. It was assumed that work-related factors at T1 affected T3 work engagement and depression through T2 work-related factors and through T2 mastery. The indirect effect via mastery was estimated by multiplying the direct effects of T1 work-related factors on T2 mastery with the direct effects of T2 mastery on T3 depression and work engagement. These indirect effects were tested using Monte Carlo simulations following the procedure of Selig and Preacher [31].

In the tested model, the cross-sectional covariances (Ψ) within the measurements T1 and T2 between work-related factors and mastery were estimated, and unstandardized covariances were constrained to be equal in both measurements. These equality constraints resulted in a more parsimonious model than allowing these covariances to differ between measurements.

For evaluating the fit of the model, the Satorra Bentler Chi Square for non-normal distributions (χ^2) was used [30]. In addition, as recommended by Bentler [32], the Comparative Fit Index (CFI), the Root Mean Square Error of Approximation (RMSEA) and the Standardized Root Mean Square Residual (SRMR) were used to evaluate the fit. Values of 0.95 and above for CFI, and values of 0.08 and below for RMSEA and SRMR indicate a good fit of the model [33].

Results

Descriptives

The characteristics of the study population are shown in Table 1. The mean age was 54.4 years (Standard Deviation (SD) 4.6) and women were slightly in the majority (55%). The population consisted of highly educated workers mostly (82.7%), since teaching is a profession requiring higher education in the Netherlands.

Persons with missing information on any of the relevant parameters during follow-up were compared to those included in the study on job demands, job resources, mastery, depression and work engagement at baseline. All p -values were non-significant (i.e., larger than 0.10).

Correlations between work-related factors, mastery, depression, and work engagement at baseline varied from very low to moderate (Pearson r range: 0.02 to 0.55; Appendix Table 4).

Finally, the contribution (Γ) of gender, age and education to the variables in the model at baseline are presented in the Appendix, Table 2. It was found that female teachers experienced more emotional demands, more social support but less autonomy, and had higher work engagement than male teachers. Moreover, highly educated teachers experienced more psychological and emotional job demands and lower autonomy than teachers with a lower educational level. With regard to age, it was found that psychological job demands were somewhat lower in older teachers.

Table 1 — Demographic information, job demands and job resources, mastery, depression and work engagement of teachers aged 45-64 years (N=549)

Variable (possible range)[#items; Cronbach's alpha]			Wave	Teachers N=549	
				Mean (SD) or Percentage (N)	Median
Age (45-64 years)			T1	54.5 (4.64)	55.0
Gender**	Male	% (n)	T1	45.0 % (247)	
	Female	% (n)	T1	55.0% (302)	
Educational level**	Low	% (n)	T1	3.6% (20)	
	Middle	% (n)	T1	13.7% (75)	
	High	% (n)	T1	82.7% (454)	
Mastery (1-5)[7 items; 0.85]			T1	3.80 (0.67)	3.86
			T2	3.75 (0.68)	3.86
			T3	3.74 (0.70)	3.71
Job demands (1-5)	Psychological job demands [4 items; 0.85]		T1	3.18 (0.79)	3.25
			T2	3.18 (0.77)	3.25
			T3	3.21 (0.82)	3.25
	Emotional job demands [3 items; 0.80]		T1	3.01 (0.70)	3.00
			T2	2.99 (0.70)	3.00
			T3	3.02 (0.73)	3.00
Job resources (1-5)	Autonomy [5 items; 0.70]		T1	3.40 (0.62)	3.40
			T2	3.38 (0.62)	3.40
			T3	3.37 (0.63)	3.40
	Support [4 items; 0.80]		T1	3.64 (0.70)	3.75
			T2	3.60 (0.73)	3.75
			T3	3.57 (0.76)	3.50
Work engagement (0-6) [6 items; 0.93]			T1	4.63 (1.02)	4.83
			T2	4.62 (1.02)	4.83
			T3	4.57 (1.13)	4.83
Depression (1-4) [10 items; 0.87]			T1	1.52 (0.49)	1.40
			T2	1.49 (0.42)	1.40
			T3	1.54 (0.48)	1.40

**for categorical variables the distribution between the two groups is unequal (Pearson Chi-Square test), for continuous variables (independent sample t-test) this group scored statistically significantly ($p < 0.01$) higher than the other group of workers (this was only tested for these three exogenous variables).

Note: SD= standard deviation, and sample size for all variables.

Table 2 — Standardized relations of T1 job demands and job resources with T2 mastery and T3 depression and work engagement

T1	T2		T3	
	Mastery ¹		Depression ²	Work Engagement ²
Psychological job demands	0.089 **	Total effect (T1-T3)	0.056 *	-0.017
		Through T2 mastery	-0.020 **	0.009 *
Emotional job demands	-0.079 *	Total effect (T1-T3)	0.075 **	-0.063 **
		Through T2 mastery	0.018 *	-0.008 *
Autonomy	0.002	Total effect (T1-T3)	-0.014	0.011
		Through T2 mastery	-0.000	0.000
Support	0.007	Total effect (T1-T3)	0.005	0.012
		Through T2 mastery	-0.002	0.001
Mastery			-0.223 **	0.096 **

Note: † $p < 0.10$; * $p < 0.05$; ** $p < 0.01$

¹Mastery at T2 was corrected for Mastery at baseline (T1)

²Depression and work engagement at T3 were corrected for depression and work engagement at baseline (T1)

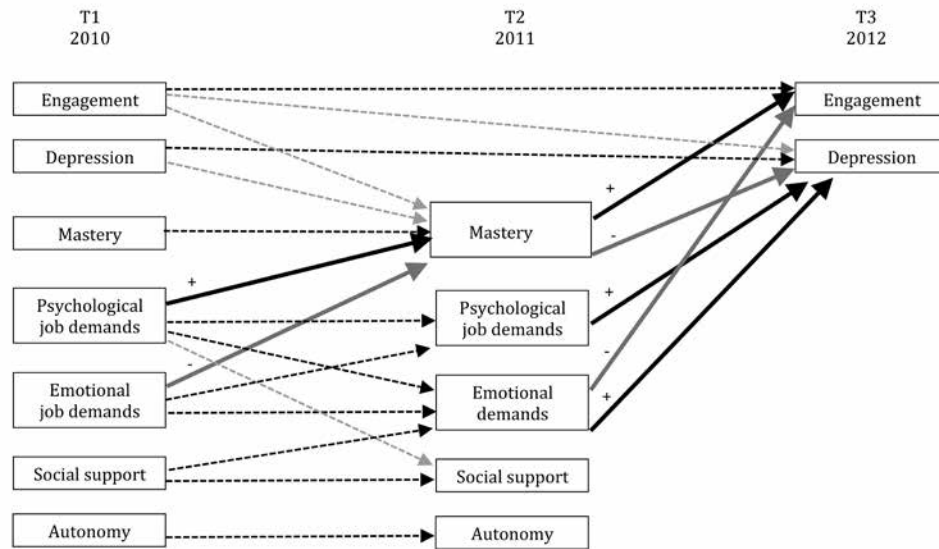
Model fit

The tested model (Figure 2) had a good fit, Satorra-Bentler χ^2 (df=40) = 53.9, $p = 0.07$, CFI = 0.99, RMSEA = 0.03, SRMR = 0.04. The longitudinal relations relevant for the aims of our study are presented in Table 2, and the pattern of significant relations is shown in Figure 3. All tested parameters in the model are presented in the Appendix.

Relation between work-related factors and mastery (aim 1)

In the left column of Table 2 the standardized effects of work-related factors at T1 on mastery at T2, controlling for mastery at T1, are presented. Psychological job demands contributed significantly to an increase in mastery one year later. Emotional job demands were associated with reduced mastery one year later. Thus *hypothesis 1a*, high job demands are related to a decrease in mastery, was rejected for psychological demands and confirmed for emotional job demands. The job resources, autonomy and social support, did not affect mastery (Table 2). *Hypothesis 1b*, high resources are related to an increase in mastery, was not supported.

Figure 3 — Relations between work-related factors, mastery, work engagement and depression



Note: Solid bold lines show the relations tested to address the research aims, broken lines show the additionally tested relations. Black lines indicate a positive association, grey lines indicate a negative association. For the relations tested to address the research aims, a plus (positive) or minus (negative) additionally indicates the direction of the association.

Relation between mastery, depression and work engagement (aim 2)

Mastery at T2 was related to lower depression at T3 and related to higher work engagement at T3, controlling for depression and work engagement at T1 (see Table 2, bottom row). *Hypothesis 2a* and *2b* were thus confirmed.

Relation of work-related factors with depression and work engagement, and the role of mastery (aim 3)

Table 2 shows how T1 work-related factors were related to T3 depression and work engagement, controlling for depression and work engagement at T1. Both the total effect (i.e., including indirect effects through work-related factors at T2) and the indirect effect via mastery at T2 are presented.

Psychological job demands influenced depression in two opposite ways. Namely, high psychological job demands at T1 was related to increased depression at T3, mostly through T2 psychological job demands. However, a smaller negative relation of T1 psychological job demands was found for depression at T3 through mastery at T2: high psychological job demands increased mastery a year later, which in turn reduced feelings of depression. This suggests that mastery partly suppressed the effect of high psychological job demands on depression.

Table 3 — Summary of the results of a path analysis including work-related factors, mastery, and indicators of sustainable employability measured at three points in time in a sample of 549 ageing teachers

Aim 1 Determine relation between work-related factors and mastery		
Hypothesis 1a	<i>Job demands</i> (T1) are related to decreased <i>mastery</i> (T2)	Partly confirmed
Hypothesis 1b	<i>Job resources</i> (T1) are related to increased <i>mastery</i> (T2)	Rejected
Aim 2 Determine relation between mastery and indicators of sustainable employability		
Hypothesis 2a	<i>Mastery</i> (T2) is related to decreased <i>depression</i> (T3)	Confirmed
Hypothesis 2b	<i>Mastery</i> (T2) is related to increased <i>work engagement</i> (T3)	Confirmed
Aim 3 Determine relation between work-related factors and indicators of sustainable employability and whether mastery mediated this relation		
Hypothesis 3a	<i>Job demands</i> (T1) are related to increased <i>depression</i> (T3)	Partly confirmed
Hypothesis 3b	<i>Job demands</i> (T1) are related to decreased <i>work engagement</i> (T3)	Partly confirmed
Hypothesis 3c	<i>Job resources</i> (T1) are related to decreased <i>depression</i> (T3)	Rejected
Hypothesis 3d	<i>Job resources</i> (T1) are related to increased <i>work engagement</i> (T3)	Rejected
Hypothesis 3e	<i>Mastery</i> (T2) mediates the longitudinal effects of <i>job demands</i> and <i>job resources</i> (T1) on <i>depression</i> and <i>work engagement</i> (T3)	Partly confirmed

The net effect of psychological job demands on change in work engagement was not significant. However, a significant positive indirect effect existed through higher mastery, namely: T1 psychological job demands was related to increased T2 mastery, which in turn showed a relation with increased T3 work engagement.

Emotional job demands were related to increased depression and lowered work engagement. These relations were partly mediated by lower mastery, that is: high emotional job demands showed a relation with decreased mastery one year later, which in turn was related to increased feelings of depression and decreased work engagement. Both indirect relations of emotional job demands on depression and work engagement through lower mastery were significant.

In sum, *hypotheses 3a* and *3b* on job demands and *hypothesis 3e* on mediation are partly supported (Table 3).

The job resources, autonomy and social support, were not related to change in work engagement or depression (*hypotheses 3c* and *3d* rejected). Neither were they related to change in mastery (*hypothesis 1b* rejected). As a result, no mediation through mastery was found for the relations between job resources and depression and work engagement (*hypothesis 3e* partly rejected; Table 3).

Discussion

The objective of the current study was to explore the interplay between mastery and job demands and job resources in influencing changes in two important precursors of sustainable employability – depression and work engagement – in a longitudinal sample of ageing teachers. Table 3 provides an overview of the aims, the hypotheses, and whether they were (partly) confirmed. Higher mastery was related to a decrease in depression and an increase in work engagement one year later (aim 2). Overall, job demands were of importance in the longitudinal paths found, whereas job resources were not. The two types of job demands were differently related to mastery and the outcomes depression and work engagement (aim 3). That is, psychological demands contributed to higher mastery, emotional job demands contributed to lower mastery (aim 1). Baseline psychological demands were related to both an increase in depression two years later (via T2 job demands) and a decrease in depression (via T2 mastery), but not to changes in work engagement. Emotional job demands were related in the expected direction to both higher depression and lower work engagement two years later and these relations were partly mediated by lower mastery.

Aim 1: Determine relation between work-related factors and mastery

The findings that both types of job demands had an opposite relation with mastery, are in line with research on challenge and hindrance demands. This will be discussed in the paragraph on aim 3, as similar relations were found between (some of the) work-related factors and depression and work engagement.

Finding no relation between job resources and mastery is not in line with previous studies. For example, Simbula et al [34] found a longitudinal relation between self-efficacy and increased job resources in teachers, as did Consiglio et al [35] cross-sectionally among call center agents. Xanthopoulou et al [18] found that self-efficacy and organizational based self-esteem increased job resources in a two-wave study among electrical engineers. This partial congruence between our finding and previous studies might be explained firstly by the type of personal factor that was studied in this article (i.e. mastery) compared to other personal factors (e.g. self-efficacy, organizational based self-esteem). Apparently, different personal factors have different relations with psychological demands and both job resources. In future studies, several personal factors could be studied simultaneously, in order to grasp the specific role of each. A second explanation could be that the current operationalization of mastery was too distant from the work situation to demonstrate a relation with job resources. Item wordings are for example: “I have little control over the things that happen to me” and “I often feel helpless in coping with life’s problems”. An employee might feel able to “control important circumstances that are currently impinging on one’s life” (as reflected in a high mastery score), but unable to control such circumstances in one’s *working* life (as reflected in a lacking rela-

tion with job resources). In future research, a measure related to mastery but closer to the work context could be adopted, such as occupational self-efficacy (e.g. [36]). A third explanation could be the place of the personal factor in the model. Some longitudinal evidence exists on personal factors as antecedents of job resources [34] and job demands [37] in teachers, and cross-sectional evidence among other occupations than teaching points in the same direction [18, 35]. Nevertheless, in the present study such reversed effects of mastery on job demands and job resources were included in the model, but none of these reversed effects proved significant (see Appendix for details). Future research should consider other roles of personal factors in the JD-R model, for example as moderators, in order to better understand how personal factors might be related to job demands, job resources and work outcomes [18].

Aim 2: Determine relation between mastery and depression and work engagement

The finding that mastery was related to depression and work engagement in the expected direction is in line with previous studies. In a two-wave study among teachers, a personal factor (i.e. self-efficacy) was related to work engagement and vice versa [38]. In a comparable study, this relation was also confirmed for burnout [37]. A relation between self-efficacy and decreased burnout was found in a cross-sectional study among more than 5000 call center agents [35]. In sum, it appears that the personal factor mastery has a direct effect on depression and work engagement, and therefore can, in its own, be considered as a potential element in interventions for work-related mental health. Theoretical studies have described that a ‘mastery experience’ could be provided by taking part in a participative intervention, consisting of shared decision making in relation to a(n occupational health) problem [39]. Empirical studies have demonstrated that mastery can be changed by cognitive behavioral strategies, replacing unconstructive perceptions with constructive ones [40]. Further research is needed to determine the elements of interventions that are effective in increasing mastery and thus improve health and sustainable employability.

Aim 3: Determine relations between work-related factors and indicators of sustainable employability and whether mastery mediated this relation

Overall job demands influenced changes in teachers’ depression and work engagement over time, whereas job resources did not. The indirect effects of the two types of job demands, psychological demands and emotional job demands, on depression and work engagement were in opposite directions. This suggests that the type of job demand matters, as has been suggested by others (e.g. [41]). In the literature, hindrance demands and challenge demands are distinguished [42]. Both types of demands deplete energy, but challenge demands might yield additional benefits in terms of personal growth or development (e.g. cognitive demands, time pressure) whereas hindrance demands do not (e.g. role ambiguity, conflicts) [41]. In the current longitudinal study, the psychological job demands seemed to act as a challenge demand, whereas emotional job demands seemed to

act as a hindrance demand. A meta-analysis has shown that challenge demands are associated with positive outcomes (e.g. job satisfaction, organizational commitment), whereas hindrance demands are associated with negative outcomes (e.g. turnover intention) [43]. However, in the current study the challenge demand (i.e. psychological job demands) seemed to be a mixed blessing: the challenge demand that reduces depression through higher mastery is offset by an increase in depression through other mechanisms (e.g. emotional demands).

In order to better understand the hindrance demand in this study, emotional job demands, the recently proposed distinction between threat and hindrance seems relevant [44]. According to Tuckey and colleagues, a job demand is a 'threat' if it is appraised as a potential personal harm or loss, and a 'hindrance' if it is appraised as an obstacle to growth and accomplishment, indirectly leading to personal harm or loss [44]. Emotional job demands in teachers (i.e. working in emotionally demanding situations, appraising the job as emotionally demanding, and becoming emotionally involved in the job) rather seemed to pose a risk of personal harm or loss (a threat), as reflected in the decrease in mastery, rather than an obstacle to growth or accomplishment (hindrance). Labeling job demands as challenge, hindrance or threat can help to determine the type of intervention that is best to apply in order to achieve optimal individual and organizational outcomes.

Contrary to expectations, job resources were not related to mastery nor to the outcomes (i.e. depression and work engagement), and therefore no mediation by mastery was found. This finding is not in line with prior research wherein a relation between job resources and personal factors was consistently found (e.g. [18, 34]). The negative effect of high (threat) job demands might overshadow the smaller positive effect of high job resources. Another explanation might be the specific study population of ageing teachers. Older teachers are generally more experienced teachers, whose feelings of mastery have accumulated over the years and that is not so easily influenced by (changes in) work-related factors.

Strengths and limitations

This study contributes to the knowledge base on the role of a personal factor in the JDR-model, which can help to strengthen the theoretical base of individual level interventions aiming to decrease work stress. We have found that relations described in cross-sectional research could not all be confirmed in a longitudinal study within a specific occupational group. This finding underlines once more the importance of longitudinal research, as acclaimed by other scholars [20]. A second strength is that we studied a specific occupational group, which provides more specific tools and insights for intervention studies in that occupational sector than research in the general working population. Another strength is the type of analysis, as we simultaneously tested different longitudinal relations, which is in line with the recommendation not to separate the JDR-model's health impairment process (job demands > burnout > health deterioration) from the health promoting process (job resources > work engagement > positive outcomes) [20]. Lastly, the addition of a personal factor to the JD-R model is

relatively new, and this study shows that it could be a relevant road to pursue. A limitation of the current study is the number (two) and the type of job resources (autonomy and social support of colleagues and supervisor). Comparable studies included three job resources, including opportunities for growth [18, 34] or availability of all necessary information [35], besides support and autonomy. Future studies could thus consider including more and different personal and job resources to expand the knowledge on specific job resources in relation to specific personal factors.

Conclusion

The current study aimed to contribute to the knowledge-base on sustainable employability, specifically within the occupational group of teachers. The JD-R model was tested in order to determine the longitudinal role of mastery in a sample of older teachers. Mastery had a direct influence on changes in depression and work engagement. Mastery partly countered the negative effect of high psychological job demands on depression, and it revealed a small positive effect on work engagement, whereas it partly explained the longitudinal relations between emotional job demands and both outcomes. Increasing mastery in ageing teachers may be considered a promising intervention strategy to decrease depression and increase work engagement in this occupational group.

References

- OECD: **How Old Are the Teachers?** *Education Indicators in Focus* 2014.
- Hooftman WE, Mars GMJ, Janssen BJM, De Vroome EMM, Bossche van den SNJ: **Nationale Enquête Arbeidsomstandigheden 2014: Methodologie en globale resultaten. [Netherlands Working Conditions Survey 2014: Methodology and overall results]**. 2015.
- Seibt R, Lützkendorf L, Thinschmidt M: **Risk factors and resources of work ability in teachers and office workers.** *International Congress Series* 2005, **1280**:310-315.
- Pas ET, Bradshaw CP, Hershfeldt PA: **Teacher- and school-level predictors of teacher efficacy and burnout: Identifying potential areas for support.** *J School Psychol* 2012, **50**(1):129-145.
- van der Klink JJJ, Bültmann U, Burdorf A, Schaufeli WB, Zijlstra FRH, Abma FI, Brouwer S, van der Wilt GJ: **Sustainable employability-definition, conceptualization, and implications: A perspective based on the capability approach.** *Scand J Work Environ Health* 2016, **42**(1):71.
- Klassen R, Wilson E, Siu AFY, Hannok W, Wong MW, Wongsri N, Sonthisap P, Pibulchol C, Buranachaitavee Y, Janssen A: **Preservice teachers' work stress, self-efficacy, and occupational commitment in four countries.** *European Journal of Psychology of Education* 2013, **28**(4):1289-1309.
- Smith A, Brice C, Collins A, McNamara R, Matthews V: **Scale of occupational stress: a further analysis of the impact of demographic factors and type of job.** 2000, **Contract Research Report 311/2000**.
- Brown J, Gilmour WH, Macdonald EB: **Ill health retirement in Scottish teachers: process, outcomes and re-employment.** *Int Arch Occup Environ Health* 2006, **79**(5):433-440.
- Weber A, Weltle D, Lederer P: **Ill health and early retirement among school principals in Bavaria.** *Int Arch Occup Environ Health* 2005, **78**(4):325-331.
- Demerouti E, Nachreiner F, Bakker AB, Schaufeli WB: **The job demands-resources model of burnout.** *J Appl Psychol* 2001, **86**(3):499-512.
- Netterstrom B, Conrad N, Bech P, Fink P, Olsen O, Rugulies R, Stansfeld S: **The relation between work-related psychosocial factors and the development of depression.** *Epidemiol Rev* 2008, **30**:118-132.
- Theorell T, Hammarstrom A, Aronsson G, Traskman Bendz L, Grape T, Hogstedt C, Marteinsdottir I, Skoog I, Hall C: **A systematic review including meta-analysis of work environment and depressive symptoms.** *BMC Public Health* 2015, **15**:738-015-1954-4.
- Halbesleben JR: **A meta-analysis of work engagement: Relationships with burnout, demands, resources, and consequences.** *Work engagement: A handbook of essential theory and research* 2010, **8**:102-117.
- Schaufeli WB, Bakker AB, Salanova M: **The measurement of work engagement with a short questionnaire: A cross-national study.** *Educational and Psychological Measurement* 2006, **66**(4):701-716.
- Leijten FR, van den Heuvel, Swenne G, van der Beek, Allard J, Ybema JF, Robroek SJ, Burdorf A: **Associations of work-related factors and work engagement with mental and physical health: a 1-year follow-up study among older workers.** *J Occup Rehabil* 2015, **25**(1):86-95
- Bakker AB, Demerouti E: **The job demands-resources model: State of the art.** *J Manage Psychol* 2007, **22**(3):309-328.
- Alarcon G, Eschleman KJ, Bowling NA: **Relationships between personality variables and burnout: A meta-analysis.** *Work Stress* 2009, **23**(3):244-263.
- Xanthopoulou D, Bakker AB, Demerouti E, Schaufeli WB: **The role of personal resources in the job demands-resources model.** *International Journal of Stress Management* 2007, **14**(2):121-141.
- Llorens S, Schaufeli W, Bakker A, Salanova M: **Does a positive gain spiral of resources, efficacy beliefs and engagement exist?** *Comput Hum Behav* 2007, **23**(1):825-841.
- Schaufeli WB, Taris TW: **A Critical Review of the Job Demands-Resources Model: Implications for Improving Work and Health.** In *Bridging Occupational, Organizational and Public Health: A Transdisciplinary Approach*. Edited by Bauer GF, Hämmig O. Dordrecht: Springer Science+Business Media; 2014:43.
- Thoits PA: **Stress and health: major findings and policy implications.** *J Health Soc Behav* 2010, **51** Suppl:S41-53.
- Pearlin LI: **The life course and the stress process: some conceptual comparisons.** *J Gerontol B Psychol Sci Soc Sci* 2010, **65B**(2):207-215.
- Ybema JF, Geuskens GA, van den Heuvel S, de Wind A, Leijten FR, Joling C, Blatter BM, Burdorf A, van der Beek, Allard J, Bongers PM: **Study on Transitions in Employment, Ability and Motivation (STREAM): The Design of a Four-year Longitudinal Cohort Study among 15,118 Persons Aged 45 to 64 Years.** *British Journal of Medicine & Medical Research* 2014, **4**(6):1383-1399.
- Karasek R, Brisson C, Kawakami N, Houtman I, Bongers P, Amick B: **The Job Content Questionnaire (JCQ): an instrument for internationally comparative assessments of psychosocial job characteristics.** *J Occup Health Psychol* 1998, **3**(4):322-355.
- Kristensen TS, Hannerz H, Høgh A, Borg V: **The Copenhagen Psychosocial Questionnaire-a tool for the assessment and improvement of the psychosocial work environment.** *Scand J Work Environ Health* 2005, **31**(6):438-449.
- Pearlin LI, Lieberman MA, Menaghan EG, Mullan JT: **The stress process.** *J Health Soc Behav* 1981, **22**(4):337-356.
- Radloff LS: **The CES-D scale a self-report depression scale for research in the general population.** *Applied psychological measurement* 1977, **1**(3):385-401.
- Jöreskog KG, Sörbom D: **PRELIS 2 User's Reference Guide: A Program for Multivariate Data Screening and Data Summarization: a Preprocessor for LISREL: Scientific Software International; 1996.**
- Jöreskog KG, Sörbom D: **LISREL 8: User's reference guide: Scientific Software International; 1996.**
- Boomsma A, Hoogland JJ: **The robustness of LISREL modeling revisited.** In *Structural equation models: Present and future. A festschrift in honor of Karl Jöreskog*. Edited by Cudeck R, Toit Sd, Sörbom D. Lincolnwood, IL: Scientific Software International; 2001:139-168.
- Selig JP, Preacher KJ: **Monte Carlo method for assessing mediation: An interactive tool for creating confidence intervals for indirect effects.** *Monte Carlo method for assessing mediation: An interactive tool for creating confidence intervals for indirect effects [Computer software]* 2008.
- Bentler P: **On tests and indices for evaluating structural models.** *Personality and Individual Differences* 2007, **42**(5):825-829.
- Hu L, Bentler PM: **Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification.** *Psychol Methods* 1998, **3**(4):424.
- Simbula S, Guglielmi D, Schaufeli WB: **A three-wave study of job resources, self-efficacy, and work engagement among Italian schoolteachers.** *European Journal of Work and Organizational Psychology* 2011, **20**(3):285-304.
- Consiglio C, Borgogni L, Alessandri G, Schaufeli WB: **Does self-efficacy matter for burnout and sickness absenteeism? The mediating role of demands and resources at the individual and team levels.** *Work Stress* 2013, **27**(1):22-42.
- Schyns B, Von Gernot C: **A new occupational self-efficacy and its relation to personality constructs and organizational variables.** *European Journal of Work and Organizational Psychology* 2002, **11**(2):219-241.
- Vera M, Salanova M, Lorente L: **The predicting role of self-efficacy in the Job Demands-Resources Model: A longitudinal study.** *Estudios de Psicología* 2012, **33**(2):167-178.
- Lorente Prieto L, Salanova Soria M, Martínez Martínez I, Schaufeli W: **Extension of the Job Demands-Resources model in the prediction of burnout and engagement among teachers over time.** 2008, **20**(3):354-360.
- Bandura A: **Cultivate self-efficacy for personal and organizational effectiveness.** In *Handbook of principles of organizational behavior: indispensable knowledge for evidence-based management*. Edited by Locke EA. Chichester, Sussex; [Hoboken]: John Wiley & Sons; 2009:120-136.
- Bandura A: **Human agency in social cognitive theory.** *Am Psychol* 1989, **44**(9):1175.
- Van den Broeck A, De Cuyper N, De Witte H, Vansteenkiste M: **Not all job demands**

are equal: Differentiating job hindrances and job challenges in the job demands-resources model. *European journal of work and organizational psychology* 2010, 19(6):735-759.

42. Cavanaugh MA, Boswell WR, Roehling MV, Boudreau JW: **An empirical examination of self-reported work stress among US managers.** *J Appl Psychol* 2000, 85(1):65.
43. Podsakoff NP, LePine JA, LePine MA: **Differential challenge stressor-hindrance stressor relationships with job attitudes, turnover intentions, turnover, and withdrawal behavior: a meta-analysis.** *J Appl Psychol* 2007, 92(2):438.
44. Tuckey MR, Searle BJ, Boyd CM, Winefield AH, Winefield HR: **Hindrances are not threats: Advancing the multidimensionality of work stress.** *J Occup Health Psychol* 2015, 20(2):131.

Appendix

Table 1 — Regression weights (Beta), standardized solution based on LISREL results

	T1 PJD	T1 EJD	T1 SS	T1 AU	T1 MA	T1 WE	T1 DE
T2 Psychological job demands	0.68 ***	0.12 ***	0.04	-0.02	-0.03	-0.00	-0.07
T2 Emotional job demands	0.10 **	0.62 ***	0.10	**	0.01	0.01	0.04
T2 Social support	-0.09 *	0.03	0.60 ***	-0.00	-0.05	0.05	-0.04
T2 Autonomy	-0.07	-0.05	-0.01	0.67 ***	0.05	0.01	0.02
T2 Mastery	0.09 **	-0.08 *	0.01	0.00	0.66 ***	-0.08 *	-0.13 ***
	T1 WE	T1 DE	T2 PJD	T2 EJD	T2 SS	T2 AU	T2 MA
T3 Work engagement	0.65 ***	-0.02	-0.02	-0.09	**	0.01	0.10 **
T3 Depression	-0.12 *	0.34 ***	0.10	0.07	-0.01	-0.01	-0.22 ***

Note: * p < .05; ** p < .01; *** p < .001; PJD = T2 Psychological job demands; EJD = Emotional job demands; SS = Social support; AU = Autonomy; MA = Mastery; WE = Work engagement; DE = Depression

Table 2 — Effects of demographic variables (Gamma), standardized solution based on LISREL results

	Gender [§]	Education	Age
T1 Psychological job demands	0.00	0.27 **	-0.10 *
T1 Emotional job demands	0.13 *	0.36 ***	-0.06
T1 Social support	0.18 ***	-0.12	-0.08
T1 Autonomy	-0.11 *	-0.21 *	-0.02
T1 Mastery	0.03	0.02	-0.05
T1 Work engagement	0.21 ***	-0.12	-0.02
T1 Depression	0.02	0.01	-0.04

Note: * p < .05; ** p < .01; *** p < .001; § Gender was coded as 1 = male, 2 = female.

Table 3 — Residual covariances (Psi), standardized solution based on LISREL results

	T1 PJD	T1 EJD	T1 SS	T1 AU	T1 MA	T1 WE	T1 DE
T1 Psychological job demands	0.92 ***						
T1 Emotional job demands	0.39 ***	0.85 ***					
T1 Social support	-0.05 *	-0.01	0.94 ***				
T1 Autonomy	-0.05 *	-0.07 **	0.06 *	0.95 ***			
T1 Mastery	-0.05 *	-0.08 ***	0.08 ***	0.04	1.00 ***		
T1 Work engagement	0.01	-0.03	0.29 ***	0.04	0.31 ***	0.94 ***	
T1 Depression	0.14 **	0.18 ***	-0.14 **	0.01	-0.53 ***	-0.44 ***	1.00 ***
	T2 PJD	T2 EJD	T2 SS	T2 AU	T2 MA	T3 WE	T3 DE
T2 Psychological job demands	0.45 ***						
T2 Emotional job demands	0.14 ***	0.51 ***					
T2 Social support	-0.05 *	-0.01	0.61 ***				
T2 Autonomy	-0.05 *	-0.07 **	0.05 *	0.52 ***			
T2 Mastery	-0.05 *	-0.08 ***	0.07 ***	0.03	0.48 ***		
T3 Work engagement						0.51 ***	
T3 Depression						-0.21 ***	0.66 ***

Note: * p < .05, ** p < .01, *** p < .001; PJD = T2 Psychological job demands; EJD = Emotional job demands; SS = Social support; AU = Autonomy; MA = Mastery; WE = Work engagement; DE = Depression

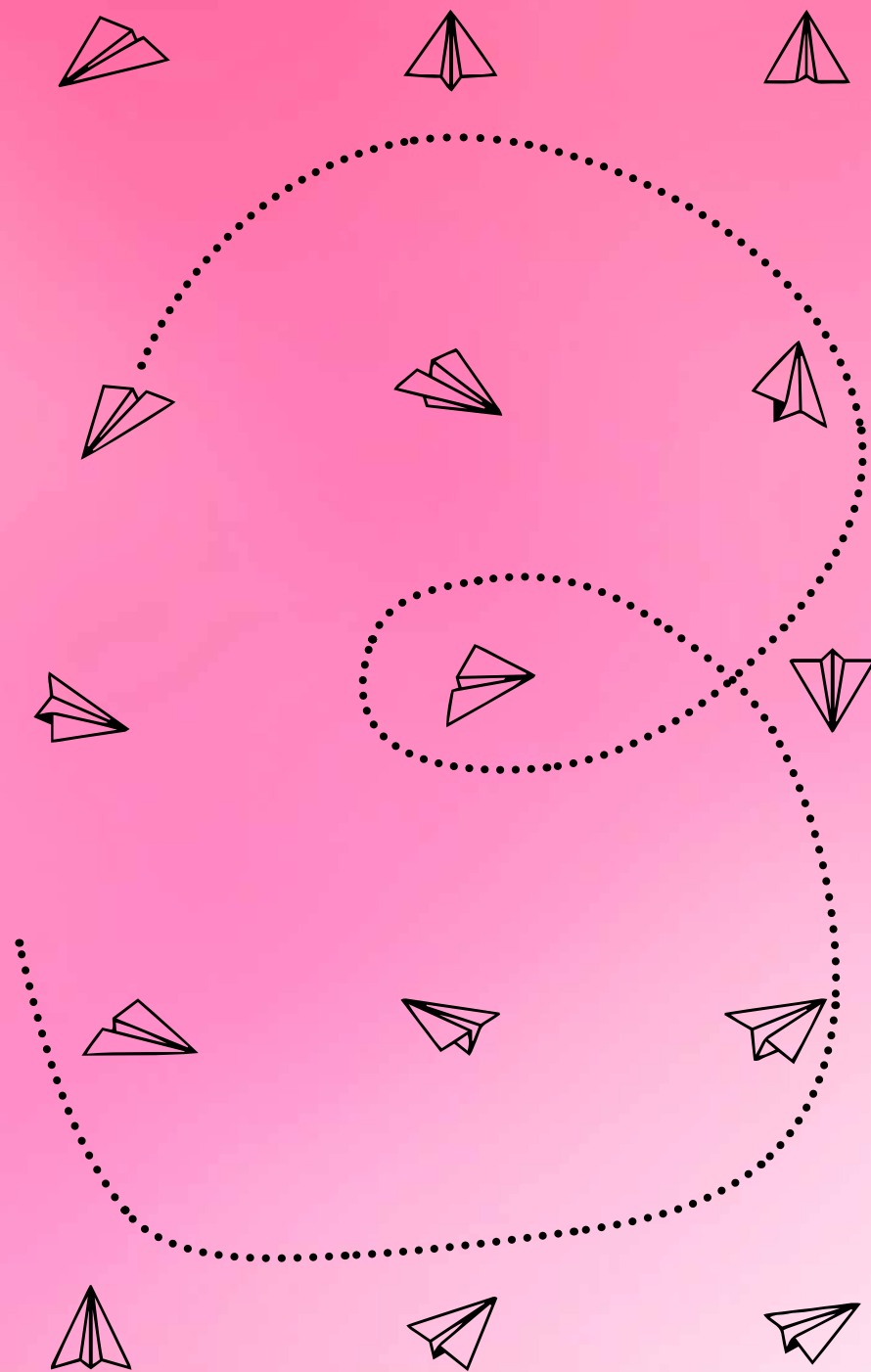
Table 4 — Pearson correlations between baseline job demands, job resources, mastery, depression, and work engagement

	1.	2.	3.	4.	5.	6.	7.
1. Psychological job demands	1	0.50**	-0.20**	-0.14**	-0.11*	0.18**	-0.05
2. Emotional job demands		1	-0.22**	-0.02	-0.19**	0.26**	-0.08
3. Autonomy			1	0.11**	0.12**	-0.06	0.08
4. Social support				1	0.16**	-0.17**	0.36**
5. Mastery					1	-0.55**	0.34**
6. Depression						1	-0.45**
7. Work engagement							1

*p<0.05, **p<0.01.

An organizational perspective





Design of the Bottom-up
Innovation project – a participatory,
primary preventive, organizational
level intervention on work-related
stress and well-being for workers in
Dutch vocational education

Roosmarijn M.C. Schelvis
Karen M. Oude Hengel
Noortje M. Wiezer
Birgitte M. Blatter
Joost A.G.M. van Genabeek
Ernst T. Bohlmeijer
Allard J. van der Beek

Abstract

Background

In the educational sector job demands have intensified, while job resources remained the same. A prolonged disbalance between demands and resources contributes to lowered vitality and heightened need for recovery, eventually resulting in burnout, sickness absence and retention problems. Until now stress management interventions in education focused mostly on strengthening the individual capacity to cope with stress, instead of altering the sources of stress at work at the organizational level. These interventions have been only partly effective in influencing burnout and well-being. Therefore, the “Bottom-up Innovation” project tests a two-phased participatory, primary preventive organizational level intervention (i.e. a participatory action approach) that targets and engages all workers in the primary process of schools. It is hypothesized that participating in the project results in increased occupational self-efficacy and organizational efficacy. The central research question: is an organization focused stress management intervention based on participatory action effective in reducing the need for recovery and enhancing vitality in school employees in comparison to business as usual?

Methods/Design

The study is designed as a controlled trial with mixed methods and three measurement moments: baseline (quantitative measures), six months and 18 months (quantitative and qualitative measures). At first follow-up short term effects of taking part in the needs assessment (phase 1) will be determined. At second follow-up the long term effects of taking part in the needs assessment will be determined as well as the effects of implemented tailored workplace solutions (phase 2). A process evaluation based on quantitative and qualitative data will shed light on whether, how and why the intervention (does not) work(s).

Discussion

“Bottom-up Innovation” is a combined effort of the educational sector, intervention providers and researchers. Results will provide insight into (1) the relation between participating in the intervention and occupational and organizational self-efficacy, (2) how an improved balance between job demands and job resources might affect need for recovery and vitality, in the short and long term, from an organizational perspective, and (3) success and fail factors for implementation of an organizational intervention.

Background

The Dutch government aspires a top five position in the global rankings for education and science [1], to ensure the competitive power of the Dutch economy. Improving the educational quality is crucial to achieve this ambition. Undisputedly, teachers and their managers play an important role in maintaining and improving the quality of education [2]. However, almost one in five workers in the Dutch educational sector (18%) suffers from work-related stress complaints, compared to one in eight workers in the Dutch working population (13%) [3]. Work-related stress is an important cause for mental health problems, such as burnout. Burnout is associated with reduced work performance (e.g [4, 5]) and its high prevalence in the educational sector thus interferes with the Dutch government’s ambition.

Work-related stress as a major problem

Burnout, as an ultimate outcome of work-related stress, is considered a prolonged response to chronic emotional and interpersonal stressors in the work context, characterized by emotional exhaustion, depersonalization and reduced personal accomplishment [6]. The work context comprises two specific sets of characteristics that influence burnout and well-being: job demands and job resources. Job demands are generally considered the physical, social or organizational aspects of the job that require sustained physical or psychological effort [7]. Job resources are the physical, social or organizational aspects of the job that may reduce job demands, help to achieve goals and stimulate learning and development [7]. A job demand, such as dealing with students with special needs, will turn into a stressor over time if job resources, such as coworker support, are insufficient or lacking [8, 9]. In the educational sector job demands have intensified at rapid pace [10], while job resources remained the same. For example, the student-teacher ratio has increased [11]; students with special needs have been integrated in the regular classes [12]; the number of accountability measures has grown, leading to numerous administrative tasks and consequent paperwork [13]; and several school reforms have been implemented in the educational sector, often even overlapping [14]. It seems likely that this intensification of job demands has contributed to the current burnout rates.

Consequences of work-related stress

Work-related stress may show as decreased vitality and increased need for recovery. These precursors of burnout have been associated with several other negative organizational outcomes, for example sickness absence and retention problems. First, sickness absence rates are relatively high in the educational sector [3]. More often than in other sectors, workers in education consider their absence a result of emotionally demanding and stressful work [3]. If a teacher falls ill, the work is often temporarily accounted for by his or her colleagues, thereby increasing the workload (i.e. a job demand) for this colleague while job

resources remain the same. This practice, although not in line with sickness replacement regulations in Dutch schools, disturbs the equilibrium between job demands and job resources of healthy colleagues. Second, a large number of teachers retire before reaching the official retirement age [15]. Between 45% [16] and 70% [17] of early retirements in teachers is accounted for by psychosomatic illness and psychological problems. Furthermore, approximately half of all novice teachers leave the sector within their first five years, as noted in a North American study [18]. Retention of both novice and experienced teachers is thus a challenge with societal implications. Burnout rates, sickness absence and lower retention rates sum up to a reduced employability of the work force, which is costly. In The Netherlands alone, work days lost due to presenteeism and sickness absence associated with mental health problems summed up to 2.7 billion Euros in 2008 [19, 20]. There is thus an urgent need for stress management interventions in the workplace. Ideally these interventions alter precursors of burnout, such as need for recovery and reduced vitality.

Interventions in education: individual-focused and secondary preventive

Stress management interventions can be classified as primary, secondary or tertiary prevention. Primary preventive interventions aim to alter the sources of stress at work (e.g. [21]). Secondary preventive interventions aim to reduce stress symptoms before they lead to health problems (e.g. [21]). Tertiary preventive interventions aim to treat health problems (e.g. [22]). Giga, Cooper and Faragher [23] found that most common stress management interventions are ‘secondary preventive’, aimed at the individual level and comprised stress management and coping techniques. The same holds true for stress management interventions in the educational sector. Until now stress management interventions in education have been ‘secondary preventive’ mostly and targeted at the individual level [24–28]. These interventions [24–28] all aimed to enhance the individual capacity of (trainee) teachers or teaching assistants to cope with stressors in the workplace, for example via mindfulness-based stress reduction or workshops on stress management skills. However, these interventions were only partly effective in influencing (dimensions of) burnout [24–28] and well-being [28]. More specifically, none of the studies influenced all three burnout dimensions positively, some influenced two dimensions (but always in differing combinations) and the long term effects were not measured. Apparently it is insufficient to reduce burnout and increase well-being in education, by focusing solely on strengthening the individual teachers’ capacity to cope with or manage stress.

The need for primary preventive organizational interventions and appropriate evaluation studies

The above leads us to the proposition that to decrease (precursors of) burnout, problems should be altered at the source, that is the (interpersonal) work context [8, 29], and targeted at the organizational level. This proposition is amplified firstly by the enormous body of research that points to the impor-

tance of the (interpersonal) work context in the development of a disbalance between demands and resources (e.g. [30, 31]).

Secondly, McVicar, Munn-Giddings, & Seebohm [32] found that primary preventive interventions can take the complexity of an organization into account when designing a preventive strategy. These interventions are therefore potentially more effective than individual level interventions [32].

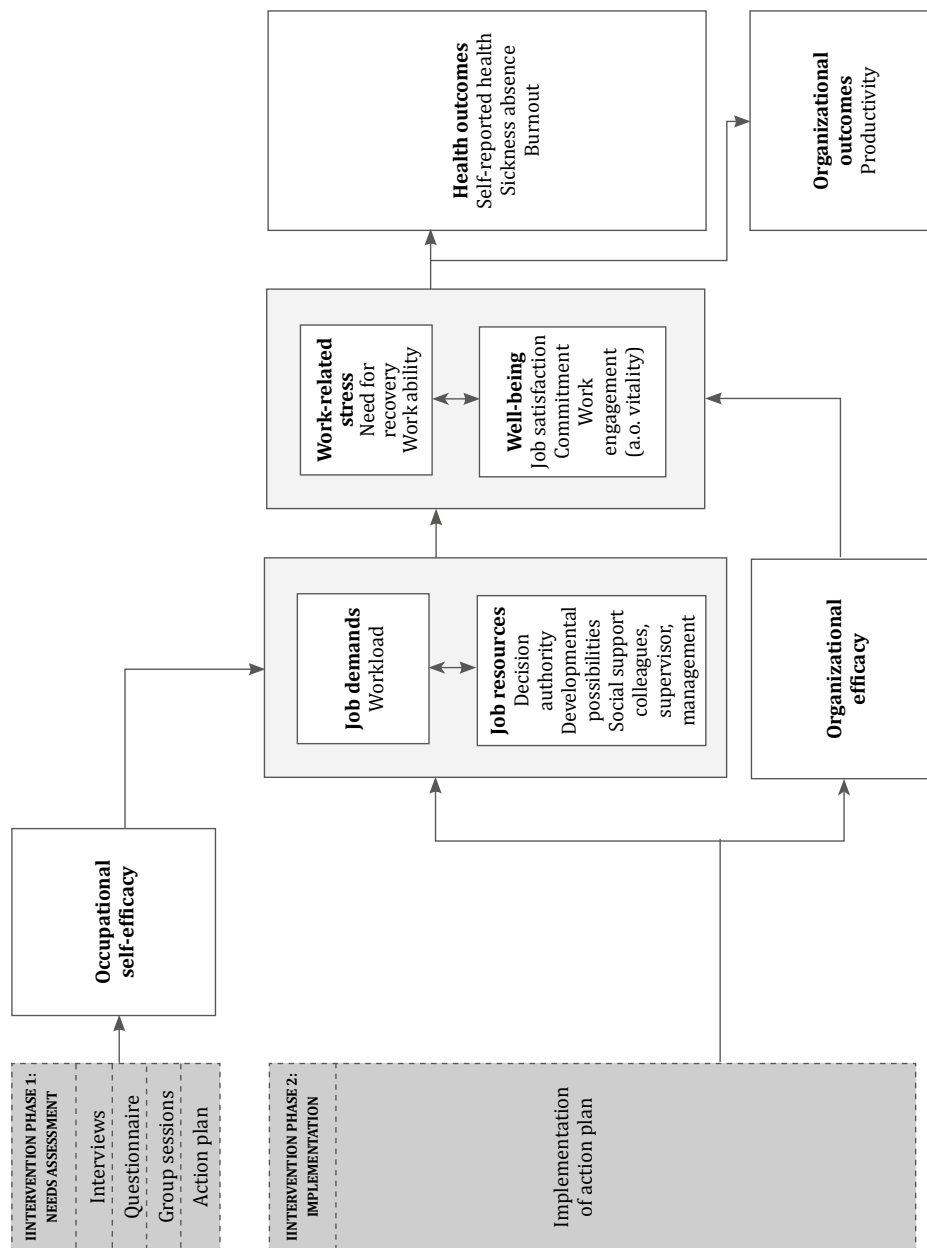
A review has suggested that if an intervention is effective, the organizational level ones are more likely to bring about positive changes than the individual level ones [33]. On the other hand, two meta-analyses on stress management interventions have failed to show substantial effects of organizational level interventions over individual level interventions [34, 35], but this has partly been explained by the underrepresentation of organizational outcome evaluation [35]. For another part, it might be explained by the finding that ‘organizational-level occupational interventions are often complex programs involving many people and several intervention components, which might [...] complicate the implementation process and the measurement of effects’ ([36], p.85). These interventions thus impose specific demands on the design of the evaluation study (e.g. monitoring the implementation process), demands that cannot be fulfilled by the gold standard design for experiments: the randomized controlled trial [37]. An organization is no laboratory where all conditions can be controlled. However, Griffiths [38] points out that occupational health interventions are still mostly regarded as experiments, set up to discover whether changes occur after manipulating a variable or introducing a particular treatment. Experiments focus on *what* works, thereby discarding to describe the processes which brought about these outcomes (*how* and *why* does it work?) [37]. Nielsen and colleagues [37] posed, that there is a lack of interventions that combine process measures (e.g. managerial support for the intervention) and effect measures (e.g. job demands). To further understand the ‘black box’ and increase the external validity (or generalizability) of interventions, the intervention ought to be evaluated by means of mixed methods [37].

The above underlines the need for appropriate evaluation of primary preventive organizational interventions. This implies for the evaluation study in the current project that: 1) the evaluation design is as rigorous as possible, 2) the implementation process is monitored by assessing process variables, 3) in the analyses it will be assessed how process variables influence intervention outcomes, and 4) (objective) organizational outcomes are measured.

Effective ingredients of primary preventive organizational interventions

The above outlines the need for primary preventive organizational interventions and a mixed methods evaluation, comprising both process and effect measures. But, what components should the intervention, or its application, comprise in order to be effective? In other words, what are effective ingredients for primary pre-

Figure 1 — Conceptual model



ventive organizational interventions in the educational sector? To the best of our knowledge, the current study is the first of that type in that sector. Therefore, we could only argue theoretically what would be the effective ingredients that bring about the desired effect. We propose hereafter that joint ownership (i.e. participation) and occupational self-efficacy play an important role in bringing about the effect on job demands and resources, and need for recovery and vitality (Figure 1).

First, the intervention should be designed in a manner that resembles the tradition of participatory action research (PAR)[39]. Essential in using PAR to design effective stress interventions in the workplace is active participation of stakeholders and (long term) collaboration between researchers and stakeholders [32]. By establishing a participatory group and making use of management's and worker's knowledge, skills and perceptions, a feeling of joint ownership of both problems and solutions is created and the participants learn-by-doing how to discuss issues in the workplace. Therefore the intervention should be considered 'bottom-up'. Nielsen and colleagues have pointed to the relative importance, but rare discussion of joint ownership [40]. This intervention will contribute to that discussion.

Second, the intervention should target occupational and organizational efficacy. Self-efficacy is 'the belief in one's own ability to master specific domains in order to produce given attainments' [41-43]. Occupational self-efficacy refers to beliefs in one's own ability in the specific domain of work. Self-efficacy can be enhanced in several manners, but the most effective way is through mastery experiences [44]. By taking part in the intervention, it is assumed that workers experience mastery and self-efficacy is thus influenced. A recent study showed that job demands and job resources partially mediated the relation between occupational self-efficacy (or: work self-efficacy) and burnout [45]. The intervention should elaborate empirically on the results of the Consiglio and colleagues article [45]. There is an intervention that comprises these supposedly effective ingredients. The intervention has been developed by a Dutch consulting firm and applied over a hundred times to public and private organizations in The Netherlands in the past decade. That intervention will be tested in the current study.

In sum: we propose an organizational level, primary preventive stress management intervention, aimed to alter the sources of work-related stress by changing the design, management and organization of work [46, 47] and to be evaluated by an effect evaluation including organizational outcomes and a process evaluation including process variables related to intervention outcomes. Both the bottom-up intervention, as well as the mixed methods design make this study innovative and a contribution to existing knowledge to the field of organizational interventions.

Study objectives

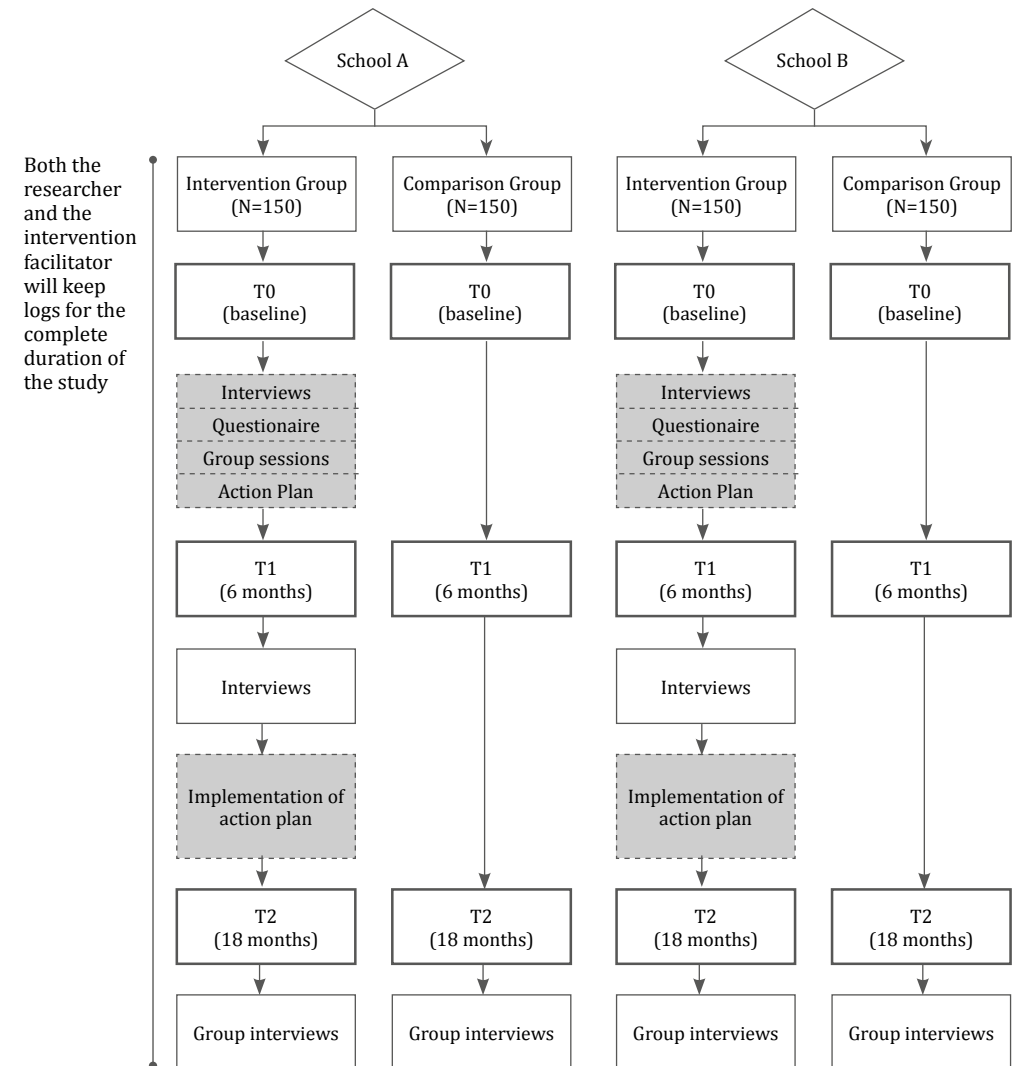
The current study tests a participatory, primary preventive organizational level intervention (i.e. a participatory action approach) that targets and engages all workers in the primary process of schools. Participation of employees and

managers is supposed to result in increased occupational self-efficacy and organizational efficacy. The application of the intervention will yield work-oriented solutions tailored to the school setting, changing (the balance between) specific job demands and job resources. By improving the balance between job demands and job resources, it is expected to improve precursors of burnout (i.e. high need for recovery, low vitality) in the long run. The central research question is thus: is an organization focused stress management intervention based on participatory action effective in reducing the need for recovery and enhancing vitality in school employees in comparison to business as usual? In this article we present the design of a controlled trial in two vocational schools in the Netherlands, wherein the participatory action approach and resulting work-oriented solutions are tested empirically.

Design

A quasi-experimental field study is conducted to determine the effectiveness of the participatory action approach (phase 1: needs assessment) and tailored work-oriented solutions (phase 2: implementation plan), compared to business as usual. The study is designed as a controlled trial (CT) with mixed methods (quantitative and qualitative) and three measurement moments: T0 at baseline (quantitative measurement), T1 at six months (quantitative and qualitative measurement) and T2 at 18 months (quantitative and qualitative measurement) (Figure 2). A CT is necessary to control for random changes, although the researchers are well aware of the fact that they are conducting a *social* experiment and that causal relations are thus embedded in complex contexts [38]. Randomization to experimental group (intervention group) or control group is practically impossible in this project, as often in organizational level workplace interventions [38], due to the aspirations of participating schools. Both schools participate in the study because they aim to solve a problem or reach a goal within a specific department of the school. The experimental groups were thus selected by the schools. To reduce the negative impact of selection bias, the control groups are selected by the researchers according to the 'general control' matching principle (or: frequency distribution control) [48]. Matching criteria are: department size (at least 150 employees), mean age of employees, and type of work (i.e. teaching vocational students and not secondary school pupils). Since the assignment to groups was out of our control, we will examine in the analyses whether propensity score matching is necessary. By applying the statistical technique of propensity score matching, the effect of the intervention can be estimated accounted for covariates that predict receiving the intervention. This way we expect to nullify potential confounding bias [49].

Figure 2 — Flow chart of design, measurements, population and intervention program



Setting

The project is conducted in two institutions for vocational education (in Dutch: *Middelbaar Beroepsonderwijs* (MBO)) in the western (Alkmaar, Hoorn and Egmond) and northern (Leeuwarden and Heerenveen) Netherlands.

Study population

The intervention is applied to one department in both schools, another department in the same school is matched by the researchers as a control group. The target group of the project are teaching and non-teaching (i.e. educational and administrative support staff) employees in two vocational education institutions and their managers. Employees who work within the vocational institution, but do not teach at a secondary vocational level are excluded from the study population (e.g. teachers in general secondary education for adults). All participants are asked to sign an informed consent at baseline.

Sample size

The sample size calculation is based on the number of cases required to detect a small (Cohen's $d = 0.2$) effect on the primary outcome vitality, as measured with the 3-item subscale of the 9-item version of the Utrecht Work Engagement Scale (UWES-9)[50]. The baseline mean vitality score (range 0-6) is assumed to be 4.01 (SD = 1.14), based on the scores of 9,679 Dutch and Belgian employees [51]. A 5% increase of the mean score on vitality in the intervention group after 12 months is considered relevant and feasible (4.21; SD 1.20).

The required sample size is then 385 (193 for both intervention and control group; thus 97 per intervention and control group per school), assuming a significance level (α) of 0.05, two-sided tests and power ($1-\beta$) of 0.80 [52]. A non-response and loss to follow-up of 35% is taken into account, so that a total sample size of 600 is needed.

The intervention

The intervention that will be tested in this study, named Heuristic Method (HM), is a participatory action approach for diagnosis, development and implementation of workplace interventions [53]. HM has been developed and applied by a Dutch consulting firm in at least 100 public and private organizations in the last decade. The consulting firm refined the intervention after each application, based on the lessons learned. Although the customers were almost always satisfied with the intervention's results, the intervention effects were never tested scientifically. The Heuristic Method is aimed at optimizing occupational self-efficacy and organizational efficacy. The purpose is to (1) use

both management's and worker's knowledge, skills and perceptions to thoroughly determine what hinders and stimulates 'healthy and happy working' in the organization, so that (2) management and workers can develop their specific work-related action plan and implementation plan that ultimately will reduce need for recovery and increase vitality. The first part of this purpose is addressed in the first phase of the intervention, named 'needs assessment'. The second part of this purpose is addressed in the second phase of the intervention, the 'implementation plan'. The needs assessment phase comprises three iterative steps led by an HM-facilitator: (1) interviews; (2) digital questionnaire; (3) group sessions, resulting in a plan of action. The components of the implementation phase can differ according to the maturity of the organization in applying organizational change processes. The minimum variant is remote counseling of the management team by a facilitator, in implementing the plan of action. Details on the application of both phases and their consequent steps in this study are provided below.

Phase 1: Needs assessment

The needs assessment is conducted in the tradition of participatory action research (PAR) [39]. Therefore, a participatory group of employees is constituted, comprising employees, a representative from the Workers Council, a staff member, a management member, the HM-facilitator and the researcher (six to eleven members in total). Its members are selected by the management team (with the exception of the facilitator and the researcher), based on their perception of the member's capacity for 'pioneering' in organizational change processes. The HM-facilitator is an expert in organizational change processes. If the intervention group is scattered among several school locations, this will be taken into account when composing the group. The participatory group is named 'Engine of Development' and becomes the project's ambassador throughout the needs assessment. The Engine of Development decides in collaboration how often they meet, but at least six times – before, during and after the three needs assessment steps.

The intervention kicks off with an information session, held after baseline measurement, led by the HM-facilitator and facilitated by the management. In the information session the HM-facilitator outlines the steps of the upcoming intervention and the researcher presents several outcomes of the baseline measurement.

Step 1: In-depth interviews

The Engine of Development approaches some prominent colleagues for an in-depth, open interview with the HM-facilitator. Prominent colleagues can be the typical optimists, pessimists, innovators, integrators or otherwise interesting employees that help the HM-facilitator grasp both initial hindrances to happy and healthy working as well as implicit norms in the intervention group. Approximately ten interviews will be held, or until saturation is reached. The HM-facilitator writes a report on his findings that is sent to all employees, after consulting the Engine of Development and the management team, respectively.

Step 2: Digital, open ended questionnaire

Based on the report of the in-depth interviews, the Engine of Development compose the questions for a digital, open ended questionnaire. The open ended questionnaire should further specify the hindrances that were found in the in-depth interviews. All employees in the intervention group are invited to take part in the questionnaire. The HM-facilitator writes a report on the findings that is sent to all employees, after consulting the Engine of Development and the management team, respectively.

Step 3: Group sessions

Based on the report of the open ended questionnaire, the Engine of Development determines several clusters of hindrances that need to be specified further in group sessions. The aim of the group sessions is not only to specify hindrances, but also to propose work-oriented solutions. All employees in the intervention group are invited to take part in the group sessions, which are chaired by the Engine of Development (except for the researcher). To ensure 'freedom of speech', the management team and staff have their own group sessions.

Result: Plan of Action

The HM-facilitator adds his own observations, experiences and assessment of (un)healthy implicit norms in the intervention group to the compressed analyses of hindrances and solutions offered in the in-depth interviews, open questionnaire and group sessions. This results in a plan of action that supposedly reflects opinions, perspectives and (feasible) wishes of employees in the intervention group (i.e. management and workers). The plan of action is sent to all employees, after consulting the Engine of Development and the management team, respectively. If the work-oriented solutions proposed also concern higher management (i.e. Executive Board or Board of Directors), then the management team is advised to inform higher management on the findings as well.

Phase 2: Implementation

In the implementation phase of the participatory action approach, the management has to take the first step. With support of the Engine of Development, the management team decides on which work-oriented solutions will be implemented. The solutions can be either evidence-based (e.g. adjusting classroom schedules), or new, tailor-made and specific to the context (e.g. adjusting the physical layout of class rooms). In any case, the management team explains to the employees which work-oriented solutions are (not) implemented and to what end. Furthermore, the method prescribes that the management team equips the plan of action with an implementation plan comprising amongst others a timeframe, a budget and the allocation of roles (e.g. the Engine of Development's role). If the management team wishes, the HM-facilitator will take up a role during the implementation phase, for example in monitoring the plan's progress or coaching the management team.

Primary and secondary outcomes

Measuring instruments: primary outcomes

Need for recovery

(Early) symptoms of fatigue at work are considered indicative of a 'need for recovery' [54]. Need for recovery after a working day is measured using a subscale of the Dutch Perception and Evaluation of Work Questionnaire (Dutch abbreviation: VBBA) [54]. The questionnaire comprises 11 dichotomous (*yes/no*) items and has proven to be valid and reliable (alpha 0.86) [54].

Vitality

Vital workers show high levels of energy and mental resilience, persist when facing difficulties and are willing to invest effort in their work [50]. Vitality is measured using the vigor subscale (3 items) of the Utrecht Work Engagement Scale-9 [55]. Responses are to be given on a seven point scale (0 = *never* to 6 = *always/every day*). The vigor subscale has shown acceptable validity and reliability in a sample across ten countries [50].

Measuring instruments: secondary outcomes

Job demands and resources

Several aspects of the job and its content are measured using subscales of the Dutch version of the Job Content Questionnaire (JCQ) [56]: psychological demands, coworker and supervisor support and decision authority. The psychological demand dimension measures the mental work load, organizational constraints on task completion and conflicting demands (5 items) [56]. The coworker and supervisor support subscales measure socio-emotional as well as instrumental support (8 items) [56]. Four identically phrased, but explorative items were added on socio-emotional and instrumental support of (higher) management. Decision authority or autonomy measures the workers' possibilities to make decisions about their work, mediated by organization factors (3 items) [56]. Response scales range from 1 = *strongly disagree* to 4 = *strongly agree*. The subscales of Dutch JCQ used in this study have shown acceptable scale reliability and validity [57].

Furthermore, possibilities for professional growth are assessed by a 6-item subscale of a Dutch questionnaire, developed for the (primary) educational sector (in Dutch: Welzijnscheck Onderwijspersoneel) [58], which has shown good divergent validity and reliability (alpha 0.87) [59]. Response scales range from 1 = *strongly disagree* to 5 = *strongly agree*.

Work ability

Work ability is based on the workers' balance between resources and demands and determines job performance now and in the near future [60]. Work ability

is measured using the Work Ability Index (WAI) [61]. The WAI is a self-report instrument and comprises seven dimensions on the physical and mental demands of work and the health and resources of the employee. For the current study, two of the seven dimensions were deemed relevant: 1) perceived current work ability, compared to lifetime best (1 item) and 2) perceived work ability related to mental job demands and perceived work ability related to physical job demands (2 items). Responses on (1) are recorded on a frequency scale from 0 (*unable to work*) to 10 (*very good*). Responses on (2) are recorded on a five-point frequency scale from 1 (*very good*) to 5 (*very bad*). Reliability and validity have been shown to be adequate in a Dutch sample (alpha 0.63 to 0.71) [62].

Job satisfaction

Job satisfaction is operationalized as workers' satisfaction with the job and its conditions. Job satisfaction is measured by two items of the Netherlands Working Conditions Survey 2010 [63]: to what extent are you, all things considered, satisfied with (1) your job, and (2) your working conditions? Response scales range from 1 = *very dissatisfied* to 5 = *very satisfied*.

Commitment

It has been shown that teacher commitment is a predictor of burnout, sickness absence and retention [64]. Therefore, commitment to work (2 items) and the organization (3 items) is measured in this study, using the Dutch questionnaire NOVA-WEBA [65, 66], which has shown moderate validity and reliability (alpha 0.68) [67]. Response scales range from 1 = *strongly disagree* to 5 = *strongly agree*.

Work engagement

Work engagement is defined as a positive, fulfilling, work-related state of mind that is characterized by vigor, dedication and absorption [55]. The Utrecht Work Engagement Scale (UWES) [55] is the most commonly used instrument to measure work engagement [68]. The 9-item version of UWES is used in this study, response scales range from 0 = *never* to 6 = *always/every day*. UWES-9 has shown good validity and reliability [69].

Health

Health is measured by asking a single item of the free version of SF-36-v2, named RAND-36 [70]. This single item measures perceived general health ("How do you rate your health in general?") on a five-point frequency-scale from 1 = *bad* to 5 = *excellent*. The subscale is considered valid and reliable (alpha 0.81) [70].

Sickness absence

Sickness absence is considered working less than normal hours or days due to illness, an incident or any other health reason. Sickness absence data will be collected in two ways: from company records as well as at baseline and follow-up measurements. At baseline and follow-up, presence, frequency and duration of sickness absence in the past 12 months is measured by three items

from the Netherlands Working Conditions Survey 2010 [63]. Furthermore, the self-reported cause of the last case of sickness absence is measured and whether this cause was attributed (fully, partly or not) to the work (NWCS 2011) [3].

Burnout

Burnout is measured with a slightly adjusted, Dutch version of the Maslach Burnout Inventory-General Survey (MBI-GS) [71], named Utrecht Burnout Scale (UBOS) [72]. This 16-item questionnaire includes the key dimensions of burnout: emotional exhaustion (feeling drained by the work), depersonalization (a cynical attitude towards the work and people working with) and reduced personal accomplishment (feeling incompetent at work). Response scales range from 0 = *never* to 6 = *every day*. Several studies have shown that the MBI-GS and its subscales are excellently reliable and valid (e.g. [73, 74]).

Inrole performance and knowledge and skills

Inrole performance is considered the achievement of work-related goals and measured by three items of the Netherlands Working Conditions Survey 2010 [63], with a response scale ranging from 1 = *strongly disagree* to 5 = *strongly agree*. Furthermore, the fit between current knowledge, skills and job tasks is measured by asking one item from NWCS 2011 [3]: "How do your knowledge and skills fit your current work?". Response scales range from 1 = *less knowledge and skills than needed*, to 2 = *it fits* to 3 = *more knowledge and skills than needed*.

Willingness and ability to prolong working life

The willingness and ability to prolong working life is measured by asking two open ended items from the Netherlands Working Conditions Survey 2010 [63]: Until what age do you (1) think you are able to continue working, and (2) want to continue working?

Productivity

Individual productivity in work is measured by a single item, based on module E of the PRODIDSQ [75]. PRODIDSQ is a scale considered to facilitate the validity of productivity costs estimates [76]. This single item measures self-rated productivity ("How would you assess your overall work performance in the past 4 weeks on a scale of 0 to 10?") from 0 = *worst quality* to 10 = *best quality*.

Measuring instruments: mediating factors

Occupational self-efficacy

Occupational self-efficacy is described as the confidence a worker has in his or her perceived ability to perform job tasks successfully [77]. The short (6 item) version of the Occupational Self-efficacy scale [78] measures the concept in a valid and reliable way (alpha 0.85) on a five-point Likert scale from 1 = *totally disagree* to 5 = *totally agree* [77].

Organizational efficacy

Organizational efficacy is defined as ‘an individual’s perception of the general capabilities of an organization’ ([79], p. 127). Van Vuuren’s seven item Organizational Efficacy Scale (OES) [79] has shown to measure the concept reliably (alpha 0.81) on a five-point Likert scale (1 = *strongly disagree*, 5 = *strongly agree*).

Measuring instruments: sociodemographic and profiling data

Sociodemographic data are collected at baseline, i.e.: gender; age; level of education; household composition; working hours per week; number of years working in current function, school and sector; main workplace location.

Profiling data are collected either at baseline or at follow-up measurements for a practical reason. A practical reason is that response rates would drop if questionnaires would be too intrusive. Profiling questions are considered stable over time, which makes the measurement moment less important.

Locus of Control

Locus of control is considered a personality trait and defined as the extent to which people believe they can influence the course of their lives [80]. The construct comprises two dimensions - internal and external locus of control - and can be measured with 8 items on a seven-point Likert-scale (1 = *totally disagree* to 7 = *totally agree*) [81]. The subscale internal locus of control has shown poor reliability (alpha 0.43) and the subscale external locus of control has shown moderate reliability (alpha 0.66) in a German population of youngsters [81]. The replication of the theoretical two factor structure in an exploratory factor analysis, indicates that the instrument’s validity might be more promising [81]. Despite the problematic reliability, the Nolte-scale was preferred over the original I-E scale by Rotter (23 items) [80], because it is more compact. The Nolte-scale is translated to Dutch by a native German speaker living in The Netherlands for long and then back-translated to German by a native Dutch speaker living in Germany for long.

Parent and pupil (mis)behavior

The extent to which employees are bothered by (mis)behavior of pupils and parents in their work, is measured by a Dutch scale designed for the (primary) educational sector, named Welzijnscheck Onderwijspersoneel [58]. Misbehavior of pupils is measured reliably (alpha 0.82) with six items on a six-point scale ranging from 1 = *not applicable* to 6 = *in a very great degree* [59]. On the same six-point scale, misbehavior of parents is measured reliably (alpha 0.78) with four items [59].

Work-life interference

Work-life interference is measured by one explorative item (“Does your work interfere with your private life?”) as well as life-work interference (“Does your private life interfere with your work?”), on a five point Likert scale ranging from 1 = *almost never* to 5 = *almost always*.

Institutional policy and educational quality

Workers’ knowledge of institutional policies is measured by one explorative item (“I am aware of the policies of my organization”) on a five-point Likert scale ranging from 1 = strongly disagree to 5 = strongly agree. The workers’ perception of the educational quality is also measured by one explorative item (“Our school prepares participants well for professional practice”) on a five-point Likert scale ranging from 1 = strongly disagree to 5 = strongly agree.

Data analysis

Effect evaluation

The effectiveness of the intervention on primary outcomes (need for recovery and vitality) and secondary outcomes at short term (T1), long term (T2) and corrected for baseline values, will be established by multilevel analyses. Repeated measurements on the worker-level and clustering of observations can thus be taken into account. The data will be analyzed at three levels: 1) worker, 2) department, and 3) school. Both crude and adjusted linear and logistic regression analyses will be conducted. The intention-to-treat principle is leading in all statistical analyses, meaning that the analyses are based on the initial treatment assignment and not on the treatment eventually received. However, per-protocol analysis will also be conducted, restricting the comparison to the ideal participants, in this study: participants that report taking part in at least two of the three steps of the needs assessment.

Multilevel analyses wherein T1 functions as the dependent variable, will be adjusted for possible confounding factors (e.g. experience, overtime). These variables will also be checked for effect modification at all measurement moments.

For all analyses, a two-tailed significance level of $p < 0.05$ will be considered statistically significant. The multilevel analyses will be conducted by means of MlwiN 2.0; linear and logistic regression analyses will be performed using SPSS 17.0 (SPSS Inc. Chicago, Illinois, USA).

Process evaluation

An extensive process evaluation will be conducted based on two complementary pillars: 1) Stecklar and Linnan’s framework [82] and its adaptations by Murta, Sanderson and Oldenburg [83], and 2) a selection of Randall, Nielsen and Tvedt’s Intervention Process Measure (IPM) [84]. The first pillar helps to answer instrumental questions concerning the intervention process (how does the intervention work?). The second pillar helps to identify participants’ appraisals of the intervention process (why does the intervention work?). A combined approach seems necessary to answer both the “how” and “why” question, since some studies have shown that an equal amount of ‘dose received’ can yield a range of heterogeneous individual appraisals of the inter-

vention [86]. And the appraisal of intervention processes may in turn influence intervention outcomes [85].

Stecklar and Linnan [82] propose seven components to determine how the intervention was implemented (Table 1): (i) recruitment (what procedures were used to interest workers and what are reasons for not participating?), (ii) reach (attendance of workers in each phase of the participatory approach and its consequent tailored measures), (iii) dose delivered (how many steps of the participatory approach were actually delivered by the facilitator?), (iv) dose received (how many steps of the participatory approach were actually followed by the worker?), (v) fidelity (was the participatory approach delivered according to protocol?), (vi) satisfaction (how satisfied are participants with the participatory approach?); and (vii) context (what organizational and environmental characteristics affect the intervention?).

Randall and colleagues [84] have shown that participants' appraisals of an intervention and its implementation can be measured quantitatively by the five scales of the Intervention Process Measure, of which four are used and adjusted to fit this study (Table 1): (a) line manager attitudes and actions, (b) exposure to components of the intended intervention, (c) employee involvement, and (d) employee readiness for change. Scores on these scales will be related to intervention outcomes (e.g. job satisfaction, well-being, and self-efficacy).

Six sources of data are used to assess the proposed process aspects (Table 1): (1) data logs by the researchers (recruitment, dose delivered, fidelity, context); (2) logs by the facilitator (dose received); (3) interviews at T1 with employees, management, participatory group and facilitator (reach, dose delivered, fidelity, satisfaction, context); (4) questionnaire at T1 (reach, dose received, fidelity, satisfaction, line manager attitudes and actions, exposure to components of the intended intervention, employee involvement and employee readiness for change); (5) separate group interviews with employees and management at T2 (reach, fidelity, satisfaction, context); (6) questionnaire at T2 (reach, fidelity, satisfaction, line manager attitudes and actions, exposure to components of the intended intervention, employee involvement and employee readiness for change). Data will be analyzed by either qualitative data software (e.g. Kwalitan) or by using a qualitative rating procedure.

Table 1 — Process evaluation components and (examples of) questions

Ref.	Components	Questions	Logs researcher	Logs facilitator	Interviews	T1	T2
[84,85]	Recruitment	what procedures were used to interest workers and what are reasons for not participating?	X				
[84,85]	Reach	attendance of workers in each phase of the participatory approach and its consequent tailored measures			X	X	X
[84,85]	Dose delivered	how many steps of the participatory approach were actually delivered by the facilitator?	X		X		
[84,85]	Dose received	how many steps of the participatory approach were actually followed by the worker?		X		X	
[84,85]	Fidelity	was the participatory approach delivered according to protocol?	X		X	X	X
[85]	Satisfaction	how satisfied are participants with the participatory approach?			X	X	X
[84,85]	Context	what organizational and environmental characteristics affect the intervention?	X		X		
[86]	Line manager attitudes and actions	e.g. "My immediate manager has done a lot to involve employees throughout the process"				X	X
[86]	Exposure to components of the intended intervention	e.g. "The project has made it easier to tackle the changes in the organization"				X	X
[86]	Employee involvement	e.g. "I had the opportunity to give my views about the change before it was implemented"				X	X
[86]	Employee readiness for change	e.g. "I looked forward to the changes brought about by the project"				X	X

Ethical considerations

The study protocol and materials are approved by TNO's Review Committee Participants in Experiments (RCPE), an internal ethics committee that assesses ethical aspects of working with participants in experiments. After review, the committee stated that in this study "the information is complete, participants can join voluntarily and an informed consent is provided". The RCPE has thus given a positive advice to the study's responsible manager, who decided to follow the positive approval by giving permission for performing the study. Hereafter we will elaborate on the information provision prior to and during the study.

Information provision prior to the study

Prior to the study, higher management of both schools signed a letter of intent to participate, cooperate and invest in kind. At the start of the project, a more detailed project plan will be presented to higher management in both schools where after working arrangements are made.

Information provision on measurement

Prior to baseline measurement employees in the intervention and control group will receive verbal and written information on the baseline measurement. Employees in the intervention group will be informed verbally by the participatory group (Engine of Development) by presentations at team meetings. Employees in the control group will be informed by their immediate supervisors during team meetings. Management in both the intervention and control group will be informed by the facilitator during a management meeting. All groups will receive a digital letter with information about the baseline measurement. Furthermore, the agenda of a team meeting will be cleared, to enable employees to fill out the questionnaire(s) during this time frame. All employees and managers participating in the baseline measurement are requested to sign an informed consent at the start of the questionnaire. By signing the informed consent participants declare amongst other things: 1) to have received information on the baseline measurement and the study, 2) to understand they can withdraw at any time from the study without reason. A similar procedure will be applied for the first follow-up (T1, 6 months) and second follow-up (T2, 18 months).

Information provision on participatory approach

Then, for both intervention groups a briefing will be held to announce the start of the participatory approach. During the briefing, the participatory approach is explained by the facilitator and the role of all participants is clarified. All questions can be asked. In addition, some of the results of the baseline measurement will be fed back during the meeting. The researchers' prior experiences in educational

institutions have learned that an enormous general scepticism towards survey research has to be overcome before support can be created. Demonstrating that results of the questionnaires are actually used for the better, helps to create support.

Information provision during the participatory approach

Continuously informing employees on the progress of each step is inherent to the participatory approach. The proceedings of each step are fed back to all staff of the intervention group, after suggestions from the participatory group and management are taken into account.

Discussion

In the educational sector, job demands have intensified at rapid pace in recent years, while job resources remained the same. The imbalance between demands and resources contributes to the development of mental health problems such as burnout. At the organizational level burnout resonates in increased sickness absence rates and problems with retention of experienced and novice teachers. Until now, most intervention studies that aimed to target these problems have been only partly effective, possibly because they focused on the individual level and applied secondary preventive interventions. Instead, it has been argued theoretically to focus on the organizational level and application of primary preventive interventions. The current research helps to translate this theoretical reasoning to empirical studies.

This is the first study to describe the test of a participatory, primary preventive organizational level intervention (i.e. a participatory action approach) on work-related stress and well-being that targets and engages all workers in the primary process of schools. The goal of this study is to determine whether the participatory action approach, which is supposed to result in tailored, work-oriented solutions on the balance between job demands and job resources, effectively influences need for recovery and vitality.

Strengths and limitations of the intervention

Less than a quarter of the intervention studies focus on primary preventive interventions, as assessed in a recent meta-analysis on stress interventions [35]. Thus, a first strength of the current intervention is the aim to alter job stress at its core (primary prevention).

Second, by making use of the participatory action approach, stakeholders at all levels are involved - teachers in the first place. This 'bottom-up' involvement of all stakeholders likely contributes to commitment to the proposed solutions. Solutions that can count on both the management and the work floor's commitment are more sustainable and thus more likely to have impact [37, 40].

A third strength is that the intervention is conducted in the same way for both schools, but the tailored workplace solutions can differ. This way we can compare different solutions on similar outcomes, further contributing to evidence-based practice.

Fourth, the project requires close collaboration between the intervention provider (i.e. the facilitator) and the researcher. They have a different task to fulfill within the project: the facilitator needs to make the intervention work, the researcher needs to make the study design work. Inevitably, their world views (or: paradigms) will meet and, maybe clash. An issue is that the intervention provider works from a 'practice-based evidence' perspective, asking himself at every step: is it useful? Is it important? Is it valid? Whereas researchers aim to contribute to evidence-based practice and therefore ask themselves: is it valid? Is it important? Is it useful? [86]. Even though the differing paradigms and activities will probably make the project difficult from time to time, the primary interest of both intervention provider and researcher is eventually the same: working happier and healthier in vocational education.

Strengths and limitations of the study

In line with earlier recommendations (e.g. [87]) the current study assesses all steps of the intervention process, that is: 1) the intended intervention, 2) intended changes in exposure or behavior (i.e. job demands and job resources), and 3) intended changes in study outcomes (i.e. need for recovery and vitality). From an (occupational) epidemiological point of view, the study can be classified as a *prevention-effectiveness study* [87, 88] as opposed to the *etiologic intervention study* (i.e. the most rigorous epidemiologic design, derived from the controlled clinical trial, studying disease and health outcomes [88]). Characteristics of the prevention-effectiveness study design are amongst others: small samples, no randomization or blinding, test of a program theory, quantitative and qualitative measures, case studies [89]. These characteristics ensure the internal validity of the study. Prevention effectiveness trials are at the core of evidence-based public health [88, 89]. The results of our study will be fed back to the vocational education council and policy makers in the field, helping them to 'practice evidence-based'. From a sociological, psychological or anthropological point of view, our study would rather be classified as a *pragmatic trial*, designed to find out 'how effective a treatment actually is in routine, everyday practice' [90]. Since we are conducting research in practice, several unforeseen events (e.g. reorganization) can take place, which might endanger the feasibility of the study. But our study is designed to reflect what happens in 'the real world' [90], maximizing external validity. This asks of the researchers and intervention providers to adjust to unexpected changes, while respecting the 'intention-to-treat-principle' (i.e. once allocated to the intervention or control group, always allocated to intervention or control group). Pragmatic trials allow for subtle variations in the intervention and research protocol, so to match the schools' context and needs. The researchers will not permit any variations in the protocol of the first phase of the intervention (needs assess-

ment), but broad variations are permitted in the second phase of the intervention. In the current study, points of view from the epidemiological, sociological and psychological disciplines are combined, so to maximize both internal and external validity [91].

A strength of the study is that both psychological outcomes (e.g. job satisfaction) as well as organizational outcomes (e.g. sickness absence) are taken into account, contrary to most intervention studies published [35]. The outcomes are assessed qualitatively (i.e. (group) interviews, observations, logs) and quantitatively (i.e. self-report measures in digital survey) and whenever possible complemented with 'objective' organizational data (e.g. sickness absence registration).

Besides strengths, the study also comprises possible limitations. A limitation of this study is the quasi-experimental design. However, the study's design – a controlled trial with departments allocated to conditions and two follow-up measurements – is as rigorous as possible in a practice-based study in (vocational) education. On the one hand, by giving higher management a vote in the choice for the intervention group, selection bias is possibly introduced. On the other hand, higher management's commitment to the study is assured and thereby relevance and feasibility of the study. A second limitation is the timeframe of the study. Behavioral and organizational changes do not come easy nor quickly. Therefore, the timing of follow-up measurements (six and eighteen months) might be too soon to establish the organizational changes and detect effects.

To conclude, Kristensen [87] reminds occupational intervention researchers of the 'simple fact that the purpose of workplaces is to produce goods and services – not to serve as arenas for intervention research'. The "Bottom-up Innovation" project will probably encounter numerous unexpected changes, but the design and research methods are chosen carefully, so to optimize both internal and external validity.

List of abbreviations

CT, controlled trial; HM, Heuristic Method (i.e. the intervention); IPM, Intervention Process Measure; JCQ, Job Content Questionnaire; MBI-GS, Maslach Burnout Inventory-General Survey; NOVA-WEBBA, Dutch questionnaire developed to identify risk factors for work stress; NWCS, Netherlands Working Conditions Survey; OES, Organizational Efficacy Scale, PAR, participatory action research; RAND-36, general health scale; PRODISQ, PROductivity and DISease Questionnaire; RCPE, Review Committee Participants in Experiments; SMI, stress management intervention; T0, baseline measurement; T1, first follow-up measurement; T2, second follow-up measurement; UBOS, Utrecht Burnout Scale; UWES-9, Utrecht Work Engagement Scale-9; VBBA, the Dutch Perception and Evaluation of Work Questionnaire; WAI, Work Ability Index.

References

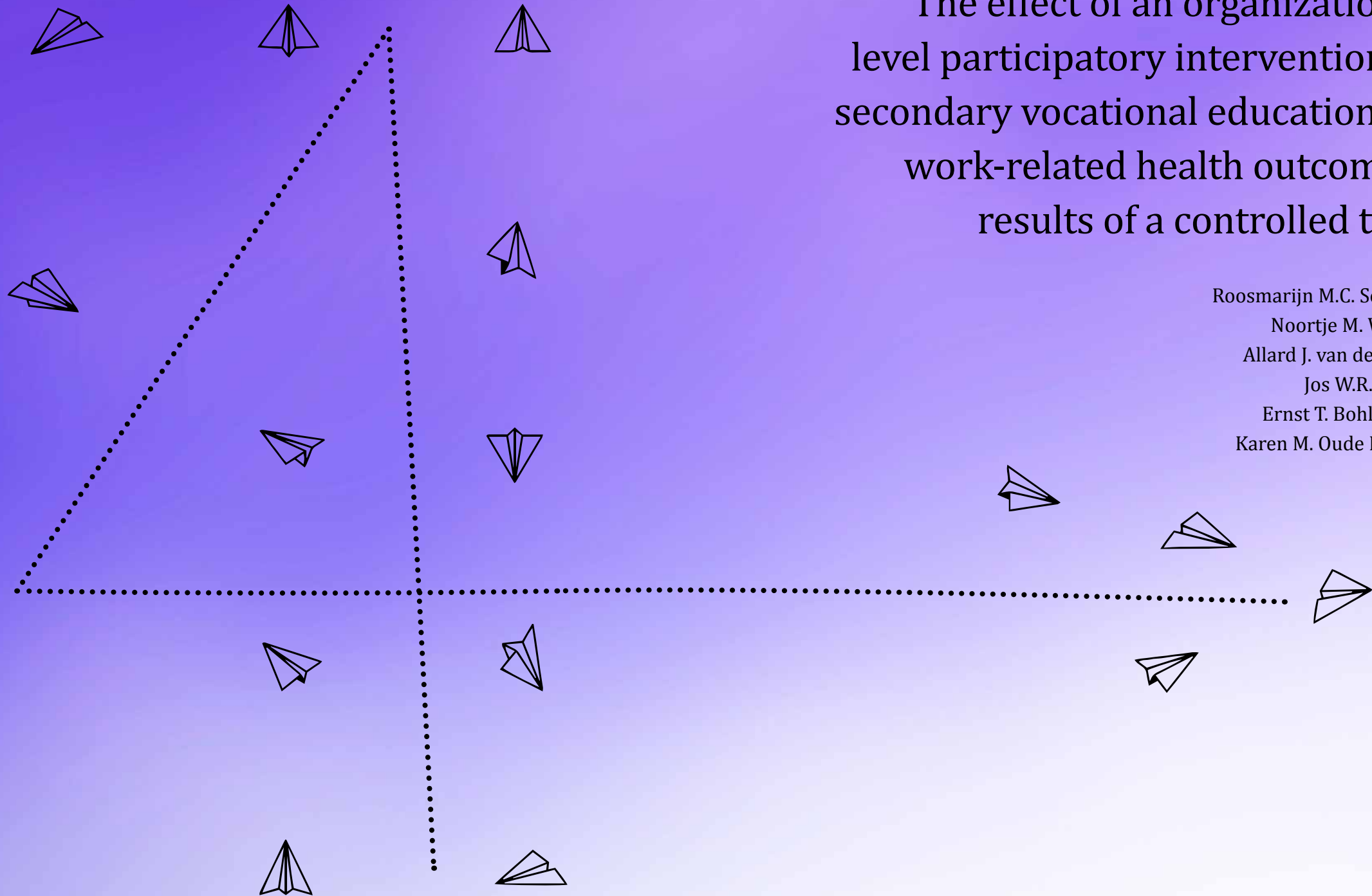
1. Rutte M, Samsom D: **Van goed naar excellent onderwijs Hoofdstuk VI.** In *Bruggen slaan, regeerakkoord VVD en PvdA*. Edited by Rutte M, Samsom D. 2012:16-17.
2. Van der Steeg M, Gerritsen S: *Teacher evaluations and pupil achievement: Evidence from classroom observations*: the Hague: CPB Netherlands Bureau for Economy Policy Analysis; 2013.
3. Koppes LLJ, De Vroome EMM, Mol MEM, Janssen BJM, van Zwieten MHJ, van den Bossche SNJ: *Netherlands Working Conditions Survey 2011: Methodology and overall results*: Hoofddorp: TNO; 2012.
4. Hughes RE: **Deciding to leave but staying: Teacher burnout, precursors and turnover.** *International Journal of Human Resource Management* 2001, **12**(2):288-298.
5. Keijsers GJ, Schaufeli WB, Le Blanc PM, Zwerts C, Miranda DR: **Performance and burnout in intensive care units.** *Work Stress* 1995, **9**(4):513-527.
6. Maslach C, Schaufeli WB, Leiter MP: **Job burnout.** *Annual Review of Psychology* 2001, **52**:397-422.
7. Demerouti E, Nachreiner F, Bakker AB, Schaufeli WB: **The job demands-resources model of burnout.** *J Appl Psychol* 2001, **86**(3):499-512.
8. Maslach C: **Job Burnout: New Directions in Research and Intervention.** *Current Directions in Psychological Science* 2003, **12**(5):189-192.
9. Hakanen JJ, Bakker AB, Schaufeli WB: **Burnout and work engagement among teachers.** *J School Psychol* 2006, **43**(6):495-513.
10. Easthope C, Easthope G: **Intensification, extension and complexity of teachers' workload.** *British Journal of Sociology of Education* 2000, **21**(1):43-58.
11. Timms C, Graham D, Cottrell D: **"I just want to teach": Queensland independent school teachers and their workload.** *Journal of Educational Administration* 2007, **45**(5):569-586.
12. Fink D: **The law of unintended consequences: the 'real' cost of top-down reform.** *Journal of Educational Change* 2003, **4**(2):105-108.
13. Howe J: **Nine 'til Three? Not Likely! Teacher: The National Education Magazine** 2005:41-43.
14. Timms C, Graham D, Caltabiano M: **Perceptions of school administration trustworthiness, teacher burnout/job stress and trust: the contribution of morale and participative decision-making.** In *Advances in Organizational Psychology*. Edited by Anonymous Bowen Hills, QLD, Australia: Australian Academic Press; 2007:135-151.
15. Seibt R, Lützkendorf L, Thinschmidt M: **Risk factors and resources of work ability in teachers and office workers.** *International Congress Series* 2005, **1280**:310-315.
16. Weber A, Weltle D, Lederer P: **Ill health and early retirement among school principals in Bavaria.** *Int Arch Occup Environ Health* 2005, **78**(4):325-331.
17. Wegner R, Berger P, Poschadel B, Manuwald U, Baur X: **Burnout hazard in teachers results of a clinical-psychological intervention study.** *Journal of Occupational Medicine and Toxicology* 2011, **6**(1).
18. Pas ET, Bradshaw CP, Hershfeldt PA: **Teacher- and school-level predictors of teacher efficacy and burnout: Identifying potential areas for support.** *J School Psychol* 2012, **50**(1):129-145.
19. De Graaf R, Tuijthof M, Dorsselaer S: *Verzuim door psychische en somatische aandoeningen bij werkenden. resultaten van de 'Netherlands Mental Health Survey and Incidence Study-2' (NEMESIS-2)*: Utrecht: Trimbos Instituut; 2011.
20. De Graaf R, Tuijthof M, Van Dorsselaer S, Ten Have M: **Comparing the effects on work performance of mental and physical disorders.** *Soc Psychiatry Psychiatr Epidemiol* 2012, **47**(11):1873-1883.
21. Murphy LR, Sauter SL: **The USA perspective: Current issues and trends in the management of work stress.** *Australian Psychologist* 2003, **38**(2):151-157.
22. Arthur AR: **Employee assistance programmes: The emperor's new clothes of stress management?** *British Journal of Guidance and Counselling* 2000, **28**(4):548-559.
23. Giga SI, Cooper CL, Faragher B: **The Development of a Framework for a Comprehensive Approach to Stress Management Interventions at Work.** *International Journal of Stress Management* 2003, **10**(4):280-296.
24. Unterbrink T, Pfeifer R, Krippel L, Zimmermann L, Rose U, Joos A, Hartmann A, Wirsching M, Bauer J: **Burnout and effort-reward imbalance improvement for teachers by a manual-based group program.** *Int Arch Occup Environ Health* 2012, **85**(6):667-674.
25. Gold E, Smith A, Hopper I, Herne D, Tansey G, Hurland C: **Mindfulness-based stress reduction (MBSR) for primary school teachers.** *J Child Fam Stud* 2010, **19**(2):184-189.
26. Franco C, Mañas I, Cangas AJ, Moreno E, Gallego J: **Reducing teachers' psychological distress through a mindfulness training program.** *Spanish Journal of Psychology* 2010, **13**(2):655-666.
27. Poulin PA, Mackenzie CS, Soloway G, Karayolas E: **Mindfulness training as an evidenced-based approach to reducing stress and promoting well-being among human services professionals.** *International Journal of Health Promotion and Education* 2008, **46**(2):72-80.
28. Zołnierczyk-Zreda D: **An intervention to reduce work-related burnout in teachers.** *International journal of occupational safety and ergonomics : JOSE* 2005, **11**(4):423-430.
29. Buunk BP, Schaufeli WB: **Reciprocity in Interpersonal Relationships: An Evolutionary Perspective on Its Importance for Health and Well-being.** *European Review of Social Psychology* 1999, **10**(1):259-291.
30. Leiter MP, Maslach C: **The impact of interpersonal environment on burnout and organizational commitment.** *J Organ Behav* 1988, **9**(4):297-308.
31. Maslach C, Goldberg J: **Prevention of burnout: New perspectives.** *Applied and Preventive Psychology* 1998, **7**(1):63-74.
32. McVicar A, Munn-Giddings C, Seebohm P: **Workplace stress interventions using participatory action research designs.** *International Journal of Workplace Health Management* 2013, **6**(1):18-37.
33. Caulfield N, Chang D, Dollard MF, Elshaug C: **A review of occupational stress interventions in Australia.** *International Journal of Stress Management* 2004, **11**(2):149-166.
34. Van der Klink JLL, Blonk RWB, Schene AH, Van Dijk FJH: **The benefits of interventions for work-related stress.** *Am J Public Health* 2001, **91**(2):270-276.
35. Richardson KM, Rothstein HR: **Effects of Occupational Stress Management Intervention Programs: A Meta-Analysis.** *J Occup Health Psychol* 2008, **13**(1):69-93.
36. Hasson H, Gilbert-Ouimet M, Baril-Gingras G, Brisson C, Vézina M, Bourbonnais R, Montreuil S: **Implementation of an organizational-level intervention on the psychosocial environment of work: Comparison of managers and employees views.** *Journal of Occupational and Environmental Medicine* 2012, **54**(1):85-91.
37. Nielsen K, Taris TW, Cox T: **The future of organizational interventions: Addressing the challenges of today's organizations.** *Work Stress* 2010, **24**(3):219-233.
38. Griffiths A: **Organizational interventions. Facing the limits of the natural science paradigm.** *Scandinavian Journal of Work, Environment and Health* 1999, **25**(6):589-596.
39. Dollard MF, Le Blanc PM, Cotton SJ: **Participatory action research as work stress interventions.** In *The individual in the changing working life*. Edited by Naswall K, Hellgren J, Sverke M. Oxford: Cambridge University Press; 2008:353-379.
40. Nielsen K, Randall R, Holten A-, González ER: **Conducting organizational-level occupational health interventions: What works?** *Work Stress* 2010, **24**(3):234-259.
41. Bandura A: *Social foundations of thought and action : a social cognitive theory*: Englewood Cliffs, N.J.: Prentice-Hall; 1986.
42. Bandura A: **Cultivate self-efficacy for personal and organizational effectiveness.** In *Handbook of principles of organizational behavior : indispensable knowledge for evidence-based management*. Edited by Locke EA. Chichester, Sussex; [Hoboken]: John Wiley & Sons; 2009:120-136.
43. Bandura A: **Health promotion by social cognitive means.** *Health Education and Behavior* 2004, **31**(2):143-164.
44. Bandura A: **Health promotion from the perspective of social cognitive theory.** *Psychology and Health* 1998, **13**(4):623-649.
45. Consiglio C, Borgogni L, Alessandri G, Schaufeli WB: **Does self-efficacy matter for burnout and sickness**

- absenteeism? The mediating role of demands and resources at the individual and team levels.** *Work Stress* 2013, **27**(1):22-42.
46. Semmer NK: **Job stress interventions and organizational at work.** In *Handbook of occupational health psychology*. Edited by Quick JC, Tetrick LE. Washington: APA; 2003:.
47. Cox T, Griffiths A, Barlow C, Randall R, Thompson T, Rial-Gonzalez E: *Organisational interventions for work stress: a risk management approach*. [Sudbury]: HSE Books; 2000.
48. Baarda,D.B.,Goede,M.P.M.de: *Basisboek Methoden en technieken: handleiding voor het opzetten en uitvoeren van kwantitatief onderzoek*: 4th ed. Groningen [etc.]: Wolters-Noordhoff; 2006.
49. Rosenbaum PR, Rubin DB: **The central role of the propensity score in observational studies for causal effects.** *Biometrika* 1983, **70**(1):41-55.
50. Schaufeli WB, Bakker AB, Salanova M: **The measurement of work engagement with a short questionnaire: A cross-national study.** *Educational and Psychological Measurement* 2006, **66**(4):701-716.
51. Schaufeli WB, Bakker AB: **Bevlogenheid: een begrip gemeten.** *Gedrag en organisatie* 2004, **17**:90-112.
52. Cohen J: *Statistical Power Analysis for the Behavioral Sciences*: Revised Edition ed. New York: Academic press; 1977.
53. Heijting E: *Proces-handboek van De Heurische Methode (DHM)*: Utrecht: Heijting en Weerts B.V.; 2011.
54. Van Veldhoven M, Broersen S: **Measurement quality and validity of the "need for recovery scale".** *Occup Environ Med* 2003, **60**(SUP-PL 1):i3-i9.
55. Schaufeli WB, Salanova M, González-Roma V, Bakker AB: **The measurement of engagement and burnout: a two sample confirmatory factor analytic approach.** *Journal of Happiness Studies* 2002, **3**:71-92.
56. Karasek R, Brisson C, Kawakami N, Houtman I, Bongers P, Amick B: **The Job Content Questionnaire (JCQ): an instrument for internationally comparative assessments of psychosocial job characteristics.** *J Occup Health Psychol* 1998, **3**(4):322-355.
57. Pelfrene E, Vlerick P, Mak RP, De Smet P, Kornitzer M, De Backer G: **Scale reliability and validity of the Karasek 'Job Demand-Control-Support model in the belstress study.** *Work Stress* 2001, **15**(4):297-313.
58. TNO Kwaliteit van Leven: *Handleiding Welzijnscheck Onderwijspersoneel*: Heerlen: Stichting Vervangingsfonds en Bedrijfsgezondheidszorg voor het Onderwijs; 2007.
59. Kraan KO, Bakhuis Roozeboom MC, Wiezer NM: *Kwaliteit 'Welzijnscheck Onderwijspersoneel' en Validiteit Afkappunten.* *TNO rapportage vertrouwelijk*: Hoofddorp: TNO; 2010.
60. Ilmarinen J, Gould R, Järviskoski A, Järvisalo J: **Diversity of work ability.** In *Dimensions of work ability: results of the health 2000 survey*. Edited by Gould R, Ilmarinen J, Järvisalo J, Koskinen S. Helsinki: Finnish Centre for Pension [u.a.]; 2008:13-24.
61. Tuomi K, Ilmarinen J, Jahkola A: *Work ability index*: 2nd revised ed ed. Helsinki: Finnish Institute of Occupational Health; 1998.
62. Radkiewicz P, Widerszal-Bazyl M: **Psychometric properties of Work Ability Index in the light of comparative survey study.** In Edited by Anonymous 2005:304-309. [Anonymous, vol 1280.]
63. Koppes LLJ, De Vroome EMM, Mol MEM, Janssen BJM, van den Bossche SNJ: *Netherlands Working Conditions Survey 2010: Methodology and overall results*: Hoofddorp: TNO; 2011.
64. Day C, Elliot B, Kington A: **Reform, standards and teacher identity: Challenges of sustaining commitment.** *Teaching and Teacher Education* 2005, **21**(5):563-577.
65. Dhondt,S.,Houtman,I.L.D.: *NIPG research questionnaire Job Content-WEBA (NOVA-WEBA): construction and a first test on reliability and validity*: Leiden: Nederlands Instituut voor Praeventieve Gezondheidszorg TNO; 1992.
66. Houtman ILD, Bloemhoff A, Dhondt S, Terwee T: *WEBA and NOVA-WEBA in relation to health and well-being of employees*: Leiden: Nederlands Instituut voor Praeventieve Gezondheidszorg TNO; 1994.
67. Kraan KO, Dhondt S, Houtman ILD, Nelemans R, De Vroome EMM: *Handleiding NOVA-WEBA. Een vragenlijst om arbeidsorganisatorische kenpunten op te sporen. [alleen voor gebruik binnen TNO Arbeid]*: Hoofddorp: TNO Arbeid; 2000.
68. Schaufeli WB: **Work engagement: What do we know and where do we go?** *Romanian Journal of Applied Psychology* 2012, **14**(1):3-10.
69. Seppälä P, Mauno S, Feldt T, Hakanen J, Kinnunen U, Tolvanen A, Schaufeli W: **The construct validity of the Utrecht Work Engagement Scale: Multisample and longitudinal evidence.** *Journal of Happiness Studies* 2009, **10**(4):459-481.
70. Van der Zee KI, Sanderman R: *Het meten van de algemene gezondheidstoestand met de RAND: een handleiding*: Groningen: Noordelijk Centrum voor Gezondheidsvraagstukken,Rijksuniversiteit Groningen; 1993.
71. Schaufeli WB, Leiter MP, Maslach CJ,S.E.: **Maslach burnout inventory—general survey.** In *Maslach burnout inventory manual*. 3rd edition. Edited by Maslach C, Jackson SE, Leiter MP. Palo Alto, Calif.: Consulting Psychologists Press; 1996:22-26.
72. Schaufeli WB, van Dierendonck D: *Handleiding van de Utrechtse Burnout Schaal (ubos)*: Lisse: Swets & Zeitlingen; 2000.
73. Bakker AB, Demerouti E, Schaufeli WB: **Validation of the Maslach Burnout Inventory - General survey: An Internet study.** *Anxiety, Stress and Coping* 2002, **15**(3):245-260.
74. Schutte N, Toppinen S, Kalimo R, Schaufeli WB: **The factorial validity of the Maslach Burnout Inventory-General Survey (MBI-GS) across occupational groups and nations.** *J Occup Organ Psychol* 2000, **73**(1):53-66.
75. Koopmanschap M, Meerding W, Evers S, Severens J, Burdorf A, Brouwer W: *Handleiding voor het gebruik van PRODISQ. Een modulaire vragenlijst over de relatie tussen ziekte en productiviteitskosten. Toepasbaar bij economische evaluaties van gezondheidszorgprogramma's voor patiënten en werknemers*: Rotterdam;Maastricht: RMEU;EMCU; 2004.
76. Koopmanschap MA: **PRODISQ: A modular questionnaire on productivity and disease for economic evaluation studies.** *Expert Review of Pharmacoeconomics and Outcomes Research* 2005, **5**(1):23-28.
77. Rigotti T, Schyns B, Mohr G: **A short version of the occupational self-efficacy scale: Structural and construct validity across five countries.** *Journal of Career Assessment* 2008, **16**(2):238-255.
78. Schyns B, Von Gernot C: **A new occupational self-efficacy and its relation to personality constructs and organizational variables.** *European Journal of Work and Organizational Psychology* 2002, **11**(2):219-241.
79. Van Vuuren M, Elving WJ, De Jong MDT, Seysel ER: **What keeps you here? Organizational efficacy and person-organization fit as antecedents organizational commitment.** In *Why work? : the contribution of value perceptions and efficacy expectations to organizational commitment*. Edited by Van Vuuren HA. S.l.: s.n.; 2006:121-147.
80. Rotter JB: **Generalized expectancies for internal versus external control of reinforcement.** *Psychol Monogr* 1966, **80**(1):1-28.
81. Weinhardt M, Schupp J: **Kontrollüberzeugungen (Locus of control).** In *Multi-Itemskalen im SOEP Jugendfragebogen*. Edited by Anonymous Berlin: Institute for Economic Research; 2011:30-36.
82. Steckler A, Linnan L: *Process evaluation. For public health intervention and research*: San Francisco: Jossey-Bass; 2002.
83. Murta SG, Sanderson K, Oldenburg B: **Process evaluation in occupational stress management programs: A systematic review.** *American Journal of Health Promotion* 2007, **21**(4):248-254.
84. Randall R, Nielsen K, Tvedt SD: **The development of five scales to measure employees' appraisals of organizational-level stress management interventions.** *Work Stress* 2009, **23**(1):1-23.
85. Randall R, Cox T, Griffiths A: **Participants' accounts of a stress management intervention.** *Human Relations* 2007, **60**(8):1181-1208.
86. Green LW: **Public health asks of systems science: To advance our evidence-based practice, can you help us get more practice-based evidence?** *Am J Public Health* 2006, **96**(3):406-409.
87. Kristensen TS: **Intervention studies in occupational epidemiology.** *Occup Environ Med* 2005, **62**(3):205-210.
88. Skov T, Kristensen TS: **Etiologic and prevention effectiveness intervention studies in occupational health.** *Am J Ind Med* 1996, **29**(4):378-381.

89. Haddix AC: *Prevention effectiveness : a guide to decision analysis and economic evaluation*: second ed. New York: Oxford University Press; 2003.
90. MacPherson H: **Pragmatic clinical trials**. *Complement Ther Med* 2004, **12**(2-3):136-140.
91. Hotopf M: **The pragmatic randomised controlled trial**. *Advances in Psychiatric Treatment* 2002, **8**(5):326-333.

The effect of an organizational level participatory intervention in secondary vocational education on work-related health outcomes: results of a controlled trial

Roosmarijn M.C. Schelvis
Noortje M. Wiezer
Allard J. van der Beek
Jos W.R. Twisk
Ernst T. Bohlmeijer
Karen M. Oude Hengel



Abstract

Background

Work-related stress is highly prevalent in the educational sector. The aim of the current study was to evaluate the effectiveness of an organizational level, participatory intervention on need for recovery and vitality in educational workers. It was hypothesized that the intervention would decrease need for recovery and increase vitality.

Methods

A quasi-experiment was conducted at two secondary Vocational Education and Training schools (N=356) with 12- and 24-months follow-up measurements. The intervention consisted of 1) a needs assessment phase, wherein staff and teachers developed actions for happy and healthy working under supervision of a facilitator, and 2) an implementation phase, wherein these actions were implemented by the management teams. Mixed model analysis was applied in order to assess the differences between the intervention and control group on average over time. All analyses were corrected for baseline values and several covariates.

Results

No effects of the intervention were found on need for recovery, vitality and most of the secondary outcomes. Two small, statistically significant effects were in unfavorable direction: the intervention group scored on average over time significantly lower on absorption (i.e. a subscale of work engagement) and organizational efficacy than the control group.

Conclusions

Since no beneficial effects of this intervention were found on the primary and most of the secondary outcomes, further implementation of the intervention in its current form is not eligible. We recommend that future organizational level interventions for occupational health 1) incorporate an elaborate implementation strategy, 2) are more specific in relating actions to stressors in the context, and 3) are integrated with secondary preventive, individual focused stress management interventions.

Background

Twenty percent of the EU workers consider their health to be at risk as a result of work-related stress [1]. Work-related stress is especially common among workers in education throughout both the eastern [2, 3] and western developed world [4]. According to a report for the Health & Safety Executive the stress levels of teachers were more than double (42%) compared to those in other occupations [5]. Also in the Netherlands, one in five employees suffer from work-related stress, according to a representative survey [6]. In secondary Vocational Education and Training (VET) this would equal to 11,174 of the currently employed 52,456 workers [7]. These workers feel emotionally drained and exhausted, especially at the end of the work day, and tired when they get up again in the morning [6]. For 6.9% of the workers in Dutch education, work-related stress results in being overworked or burned out, including long term sick leave [6].

Work-related stress may show in workers as decreased vitality and increased need for recovery, the primary outcomes of this study. High vitality is defined as having high levels of energy and resilience, persisting in the face of difficulties, and willingness to invest effort in the work [8]. Need for recovery is the extent to which employees experience problems in the recovery of efforts at work, and is hence indicative of early symptoms of fatigue at work [9]. Work-related stress and reduced well-being can result from an imbalance between two types of workplace characteristics: job demands and job resources. The physical, social or organizational aspects of the job that require sustained physical or psychological effort are considered job demands, whereas job resources are those work aspects that may reduce job demands, help to achieve goals and stimulate learning and development [10]. Over the years, job demands have intensified in the educational sector [11], while job resources remained the same. Examples of increasing job demands are the growth in accountability measures [12] or the integration of students with special needs in the regular classes [13]. Job demands and job resources have to be balanced in order to prevent stress [14, 15]. A job demand, such as dealing with students with special needs, will turn into a stressor over time if job resources, such as coworker support, are insufficient or lacking [16, 17]. The imbalance between demands and resources likely contributed to the current work-related stress prevalence.

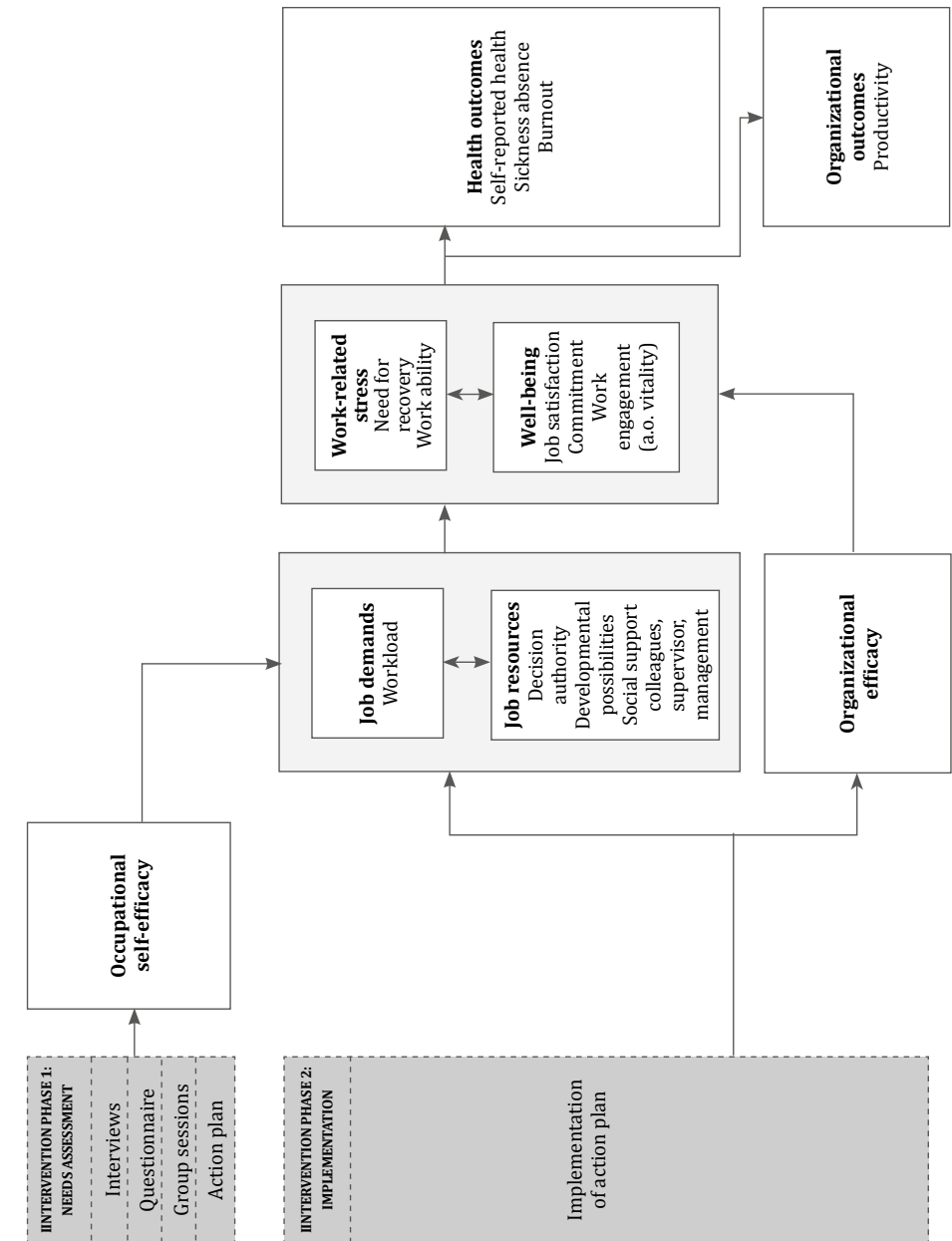
An imbalance between job demands and job resources can be restored by primary preventive interventions, which aim to alter the source of the stress at work. However, an overview of stress management interventions showed that typical stress management interventions mostly aimed to reduce stress symptoms (i.e. secondary prevention)[18]. Moreover, the existing interventions were targeted at the individual level and comprised stress management and coping techniques [18], whereas a review demonstrated that the organizational level is to be preferred for implementing stress management interventions, because organizational level interventions are more likely to bring about positive changes than the individual level interventions [19]. To date,

the interventions targeted at workers in education primarily aimed to enhance the individual capacity of (trainee) teachers or teaching assistants to cope with stressors in the workplace, for example via mindfulness-based stress reduction or workshops on stress management skills [20-24]. These interventions were only partly effective in influencing work-related stress or (dimensions of) burnout [20-24] and well-being [24].

An organizational level intervention focuses on changing stressors in the work environment, rather than changing the response of employees to stressors, and the change consists of altering some aspect of the organization (e.g. roles, structure) [25]. However, more is needed than just applying a primary preventive, organizational level intervention to render effective results [26, 27]. First, participation of stakeholders is acknowledged as one of the most desirable intervention strategies [28], since it leads to a feeling of joint ownership of both problems and solutions and thereby increases implementation and long term adherence. Secondly, self-efficacy beliefs of the target group are of importance for interventions targeted at changing the root-cause of stress [29]. Self-efficacy is 'the belief in one's own ability to master specific domains in order to produce given attainments' [30, 31]. High self-efficacy would help workers to create a 'control over circumstances mindset' [32]. The most effective way to enhance self-efficacy is by providing a mastery experience, and it was assumed that taking part in the first phase of the intervention leads to this experience of mastery.

The aim of the current study was to evaluate the long term effectiveness of an organizational level, primary preventive, participatory intervention on need for recovery and vitality primarily. The hypothesized order of expected changes is that participating in the intervention's first phase (needs assessment) would result directly in participant's increased occupational self-efficacy (proximal effect, Figure 1). Implementation of intervention activities (the intervention's second phase) would increase organizational efficacy and job resources (i.e. decision authority, developmental possibilities and various forms of social support) and reduce job demands (i.e. psychological demands), these are the expected intermediate effects (Figure 1). And if the balance between job demands and job resources is restored, distal effects are supposedly to be found on work-related stress constructs (i.e. need for recovery and work ability) and well-being constructs (i.e. work engagement including vitality, job satisfaction and commitment; Figure 1).

Figure 1 — Model representing the logic order of expected changes



Methods

Study design

The effectiveness of the intervention was evaluated in a controlled trial with a matched control group and three measurements (T0, T1, and T2). First follow-up measurement was conducted at 12 months after baseline (T1) and second follow-up at 24 months after baseline (T2). Data were collected via a questionnaire constructed with online survey software, participants received a link to the questionnaires in their mailbox. To increase the response rate at T1 and T2, an incentive (i.e. a warehouse gift card) was sent to respondents in the intervention and control group.

This trial was registered in the Netherlands Trial Register (NTR3284). The study has been approved by TNO's Review Committee Participants in Experiments (RCPE), an internal ethics committee that assesses ethical aspects of working with participants in experiments. The RCPE advised positively on the study to the responsible manager since the committee perceived "the information to be complete, participants can join voluntarily and an informed consent is provided" [33]. The manager decided to follow the RCPE's approval by permitting the study.

Study population

Two Vocational Education and Training (VET) schools were recruited via a mailing by the sector organization, The Netherlands Association of VET Colleges. A high sickness absence rate within a certain department was the most important reason to participate in this study, according to the Executive Boards of both schools. Therefore these two departments were selected as intervention groups by the schools. According to the directors of departments that were selected as intervention groups, their concerns about the situation in their department and a notion of their employees' diminishing happiness at work were important deliberations to participate. The researchers matched a control group within the same school to each intervention group, based on department size, mean age and type of work. In total, four departments were included. Since the intervention and control groups were situated in different locations we consider diffusion of treatment effects to be unlikely. All teaching and non-teaching (i.e. educational and administrative support staff) employees and their managers in these departments were invited to participate in the study. Employees who worked within the school, but did not teach at a secondary vocational level were excluded. Informed consent was obtained from all individual participants included in the study.

Matching, blinding and sample size

In each school, one department was selected as the experimental group by the participating schools, since their motive to participate in this study was to solve a problem or reach a goal within a specific department. To reduce the negative impact of selection bias, the control groups were obtained according

to the 'general control' matching principle [34] on the criteria: department size (at least 150 employees), age composition of staffing, and type of work (i.e. teaching vocational students and not secondary school pupils). Blinding of the participants and intervention providers was impossible due to the participatory nature of the intervention.

The sample size calculation was based on the number of cases required to detect an effect (Cohen's $d = 0.2$) on the primary outcome vitality, as measured with the 3-item subscale 'vigor' of the Utrecht Work Engagement Scale-9 (UWES-9) [35]. With a power of 80%, a two-sided alpha of 5%, the required sample size is 385, which translates to 193 participants per school and 97 per group. With an expected loss to follow-up of 35% over 24 months, a total sample size of 600 was needed at baseline. The sample size calculation has been described extensively elsewhere [33].

Intervention

The intervention was a participatory action approach applied at the organizational level, named the Heuristic Method (HM). HM was developed by a Dutch consultancy firm and piloted over a hundred times in public and private organizations before evaluation within this trial. The consultancy firm refined the intervention after each application, based on the lessons learned. Although the customers were almost always satisfied with the intervention's results, the intervention effects were never tested scientifically. The intervention consisted of two 12-month phases: (i) a phase of needs assessment, and (ii) an implementation phase.

In the needs assessment phase, staff and teachers developed actions to 'work happily and healthily', under supervision of an HM facilitator. The HM facilitator held expertise in organizational change processes, and he used the management's and workers' knowledge, skills and perceptions to thoroughly determine what hindered and facilitated 'healthy and happy working' in the organization. A participatory work group was formed, its members were ambassadors of the project and assisted the facilitator (e.g. by approaching interviewees or by proof reading reports).

The HM facilitator then led three iterative steps to complete the needs assessment by: (i) approximately ten one-hour interviews with typical optimistic and typical critical teachers and staff; (ii) a digital open-ended questionnaire for all workers; and (iii) group sessions with all teams, chaired by members of the participatory group. The result of each step in the intervention determined the content of the following step. Reports of each step were written by the HM facilitator and were shared with all workers, starting with the participatory work group, then the management team, and finally with all workers. The third and last report contained the facilitator's advice to the management team on intervention activities to be implemented in the next phase. Examples of these activities were: creating a staff room or implementing performance assessment policies (see Table 1 for an overview of all problems, goals and intervention activities).

Table 1 — Results of the needs assessment and translation into action plan

	Main content of advisory report delivered by facilitator	Main content of action plan ^a constructed by management team
School A	(i) professionalize the teams; (ii) professionalize the management; (iii) improve the administrative support and facilities.	<p>The director, assisted by an HM consultant, translated the recommendations into an action plan with three goals, six changes and a set of quick wins.</p> <p>Goals: i) unambiguous management control; ii) competence and professionalism in the teams, and iii) adequate facilities</p> <p>Changes: (i) compliance to the workload policy, (ii) structured performance reviews; (iii) a continuous dialogue on the organization of the educational programs; (iv) a leading team activities plan; (v) weekly work meetings; and (vi) personalized competence development plans.</p> <p>Quick wins: create adequate facilities by creating a staff room at both locations; place extra walls in some classrooms; place beamers in all class rooms; improve the service by the facilitation services office.</p>
School B	(i) create adequate and effective management control by installing a management team that is approachable, coaching, and leading; (ii) make teams the central executive units by developing a team program; (iii) eliminate cumbersome administrative procedures.	<p>The directors of the management team decided to integrate the facilitator's recommendations in the annual agreements (i.e. a management contract) she made with the Executive Board, instead of writing a separate action plan. A coach was attracted to support teams in a previously initiated change towards becoming self-managing.</p> <p>Goals were formulated in four headlines: i) strategy; ii) education; iii) personnel; iv) organization; and v) business operations.</p> <p>The most important change per headline was: i) alliances with partners in the region are closed; ii) the curriculum of two educations are reconstructed into units of learning; iii) performance review policies are implemented; iv) teams function as self-managing units; and v) a multi-annual housing plan is developed.</p> <p>No quick wins were formulated.</p>

^a Action plan was termed 'Management Contract' in school B.

In the implementation phase, the intervention activities were implemented by the management teams at both schools. HM prescribed that the management team translated the facilitator's advisory report into an action plan, containing an implementation plan, comprising at least a timeframe, a budget and the allocation of roles. Assistance by the HM facilitator could be provided if the management had the means to temporarily hire a consultant.

Primary outcome measures

Primary outcomes were an indicator of work-related stress (i.e. need for recovery) and well-being (i.e. vitality).

Need for Recovery

The concept was assessed using a subscale of the Dutch Perception and Evaluation of Work Questionnaire [9]. The scale comprises 11 dichotomous (yes/no) statements such as "My job makes me feel rather exhausted at the end of a work day". The need for recovery scale ranges from 0 to 100, calculated as the number of points (1=yes, 0=no) divided by the number of questions answered, multiplied by 100. Higher scores indicate a higher need for recovery, which is unfavorable. The questionnaire has proven to be valid and reliable (Cronbach's alpha 0.86) [9]. In the current study, internal consistency was excellent (Cronbach's alpha: 0.89).

Vitality

Vitality was assessed using the 3-item vigor subscale of the Utrecht Work Engagement Scale-9 (UWES-9; e.g. "At my job, I feel strong and vigorous") [35]. Response scales range from 0 (never) to 6 (always/every day). The subscale scores were obtained by calculating the mean (range 0-6). Higher scores are indicative of more vigor. The total UWES-9 has shown good validity and reliability [36], as was the case for the subscale in this study (Cronbach's alpha: 0.87).

Secondary outcome measures

Several categories of secondary outcomes were measured: job demands, job resources, indicators of work-related stress, well-being and efficacy. Job demands were operationalized as psychological demands and job resources as decision authority, developmental possibilities and various forms of social support. Work-related stress was indicated as reduced work ability, well-being was indicated by work engagement, job satisfaction and commitment. Two efficacy or competence measures were taken into account: occupational self-efficacy and organizational efficacy.

Psychological demands

A five item subscale of the Dutch version of the Job Content Questionnaire (JCQ) was used to measure psychological demands, e.g. "My job requires that I work very fast". Scale reliability and validity was acceptable upon construction [37], as was the case in the current study (Cronbach's alpha: 0.68). The response scale ranged from 1 (strongly disagree) to 4 (strongly agree), and the scale score was calculated as the sum of the individual items (range 4-16)[37]. Higher scores indicate higher job demands, which is unfavorable.

Decision authority

The Dutch version of the Job Content Questionnaire (JCQ) was used to assess decision authority. The three item subscale comprised items such as "My job allows me to make many decisions myself". Scale reliability and validity was acceptable upon construction [37], in the current study it was good (Cronbach's alpha: 0.79). The response scale ranges from 1 (strongly disagree) to 4 (strongly

agree). The scale score was obtained by summing all the individual items (range 3-12). Higher scores indicate higher decision authority, which is positive.

Developmental possibilities

This concept was assessed with a subscale of the Dutch Well-being Checklist for Education [38], comprising four items, for example: “My work gives me the opportunity to learn new things”. The scale has shown good reliability (alpha 0.87) [39]. In the current study, internal consistency was acceptable (Cronbach’s alpha: 0.77). The response scale ranged from 1 (strongly disagree) to 5 (strongly agree), and the summed scale score ranged from 4-20. The higher the scale score, the more developmental possibilities were perceived, which is favorable.

Social support

The social support of colleagues, supervisor and management was measured using a modified version of two subscales of the Dutch version of the Job Content Questionnaire (JCQ) [37]. Each of these three subscales comprises three items, such as “My colleagues/my supervisor/the management help(s) to get the job done”. In the current study, internal consistency of the respective scales was excellent (Cronbach’s alpha: 0.99; 0.98; and 0.98). The response scales range from 1 (strongly disagree) to 4 (strongly agree). A scale score was obtained by summing the three individual items (range 3-12). Higher scores are indicative of more social support, which is positive.

Work ability

Work ability was measured using two of the seven dimensions of the Work Ability Index (WAI) [40]. Several studies have indicated that the first dimension, current work ability compared to lifetime best, could be used as an indicator of the status and progress of work ability [41, 42]. Reliability and validity of this scale have been shown to be adequate in a Dutch sample (Cronbach’s alpha 0.63 to 0.71) [43]. The scale comprises a question on perceived current work ability compared to lifetime best, measured on a frequency scale from 0 (unable to work) to 10 (very good). To additionally gain insight into work ability in relation to job demands, the second dimension of the WAI was added. This dimension comprises two questions on perceived work ability in relation to mental and physical job demands, recorded on a five-point frequency scale from 1 (very bad) to 5 (very good).

The combined scale score (range 2-20) was calculated as the sum of the score on current work ability and the weighted scores on the demands, according to the nature of the work. Higher scores indicate higher work ability, which is favorable. In the current study, internal consistency of the combined scale was good (Cronbach’s alpha: 0.76).

Job satisfaction

Two items of the Netherlands Working Conditions Survey 2010 [44] were measured to determine level of job satisfaction, namely: “to what extent are you, all things considered, satisfied with your job” and “[...], satisfied with your working conditions?” Response scales range from 1 (very dissatisfied) to 5 (very satisfied). The items were combined into one scale, showing an acceptable internal

consistency (Cronbach’s alpha: 0.70). The scale score was calculated as the mean of the two items (range 1-5), with higher scores indicating higher satisfaction.

Commitment to work and the organization

This concept was assessed using five items of the Dutch NOVA-WEBA questionnaire [45, 46], such as “My work means a lot to me” and “I feel perfectly at home in this organization”. Response scales range from 1 (strongly disagree) to 5 (strongly agree), with the scale score calculated as the mean of the score of all five items (range 1-5). Higher mean scores indicate higher commitment. Validity and reliability were moderate in an earlier report (Cronbach’s alpha 0.68) [47]. In the current study, internal consistency was acceptable (Cronbach’s alpha: 0.73).

Work engagement

Work engagement was assessed using the Utrecht Work Engagement Scale-9 (UWES-9), with the 3-item subscales vigor (see primary outcome vitality), dedication (e.g. “I am proud of the work that I do”), and absorption (e.g. “I am immersed in my work”) [35]. Response scales range from 0 (never) to 6 (always/every day). The scale and subscale scores were obtained by calculating the mean (range 0-6). Higher scores are indicative of higher work engagement. UWES-9 has shown good validity and reliability [36], as was the case in the current study (Cronbach’s alpha: 0.87).

Occupational self-efficacy

A modified version of the short Occupational Self-Efficacy Scale, comprising six items, was used to measure occupational self-efficacy (e.g. “Whatever happens in my work, I can usually handle it” [48]. Internal consistency was excellent in the current study (Cronbach’s alpha: 0.85), as was the case in other studies (Cronbach’s alpha 0.85) [49]. The response scale ranged from 1 (strongly disagree) to 5 (strongly agree) and a scale score was obtained by summing all individual items (range 6-30). A higher score indicates higher self-efficacy, which is favorable.

Organizational efficacy

This concept was assessed using the Organizational Efficacy Scale, comprising seven items, e.g.: “To what extent do you think your organization is able to deliver services of the highest quality?” [50]. The questionnaire was valid and reliable in previous studies (alpha 0.81) [50], internal consistency was excellent (Cronbach’s alpha: 0.89) in the current study. The response scale ranges from 1 (strongly disagree) to 5 (strongly agree). A total scale score was obtained by summing all individual items (range 7-35), so that a higher scores indicates higher organizational efficacy, which is favorable. Contrary to all other measures, organizational efficacy was measured at first and second follow-up only.

Covariates

Data on potential effect modifiers or confounders were collected at baseline, including age (in years), gender (male, female), school location (one of 12 locations), highest level of education (secondary school, vocational, professional or academic), function (teacher, teaching assistant, support staff, or management staff), and job tenure (in years).

Statistical analyses

All analyses were performed according to the intention to treat principle (i.e. the analyses are based on the initial treatment assignment), using IBM SPSS Statistics 22.

Baseline differences between the intervention and control group were checked by performing regression analyses for all outcomes and independent samples *t*-tests for all continuous variables and Pearson Chi-square tests for the dichotomous variable describing individual characteristics of the sample.

Selective attrition was checked by conducting loss to follow-up analyses. With independent samples *t*-tests, baseline scores of participants at first and/or second follow-up were compared to baseline scores of participants who did not fill out first and/or second follow-up (*p*-value <0.05).

To assess the effect of the intervention, linear mixed models with a two level structure was used, i.e. repeated measures were clustered within workers. Mixed models are especially suitable for longitudinal datasets containing correlated and unbalanced data [51, 52]. For each outcome variable, a crude model was built (i.e. difference between intervention and control group on average over time, corrected for the baseline value of the outcome [53]) as well as an adjusted model (i.e. the crude model, including adjustment for possible confounders age, gender, school location, and educational level). For organizational efficacy data were gathered only at first and second follow-up, hence linear regression analyses were conducted adjusting for the score on first follow-up measurement and for possible confounders (i.e. age, gender, school location, and educational level).

Two additional analyses were performed. First, time and the interaction between group and time were added to the adjusted mixed model in order to investigate whether the intervention effect was different over time (with a *p*-value <.05 indicating an interaction effect). And secondly, we compared high compliers in phase 1 (participation in two or three of the intervention's first phase elements) to the control group on the primary and secondary outcomes, while correcting for baseline values and covariates.

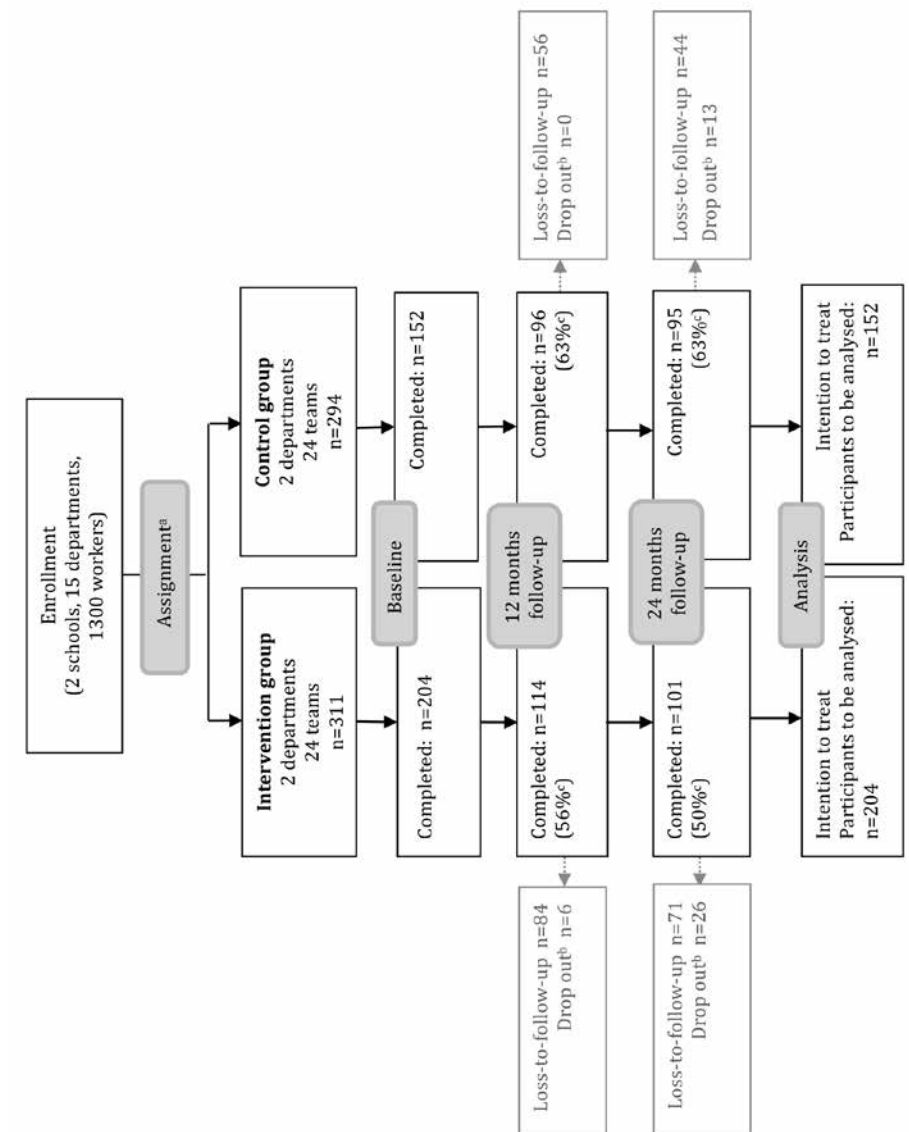
Results

Participant flow

The two schools were recruited in 2011. Figure 2 outlines the complete flow of participants through the study: of the 605 eligible workers from four departments, 356 (59%) completed the baseline questionnaire in February or June 2012. Between February 2013 and June 2014 the follow-up measurements were conducted. After 12 months, 210 participants completed the questionnaire (59%) and 6 participants dropped out due to termination of employment (Fig-

ure 2). After 24 months, 196 participants (55%) completed the questionnaire and 39 dropped out due to termination of employment (Figure 2). Following the intention to treat principle, the total number of employees to be analyzed is 204 for the intervention group and 152 for the control group (Figure 2). Loss to follow-up analyses did not show any selective attrition of participants.

Figure 2 — Flow diagram of the participants through the measurement moments of the trial



^a Assignment was based on matching criteria: department size, age composition, and type of work.

^b The reason for drop out (i.e. discontinuing intervention) was in all cases termination of employment.

^c Percentages are response percentages compared to baseline.

Baseline characteristics

The baseline characteristics of the study population are presented in Table 2. Most of the teams were represented in the baseline sample (20 out of 24 in the intervention group, 21 out of 24 in the control group). Both the intervention and control group consisted mainly of highly educated workers (85.8% and 77.0%, respectively) and teachers (78.4% and 65.1%, respectively). However, the intervention group comprised more women, was of higher age, and had more years of service in education.

Table 3 shows the means and standard deviations of baseline measurements. Significant differences existed between the intervention and control group on most of the outcomes (except for secondary outcomes work ability, absorption, social support colleagues and supervisor), in favor of the control group (Table 3).

Table 2 — Individual characteristics at baseline

	Total sample <i>N</i> =356	Intervention group <i>N</i> =204	Control group <i>N</i> =152	p-value ^a
Number of departments	4	2	2	-
Number of teams	41	20	21	-
Number of school locations	4	2	2	-
Gender (female) (%)	55.9%	65.2%	43.4%	.00*
Age (years) ^b [mean (SD ^c)]	50.7 (9.2)	52.5 (8.5)	48.7 (9.5)	.01**
Tenure (years) [mean (SD ^c)]	18.3 (11.5)	20.3 (11.4)	15.6 (11.2)	.00***
Educational level (%)				.09
Secondary school	6.2%	5.4%	7.2%	
Vocational	11.8%	8.8%	15.8%	
Professional or academic	82.0%	85.8%	77.0%	
Function (%)				.03
Teacher	72.8%	78.4%	65.1%	
Teaching assistant	7.6%	4.9%	11.2%	
Support staff	13.2%	10.8%	16.4%	
Management staff	6.5%	5.9%	7.2%	

^a Gender, education, and function tested with Chi-square test, age and tenure tested with an independent samples *t*-test.

^b Age based on *n*=182 due to missings on this voluntary question.

^c SD is standard deviation.

Table 3 — Means and standard deviations at baseline, and at 12-month and 24-month follow-up

	Intervention Group		Control group		p-value ^b
	<i>n</i>	mean (SD ^c)	<i>n</i>	mean (SD ^c)	
Primary outcomes					
Need for recovery (0-100)					
Baseline	204	41.7 (33.6)	152	31.5 (30.7)	0.00*
12 months	112	47.5 (32.4)	92	36.1 (31.4)	
24 months	101	45.2 (33.5)	94	43.0 (33.0)	
Vitality (0-6)					
Baseline	204	4.2 (1.3)	152	4.5 (1.1)	0.00*
12 months	113	4.0 (1.3)	92	4.5 (0.9)	
24 months	101	4.1 (1.2)	95	4.3 (1.0)	
Secondary outcomes					
Psychological demands (4-16)					
Baseline	204	14.3 (2.2)	152	13.6 (2.0)	0.00*
12 months	114	14.2 (2.0)	96	13.6 (1.9)	
24 months	101	14.3 (2.5)	95	14.3 (1.9)	
Decision authority (3-12)					
Baseline	204	8.4 (1.4)	152	8.9 (1.5)	0.00*
12 months	114	8.0 (1.4)	96	8.7 (1.5)	
24 months	101	8.3 (1.5)	95	8.8 (1.4)	
Developmental possibilities (4-20)					
Baseline	204	13.3 (2.7)	152	14.1 (2.9)	0.00*
12 months	114	13.4 (2.6)	96	14.1 (2.6)	
24 months	101	13.6 (2.9)	95	14.2 (2.5)	
Social support colleagues (3-12)					
Baseline	204	9.3 (1.1)	152	9.3 (1.1)	0.96
12 months	111	9.0 (1.0)	91	9.1 (0.9)	
24 months	101	9.0 (0.8)	94	9.3 (1.1)	
Social support supervisor (3-12)					
Baseline	204	8.2 (1.5)	152	8.1 (1.7)	0.99
12 months	111	7.7 (1.6)	91	7.9 (1.6)	
24 months	101	7.9 (1.7)	94	7.6 (1.9)	
Social support management (3-12)					
Baseline	204	7.2 (1.7)	152	7.6 (1.6)	0.02*
12 months	111	6.8 (1.6)	91	7.2 (1.8)	
24 months	101	6.8 (2.0)	94	7.3 (1.8)	

Work ability (2-20)						
Baseline	204	15.3 (2.7)	152	15.9 (2.0)	0.02*	
12 months	108	15.4 (2.3)	91	16.1 (2.1)		
24 months	99	15.3 (2.3)	91	15.4 (2.1)		
Job satisfaction (1-5)						
Baseline	204	3.3 (0.8)	152	3.7 (0.7)	0.00*	
12 months	107	3.5 (0.7)	90	3.8 (0.7)		
24 months	99	3.3 (0.8)	91	3.6 (0.7)		
Commitment (1-5)						
Baseline	204	3.6 (.5)	152	3.8 (0.5)	0.00*	
12 months	111	3.6 (.5)	90	3.8 (0.6)		
24 months	101	3.4 (.7)	94	3.8 (0.5)		
Work engagement (0-6)						
Baseline	204	4.0 (1.2)	152	4.3 (1.0)	0.00*	
12 months	113	3.9 (1.2)	92	4.4 (0.9)		
24 months	101	3.9 (1.2)	95	4.2 (1.0)		
Dedication						
Baseline	204	4.1 (1.3)	152	4.6 (1.1)	0.00*	
12 months	113	4.1 (1.3)	92	4.6 (0.9)		
24 months	101	4.1 (1.4)	95	4.5 (1.0)		
Absorption						
Baseline	204	3.7 (1.4)	152	3.9 (1.2)	0.00*	
12 months	113	3.7 (1.4)	92	4.0 (1.1)		
24 months	101	3.6 (1.4)	95	3.9 (1.1)		
Occupational self-efficacy (5-30)						
Baseline	204	23.5 (3.2)	152	23.9 (2.7)	0.02*	
12 months	113	22.5 (3.0)	92	22.8 (3.1)		
24 months	101	23.0 (3.4)	95	22.9 (2.9)		
Organizational efficacy ^b (7-35)						
12 months	111	19.8 (4.8)	91	22.1 (4.6)		
24 months	101	19.7 (4.8)	94	22.0 (4.9)		

^aSD is standard deviation.

^bAll variables are tested with a regression analysis corrected for school.

^cNot measured at baseline.

Effectiveness of the intervention

The intervention effects on primary and secondary outcomes are presented in Table 4. No significant intervention effects were found on the primary outcomes need for recovery (the difference between the groups on average over time $\beta=-3.2$; 95% CI -12.1 ; 5.7) and vitality ($\beta=0.1$; 95% CI -0.3 ; 0.4). For most of the secondary outcomes no intervention effect was found either, except for absorption (a subscale of work engagement) and organizational efficacy. For absorption, a significant intervention effect in unfavorable direction was found. The intervention group scored on average over time significantly lower on absorption than the control group ($\beta=-0.3$; 95% CI -0.6 ; -0.0). For organizational efficacy, a significant effect in unfavorable direction was found. The intervention group scored on average over time, significantly lower on organizational efficacy than the control group ($\beta=-2.2$; 95% CI -3.9 ; -0.5).

Significant interactions between group and time (i.e. effect of the intervention from baseline to T1) were observed on the primary outcomes need for recovery ($p=.036$) and vitality ($p=.018$) and the secondary outcomes social support of supervisor ($p=.048$) and work ability ($p=.013$). The interaction for need for recovery was negative ($\beta=-10.97$; 95% CI -21.91 ; -7.4), whereas positive interactions were found for vitality ($\beta=.44$; 95% CI .07 ; .81), social support supervisor ($\beta=.56$; 95% CI .01 ; 1.11) and workability ($\beta=1.12$; 95% CI .24 ; 2.00). This means that the 'effects' for need for recovery, vitality, social support of supervisor and work ability are stronger on T1 than on average over time. On the second additional analysis one effect was found: the high compliers scored on average over time significantly higher ($p=.00$) on occupational self-efficacy than the control group ($\beta=1.24$; 95% CI 0.06 ; 2.42).

Table 4 — Intervention effects on primary and secondary outcomes

	Crude model			Adjusted model ^b		
	Regression coefficient ^a	95% CI	p-value	Regression coefficient	95% CI	p-value ^c
Primary outcomes						
Need for recovery (0-100)	-0.486	-6.182 ; 5.209	0.867	-3.170	-12.067 ; 5.726	0.482
Vitality (0-6)	-0.010	-0.221 ; 0.200	0.922	0.059	-0.250 ; 0.368	0.707
Secondary outcomes						
Psychological demands (4-16)	0.016	-0.396 ; 0.428	0.939	-0.133	-0.668 ; 0.403	0.625
Decision authority (3-12)	-0.262	-0.544 ; 0.021	0.070	0.025	-0.387 ; 0.437	0.904
Developmental possibilities (6-30)	-0.432	-1.004 ; 0.141	0.139	-0.445	-1.339 ; 0.447	0.325
Social support colleagues (3-12)	-0.174	-0.365 ; 0.017	0.074	-0.156	-0.417 ; 0.103	0.236
Social support supervisor (3-12)	0.068	-0.278 ; 0.415	0.699	0.020	-0.484 ; 0.524	0.938
Social support management (3-12)	-0.259	-0.633 ; 0.115	0.174	-0.357	-0.834 ; 0.120	0.141
Work ability (1-10)	-0.173	-0.627 ; 0.280	0.452	0.134	-0.492 ; 0.761	0.672
Job satisfaction (1-5)	-0.124	-0.279 ; 0.030	0.115	-0.148	-0.366 ; 0.070	0.183
Commitment (1-5)	-0.151	-0.271 ; 0.032	0.013*	-0.163	-0.332 ; 0.006	0.058
Work engagement (0-6)	-0.037	-0.227 ; 0.154	0.706	-0.099	-0.360 ; 0.162	0.453
Dedication (0-6)	-0.055	-0.279 ; 0.169	0.629	-0.172	-0.471 ; 0.125	0.254
Absorption (0-6)	-0.132	-0.343 ; 0.078	0.216	-0.288	-0.576 ; -0.001	0.049*
Occupational self-efficacy (5-30)	0.149	-0.466 ; 0.763	0.634	0.065	-0.855 ; 0.985	0.889
Organizational efficacy ^c (7-35)	0.165	-1.055 ; 1.386	0.790	-2.21	-3.906 ; -0.507	0.012*

Note. The correlation of repeated measurements within the individual (the personal ID level) is taken into account in the mixed model analyses. The clustering effect of workplaces/teams is taken into account by correcting for school location, by adding three dummy variables to the model.

^a The regression coefficient indicates the difference between the intervention and the control group on average over time, corrected for baseline value of the particular outcome.

^b Adjusted for age, gender, school location, and education level. The correlation of repeated measurements within the individual (the personal ID level) is taken into account in the mixed model analyses.

^c Measured for the first time at T1; regression coefficient is an unstandardized B.

Discussion

The current study aimed to evaluate the long term effectiveness of an organizational level, primary preventive, participatory intervention on need for recovery and vitality. Contrary to the hypothesis, the results showed no effects of the intervention on the aforementioned primary outcomes. For most secondary outcomes no effects were found either. However, statistically significant effects on two of the secondary outcome measures were in unfavorable direction (i.e. absorption as a subscale of work engagement, and organizational efficacy).

At least four aspects of the current study could explain the lack of effect. Firstly, we measured a wide range of positive and negative outcomes, but all measures were collected at the individual level. One could argue that an organizational level intervention requires organizational level collection of data to detect an effect, such as sickness absence registrations [27], team performance indicators or company results. A second reason regarding the type of outcomes could be that we defined and operationalized the outcomes before the trial. However, the exact type, content and implementation of actions was developed during the intervention. Therefore, the relation between actions taken and measures was possibly too distant to detect an effect. Third, the process evaluation demonstrated that implementation of the intervention's first phase (needs assessment) was rather good, whereas the implementation of the actual changes in phase two (implementation phase) was poor in both schools [54]. Based on the level of implementation we expected to notice effects directly after intervention phase 1, at first follow-up, but these effects were only found for need for recovery, vitality, social support of supervisor and work ability. This finding should be interpreted with caution though, because it might as well be explained by a 'ceiling effect' in the high scores of the intervention group at baseline. For example, the baseline score of the intervention group for 'need for recovery' was 41.7. This is not only almost ten points higher than the control group, it is also higher than the mean score of around 30 points found in other studies (e.g. [9, 55, 56]). Such a high score at baseline makes an increase not likely. The fact that the improvement at T1 was not found for all outcomes might be due to the medium to low levels of satisfaction with the intervention. Hence, the lack of effect could be due to implementation failure. In post hoc analyses we tested for implementation failure and the effect found on occupational self-efficacy suggests that if the intervention would have been implemented as planned and the dose received would have been high enough for all, participants indeed might get a mastery experience out of taking part. Which in turn might lead to an increase in occupational self-efficacy. However, to reach this high dose received, the intervention's implementation strategy ought to be revised so to ensure participation throughout both phases of the intervention (e.g. by planning all intervention elements during working hours). Fourth, the lack of effectiveness could be due to theory failure, it could have been that the theory behind the intervention did not address the problem righteously. In future participatory intervention studies researchers could consider constructs that are 'closer' to the actual implementation process as outcomes (e.g. participation, readiness for change).

Comparison with earlier studies

Although the (partial) lack of effect was contrary to our expectations it is in line with some existing evidence on organizational-level interventions in education. A recent Cochrane review of organizational level interventions in (primary and secondary) education found only low-quality evidence that organizational interventions lead to improvements in teacher well-being and retention rates [57]. Low quality could for example be due to small numbers of participants or a lacking control group. However, the review included only four studies and in two cases teacher well-being was measured as a side effect of a student's intervention, limiting the generalizability of the review's outcomes.

The low or mediocre quality of evidence for organizational level interventions was also found in studies conducted outside of the educational domain. For example, the review by Montano and colleagues [58] included studies in health care, manufacturing and civil service mainly. The review demonstrated that comprehensive interventions, simultaneously addressing material, organizational, and work-time conditions, were more successful than single interventions. As a second example, an elaborate Cochrane review of stress management interventions of any type, conducted in health care, demonstrated that of the organizational-level interventions only changing work schedules may lead to a reduction of stress [59].

The current study adds to the existing body of evidence on the (partial) ineffectiveness of organizational level interventions for employee health. The evidence is considered to be relatively strong, since the design, with three measurements, was longitudinal as recommended in Michie and Williams' review [60] and followed participants for a longer period than in most studies [57]. Secondly, a complex intervention framework was used as recommended for this target group and these outcomes [57]. Thirdly, validated measures were used for the operationalization of the concepts. And lastly, the theoretical concepts focused both on positive and negative work-related aspects, hence health protective and health promotive effects could be detected.

Limitations of the current study

Some limitations of the current study need to be considered before generalizing the findings. Firstly, as a result of the long follow-up period of 12 and 24 months, loss to follow-up and drop out due to the termination of employment contracts were quite high. This probably affected the statistical power to detect changes. Secondly, although the matching was performed as effectively as possible, significant differences between the intervention and control groups persisted at baseline. This group difference was dealt with by correcting for baseline differences in all analyses [53]. A related, third limitation is the lack of randomization in this controlled trial: unknown confounding variables could be unevenly distributed over groups, threatening the internal validity. As has been described in the literature as a common challenge, the schools wanted to participate under the condition of choosing the intervention group [61]. Future studies of this type could consider alternative designs, such as the

stepped wedge approach for selecting the order of groups receiving treatment, or methods, such as propensity score matching, to overcome the possible bias resulting from non-randomization [61].

Recommendations for future research and practice

As described above, the current study already met some of the most important recommendations that were based on reviews of organizational level interventions and still no effects were found. There are at least three ways to further improve organizational level interventions. Montano and colleagues (2014) point to optimization of the implementation process as the strategy towards successful organizational-level interventions. The current study was conducted in daily practice and implementation suffered in this environment. The implementation is of utmost importance, since determinants of successful intervention [62] overlap with determinants of work-related stress (e.g. such as participation in decision making). By not implementing correctly, the facilitator or researcher could actually be adding a stressor to the work environment. The implementation strategy of this intervention should thus be revised (e.g. [63]) before the intervention can be recommended.

A second manner to improve effectiveness of organizational level interventions has been suggested by Ruotsalainen and colleagues: the interventions need to be more specific in their focus on stressors in order to be more effective [59]. In the current study the link between stressors formulated by all workers in phase 1 and actions taken by the management in phase 2 was unclear to most workers, which possibly hindered the effectiveness. This could be prevented by redesigning the implementation strategy in this intervention so as to incorporate participation as a central element in phase 2 as well, instead of leaving phase 2 to the management team in the schools.

Thirdly, organizational level interventions focusing on the root-cause of stress could be integrated with secondary preventive, individual focused stress management interventions. This simultaneous intervention on root-cause and early symptoms, could create a feeling of shared responsibility between organization and individual employee for occupational health. An integrated approach to workplace mental health might be more effective than one of the prevention types or levels alone [64].

Conclusions

To our knowledge, this is one of the first primary preventive, organizational level intervention studies targeted at all workers in education. Reviews have shown the potential of this type of intervention [26, 27], especially if participation and mastery experiences are incorporated in the intervention strategy [28, 29]. Until now intervention studies that aimed to improve teacher well-being were secondary preventive and targeted at the individual level mostly [20, 21]. Unique is the content of the intervention; we evaluated a practice-based intervention that had been applied and redesigned over a hundred times for differing organizations, according to the consultancy firm which developed the intervention. However, the results of this evaluation showed no effects of this type of intervention on the primary outcomes. Two small, statistically significant effects on secondary outcomes absorption and organizational efficacy appeared to be in unfavorable direction. Post-hoc analyses showed that high compliers with the first phase of the intervention, scored on average over time significantly higher on occupational self-efficacy than the control group. Suggesting that if the 'dose' is high enough (i.e. implementation is sufficient), the intervention might offer participants a mastery experience which affects occupational self-efficacy. The intervention program in its current form lacks a sufficient implementation strategy and is therefore not recommended. Organizational level participatory interventions for occupational health should incorporate an elaborate implementation strategy and be more specific in relating the actions taken to the stressors in the context. Future intervention studies aiming to improve occupational health should consider integrating organizational level, primary preventive elements with individual, secondary preventive elements, in order to be effective [64].

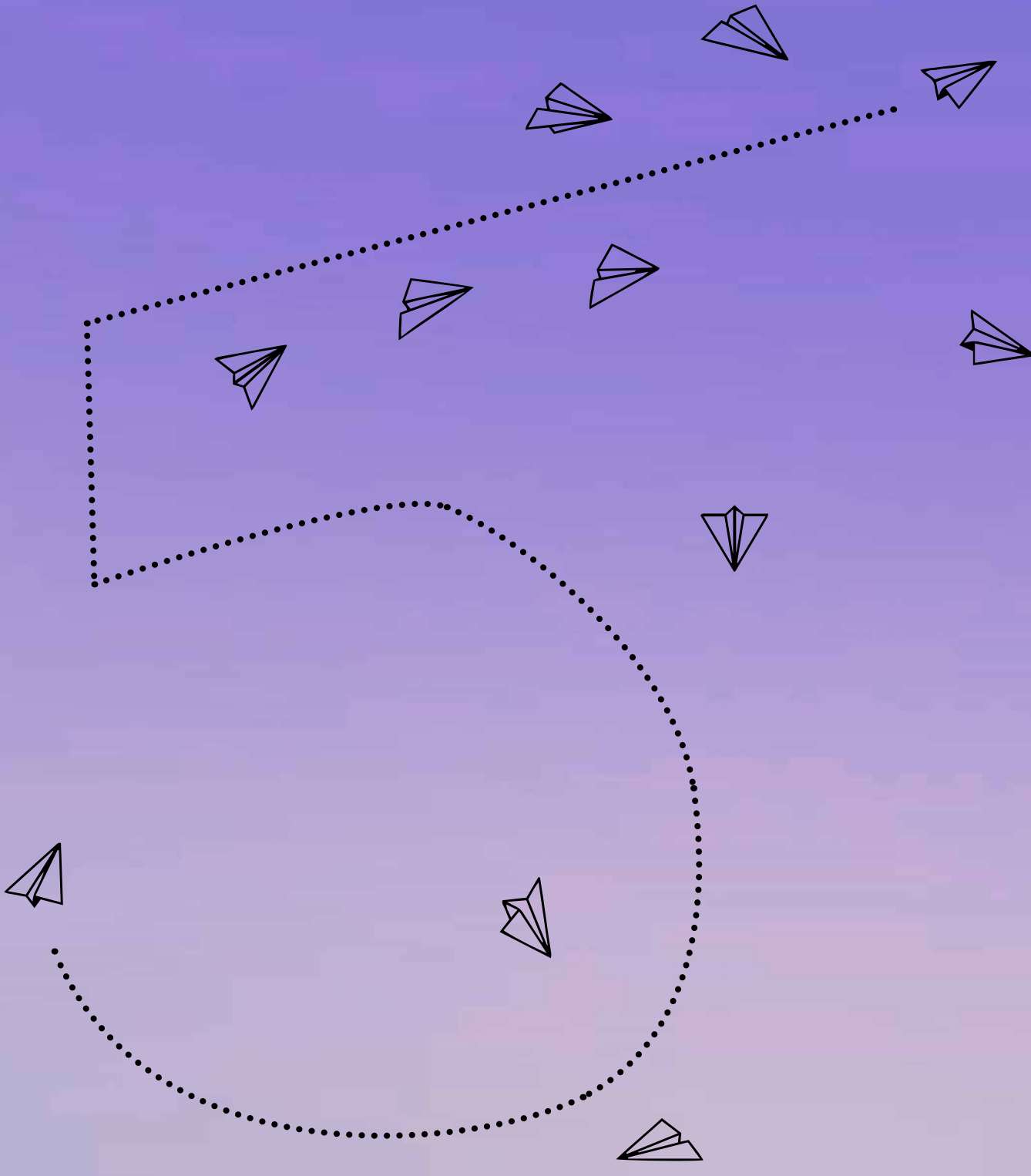
List of abbreviations

HM: Heuristic method (i.e. the intervention); JCQ: Job content questionnaire; NOVA-WEBA: Dutch questionnaire developed to identify risk factors for work stress; RCPE: Review committee participants in experiments; T0: Baseline measurement; T1: First follow-up measurement; T2: Second follow-up measurement; UWES-9: Utrecht work engagement scale-9; VET: Vocational Education and Training school; WAI: Work ability index.

References

- European Foundation for the Improvement of Living: *Fourth European working conditions survey*: European Foundation for the Improvement of Living and Working Conditions; 2007.
- Yang X, Wang L, Ge C, Hu B, Chi T: **Factors associated with occupational strain among Chinese teachers: a cross-sectional study**. *Public Health* 2011, **125**(2):106-113.
- Al-Mohannadi A, Capel S: **Stress in physical education teachers in Qatar**. *Social Psychology of Education* 2007, **10**(1):55-75.
- Klassen R, Wilson E, Siu AFY, Hannok W, Wong MW, Wongsri N, Sonthisap P, Pibulchol C, Buranachaitavee Y, Jansem A: **Preservice teachers' work stress, self-efficacy, and occupational commitment in four countries**. *European Journal of Psychology of Education* 2013, **28**(4):1289-1309.
- Smith A, Brice C, Collins A, McNamara R, Matthews V: **Scale of occupational stress: a further analysis of the impact of demographic factors and type of job**. 2000, *Contract Research Report 311/2000*.
- van Zwieten MHJ, de Vroome EMM, Mol MEM, Mars GMJ, Koppes LLJ, van denBossche SNJ: *Nationale Enquête Arbeidsomstandigheden 2013: Methodologie en globale resultaten*. [Netherlands Working Conditions Survey 2013: Methodology and overall results]. Hoofddorp: TNO; 2014.
- The Netherlands Association of VET Colleges (2012) VET in facts and figures: employees. <http://www.mboraad.nl/?page/1642332/Medewerkers.aspx>. Accessed October 15, 2015.
- Schaufeli WB, Bakker AB, Salanova M: **The measurement of work engagement with a short questionnaire: A cross-national study**. *Educational and Psychological Measurement* 2006, **66**(4):701-716.
- Van Veldhoven M, Broersen S: **Measurement quality and validity of the "need for recovery scale"**. *Occup Environ Med* 2003, **60**(SUPPL. 1):i3-i9.
- Demerouti E, Nachreiner F, Bakker AB, Schaufeli WB: **The job demands-resources model of burnout**. *J Appl Psychol* 2001, **86**(3):499-512.
- Easthope C, Easthope G: **Intensification, extension and complexity of teachers' workload**. *British Journal of Sociology of Education* 2000, **21**(1):43-58.
- Howe J: **Nine 'til Three? Not Likely! Teacher: The National Education Magazine** 2005:41-43.
- Fink D: **The law of unintended consequences: the 'real' cost of top-down reform**. *Journal of Educational Change* 2003, **4**(2):105-108.
- Karasek Jr RA: **Job demands, job decision latitude, and mental strain: Implications for job redesign**. *Adm Sci Q* 1979, :285-308.
- Demerouti E, Nachreiner F, Bakker AB, Schaufeli WB: **The job demands-resources model of burnout**. *J Appl Psychol* 2001, **86**(3):499-512.
- Maslach C: **Job Burnout: New Directions in Research and Intervention**. *Current Directions in Psychological Science* 2003, **12**(5):189-192.
- Hakanen JJ, Bakker AB, Schaufeli WB: **Burnout and work engagement among teachers**. *J School Psychol* 2006, **43**(6):495-513.
- Giga SI, Cooper CL, Faragher B: **The Development of a Framework for a Comprehensive Approach to Stress Management Interventions at Work**. *International Journal of Stress Management* 2003, **10**(4):280-296.
- Caulfield N, Chang D, Dollard MF, Elshaug C: **A review of occupational stress model of burnout**. *J Appl Psychol* 2004, **11**(2):149-166.
- Unterbrink T, Pfeifer R, Krippel L, Zimmermann L, Rose U, Joos A, Hartmann A, Wirsching M, Bauer J: **Burnout and effort-reward imbalance improvement for teachers by a manual-based group program**. *Int Arch Occup Environ Health* 2012, **85**(6):667-674.
- Gold E, Smith A, Hopper I, Herne D, Tansey G, Hulland C: **Mindfulness-based stress reduction (MBSR) for primary school teachers**. *J Child Fam Stud* 2010, **19**(2):184-189.
- Franco C, Mañas I, Cangas AJ, Moreno E, Gallego J: **Reducing teachers' psychological distress through a mindfulness training program**. *Spanish Journal of Psychology* 2010, **13**(2):655-666.
- Poulin PA, Mackenzie CS, Soloway G, Karayolas E:

- Mindfulness training as an evidenced-based approach to reducing stress and promoting well-being among human services professionals.** *International Journal of Health Promotion and Education* 2008, **46**(2):72-80.
24. Zolnierczyk-Zreda D: **An intervention to reduce work-related burnout in teachers.** *Int J Occup Saf Ergon* 2005, **11**(4):423-430.
25. Lamontagne AD, Keegel T, Louie AM, Ostry A, Landsbergis PA: **A systematic review of the job-stress intervention evaluation literature, 1990-2005.** *Int J Occup Environ Health* 2007, **13**(3):268-280.
26. Van der Klink JJJ, Blonk RWB, Schene AH, Van Dijk FJH: **The benefits of interventions for work-related stress.** *Am J Public Health* 2001, **91**(2):270-276.
27. Richardson KM, Rothstein HR: **Effects of Occupational Stress Management Intervention Programs: A Meta-Analysis.** *J Occup Health Psychol* 2008, **13**(1):69-93.
28. Nielsen K, Randall R, Holten A-, González ER: **Conducting organizational-level occupational health interventions: What works?** *Work Stress* 2010, **24**(3):234-259.
29. Bandura A: *Social foundations of thought and action : a social cognitive theory*: Englewood Cliffs, N.J.: Prentice-Hall; 1986.
30. Bandura A: **Cultivate self-efficacy for personal and organizational effectiveness.** In *Handbook of principles of organizational behavior : indispensable knowledge for evidence-based management*. Edited by Locke EA. Chichester, Sussex; [Hoboken]: John Wiley & Sons; 2009:120-136.
31. Bandura A: **Health promotion by social cognitive means.** *Health Education and Behavior* 2004, **31**(2):143-164.
32. Consiglio C, Borgogni L, Alessandri G, Schaufeli WB: **Does self-efficacy matter for burnout and sickness absenteeism? The mediating role of demands and resources at the individual and team levels.** *Work Stress* 2013, **27**(1):22-42.
33. Schelvis RMC, Oude Hengel KM, Wiezer NM, Blatter BM, Van Genabeek JAGM, Bohlmeijer ET, Van Der Beek AJ: **Design of the Bottom-up Innovation project - A participatory, Primary preventive, Organizational level intervention on work-related stress and well-being for workers in dutch vocational education.** *BMC Public Health* 2013, **13**(1).
34. Baarda, D.B., Goede, M.P.M.de: *Basisboek Methoden en technieken: handleiding voor het opzetten en uitvoeren van kwantitatief onderzoek*: 4th ed. Groningen [etc.]: Wolters-Noordhoff; 2006.
35. Schaufeli WB, Salanova M, González-Roma V, Bakker AB: **The measurement of engagement and burnout: a two sample confirmatory factor analytic approach.** *Journal of Happiness Studies* 2002, **3**:71-92.
36. Seppälä P, Mauno S, Feldt T, Hakanen J, Kinnunen U, Tolvanen A, Schaufeli W: **The construct validity of the Utrecht Work Engagement Scale: Multisample and longitudinal evidence.** *Journal of Happiness Studies* 2009, **10**(4):459-481.
37. Karasek R, Brisson C, Kawakami N, Houtman I, Bongers P, Amick B: **The Job Content Questionnaire (JCQ): an instrument for internationally comparative assessments of psychosocial job characteristics.** *J Occup Health Psychol* 1998, **3**(4):322-355.
38. TNO Kwaliteit van Leven: *Handleiding Welzijnscheck Onderwijspersoneel*: Heerlen: Stichting Vervangingsfonds en Bedrijfsgezondheidszorg voor het Onderwijs; 2007.
39. Kraan KO, Bakhuys Roozeboom MC, Wiezer NM: *Kwaliteit 'Welzijnscheck Onderwijspersoneel' en Validiteit Afkappunten*. Hoofddorp: TNO; 2010.
40. Tuomi K, Ilmarinen J, Jahkola A: *Work ability index: 2nd revised ed ed*. Helsinki: Finnish Institute of Occupational Health; 1998.
41. Ahlstrom L, Grimby-Ekman A, Hagberg M, Dellve L: **The work ability index and single-item question: Associations with sick leave, symptoms, and health - A prospective study of women on long-term sick leave.** *Scand J Work Environ Health* 2010, **36**(5):404-412.
42. El Fassi M, Bocquet V, Majery N, Lair ML, Couffignal S, Mairiaux P: **Work ability assessment in a worker population: Comparison and determinants of Work Ability Index and Work Ability score.** *BMC Public Health* 2013, **13**(1).
43. Radkiewicz P, Widerszal-Bazyl M: **Psychometric properties of Work Ability Index in the light of comparative survey study.** *International Congress Series* 2005, **1280**:304-309.
44. Koppes LLJ, De Vroome EMM, Mol MEM, Janssen BJM, van den Bossche SNJ: *Netherlands Working Conditions Survey 2010: Methodology and overall results*. Hoofddorp: TNO; 2011.
45. Dhondt, S., Houtman, I.L.D.,: *NIPG research questionnaire Job Content-WEBA (NOVA-WEBA): construction and a first test on reliability and validity*: Leiden: Nederlands Instituut voor Praeventieve Gezondheidszorg TNO; 1992.
46. Houtman ILD, Bloemhoff A, Dhondt S, Terwee T: *WEBA and NOVA-WEBA in relation to health and well-being of employees*: Leiden: Nederlands Instituut voor Praeventieve Gezondheidszorg TNO; 1994.
47. Kraan KO, Dhondt S, Houtman ILD, Nelemans R, De Vroome EMM: *Handleiding NOVA-WEBA. Een vragenlijst om arbeidsorganisatorische knelpunten op te sporen*. Hoofddorp: TNO Arbeid; 2000.
48. Schyns B, Von Gernot C: **A new occupational self-efficacy and its relation to personality constructs and organizational variables.** *European Journal of Work and Organizational Psychology* 2002, **11**(2):219-241.
49. Rigotti T, Schyns B, Mohr G: **A short version of the occupational self-efficacy scale: Structural and construct validity across five countries.** *Journal of Career Assessment* 2008, **16**(2):238-255.
50. Van Vuuren M, Elving WJ, De Jong MDT, Seysel ER: **What keeps you here? Organizational efficacy and person-organization fit as antecedents organizational commitment.** In *Why work? The contribution of value perceptions and efficacy expectations to organizational commitment*. Edited by Van Vuuren HA. 2006:121-147.
51. Twisk JWR: **Longitudinal data analysis. A comparison between generalized estimating equations and random coefficient analysis.** *Eur J Epidemiol* 2004, **19**(8):769-776.
52. Twisk JW: *Applied multilevel analysis: a practical guide*: Cambridge University Press; 2006.
53. Twisk J, Proper K: **Evaluation of the results of a randomized controlled trial: how to define changes between baseline and follow-up.** *J Clin Epidemiol* 2004, **57**(3):223-228.
54. Schelvis RM, Wiezer NM, Blatter BM, van Genabeek JA, Oude Hengel KM, Bohlmeijer ET, van der Beek, Allard J: **Evaluating the implementation process of a participatory organizational level occupational health intervention in schools.** *BMC Public Health* 2016, **16**(1):1212.
55. van Berkel J, Boot CR, Proper KI, Bongers PM, van der Beek, Allard J: **Effectiveness of a worksite mindfulness-related multi-component health promotion intervention on work engagement and mental health: results of a randomized controlled trial.** *PLoS one* 2014, **9**(1):e84118.
56. Coffeng JK, Boot CR, Duijts SF, Twisk JW, van Mechelen W, Hendriksen IJ: **Effectiveness of a worksite social & physical environment intervention on need for recovery, physical activity and relaxation; results of a randomized controlled trial.** *PLoS one* 2014, **9**(12):e114860.
57. Naghieh A, Montgomery P, Bonell CP, Thompson M, Aber JL, Milner AJ, Keegel T, Smith PM: **Workplace mental health: developing an integrated intervention approach.** *BMC Psychiatry* 2014, **14**(1):1.
58. Montano D, Hoven H, Siegrist J: **Effects of organisational-level interventions at work on employees' health: a systematic review.** *BMC Public Health* 2014, **14**:135-2458-14-135.
59. Ruotsalainen JH, Verbeek JH, Mariné A, Serra C: **Preventive occupational stress in healthcare workers.** *Cochrane Database Syst Rev* 2014, **Dec 8**(12).
60. Michie S, Williams S: **Reducing work related psychological ill health and sickness absence: a systematic literature review.** *Occup Environ Med* 2003, **60**(1):3-9.
61. Schelvis R, Oude Hengel K, Burdorf A, Blatter B, Strijk J, van der Beek A: **Evaluation of occupational health interventions using a randomized controlled trial: challenges and alternative research designs.** *Scand J Work Environ Health* 2015, **41**(5):491-503.
62. Durlak JA, DuPre EP: **Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation.** *Am J Community Psychol* 2008, **41**(3-4):327-350.
63. Wierenga D, Engbers LH, Van Empelen P, van Mechelen W: **A 7-Step Strategy for the Implementation of Worksite Lifestyle Interventions: Helpful or Not?** *J Occup Environ Med* 2016, **58**(5):e159-65.
64. LaMontagne AD, Martin A, Page KM, Reavley NJ, Noblet AJ, Milner AJ, Keegel T, Smith PM: **Workplace mental health: developing an integrated intervention approach.** *BMC Psychiatry* 2014, **14**(1):1.



Evaluating the implementation process of a participatory organizational level occupational health intervention in schools

Roosmarijn M.C. Schelvis
Noortje M. Wiezer
Birgitte M. Blatter
Joost A.G.M. van Genabeek
Karen M. Oude Hengel
Ernst T. Bohlmeijer
Allard J. van der Beek

Abstract

Background

The importance of process evaluations in examining how and why interventions are (un)successful is increasingly recognized. Process evaluations mainly studied the implementation process and the quality of the implementation (fidelity). However, in adopting this approach for participatory organizational level occupational health interventions, important aspects such as context and participants perceptions are missing. Our objective was to systematically describe the implementation process of a participatory organizational level occupational health intervention aimed at reducing work stress and increasing vitality in two schools by applying a framework that covers aspects of the intervention and its implementation as well as the context and participants perceptions.

Methods

A program theory was developed, describing the requirements for successful implementation. Each requirement was operationalized by making use of the framework, covering: initiation, communication, participation, fidelity, reach, communication, satisfaction, management support, targeting, delivery, exposure, culture, conditions, readiness for change, and perceptions. The requirements were assessed by quantitative and qualitative data, collected at 12 and 24 months after baseline in both schools (questionnaire and interviews) or continuously (logbooks).

Results

The intervention consisted of a needs assessment phase and a phase of implementing intervention activities. The needs assessment phase was implemented successfully in school A, but not in school B where *participation* and *readiness for change* were insufficient. In the second phase, several intervention activities were implemented at school A, whereas this was only partly the case in school B (*delivery*). In both schools, however, participants felt not involved in the choice of intervention activities (*targeting, participation, support*), resulting in a negative *perception* of and only partial *exposure* to the intervention activities. *Conditions, culture, and events* hindered the implementation of intervention activities in both schools.

Conclusions

The framework helped us to understand why the implementation process was not successful. It is therefore considered of added value for the evaluation of implementation processes in participatory organizational level interventions, foremost because of the context and mental models dimensions. However, less demanding methods for doing detailed process evaluations need to be developed. This can only be done if we know more about the most important process components and this study contributes to that knowledge base.

Background

Work-related stress is highly prevalent among teachers [1, 2] in different countries throughout both the eastern and western developed world [3]. In the Netherlands almost one in five teachers suffers from burnout complaints, compared to one in eight employees in the general working population [4]. Most of the interventions to prevent work-related stress in education aim to increase the resources of the individual to deal with the demands of the job [5-10]. However, these interventions were only partially effective in influencing (dimensions of) burnout [5-7, 9, 10] and well-being [10]. Explanations for this lack of effectiveness could be the level the intervention is aimed at (i.e. secondary or tertiary prevention), whereas it is proposed that problems should also be addressed at the source (primary prevention) and organizational level to sustainably decrease work-related stress. An effectiveness study of a primary preventive strategy for schools has demonstrated that this approach can indeed help to decrease burnout and to increase efficacy in teachers [11]. However, two meta-analyses of stress management interventions were thus far unable to demonstrate that primary interventions are more effective than secondary or tertiary interventions [12, 13]. This might be due to insufficient or partial implementation of primary, organizational level interventions, which might be explained by the lack of a proper implementation strategy [14-16].

Since it has been suggested that the implementation process can moderate or mediate the potential effects of complex organizational interventions on health or well-being [17], it seems important to study this process. The advantages of a process evaluation are that it helps the interpretation of outcomes [18, 19], sheds light on successes and failures of an intervention [17, 20] and thus shows what parts of the interventions should be improved in replication studies [17]. Lastly, it allows to draw inferences about future applicability in the current setting and about generalizability and transferability to other settings [21-23]. The importance of process evaluations in examining how and why interventions are (un)successful is increasingly recognized [24].

However, a recent review showed that process evaluations of stress management interventions are conducted in an explorative manner mostly, instead of using a theoretical framework [25]. Several models for the evaluation of implementation processes are available though [17, 26-29]. Previous public health intervention studies with comparable outcome measures as in the current study applied the Steckler and Linnan approach [26] to evaluate the implementation process [30, 31]. This process evaluation framework examines context, reach, dose delivered, dose received, fidelity, implementation, recruitment and satisfaction at the individual level. However, we also need to include particular aspects in our process evaluation that seem to be especially relevant for understanding implementation processes in participatory interventions in constantly changing organizations. These aspects are a close examination of the organizational context and participant's perceptions of the intervention. The first aspect, context, is often narrowly defined as the events that hindered

the implementation, whereas the broader organizational context, encompassing also the organization's culture and capacity to implement the intervention is often also of influence in this type of intervention [32]. The second aspect, the perception of the intervention, may be even more important than actual exposure to the intervention [33, 34], though few studies have actually measured exposure to primary organizational stress interventions, and linked the exposure patterns to outcomes. Studying exposure as well as the organizational context and participant's perceptions is possible using Nielsen and Randall's framework [17], which is developed specifically for organizational level occupational health interventions and thus best suited the intervention that is evaluated here. The framework can be applied to quantitatively and qualitatively assess three themes of process components: (i) intervention design and implementation; (ii) context; and (iii) participant's mental models. To our knowledge, this is one of the first studies in which the framework is applied to evaluate the quality of an implementation process.

The objective of this article is thus to systematically evaluate the implementation process of a primary preventive, participatory, organizational level intervention in two schools, by applying the Nielsen and Randall framework that addresses the intervention, the context and participants' mental models. The research question is: does the use of this process evaluation framework help us understand why or why not the implementation was successful?

Methods

The current process evaluation was performed alongside a controlled trial among employees in two secondary vocational education and training (VET) schools, investigating the effectiveness of an intervention on vitality and need for recovery. Detailed information on the methods, procedures and intervention can be found in the protocol article [36]. The project was conducted in two institutions for vocational education in the west (school A) and north (school B) of the Netherlands.

The study protocol and materials were approved by TNO's Review Committee Participants in Experiments, which is an internal ethics committee that assesses ethical aspects of involving participants in scientific experiments. All participants signed an informed consent before the first measurement.

Study population

The study population for this process evaluation consisted of teaching and non-teaching employees (i.e. educational and administrative support staff) and their managers in the intervention departments of both schools (school A, N=150; school B, N=161), including the senior management (i.e. Executive Board) and two intervention facilitators. The few employees within the inter-

vention departments teaching in general secondary education for adults only were excluded, because they were only administratively part of the intervention departments. In practice, they worked with and belonged to an interdepartmental group of teachers in secondary education for adults.

The intervention

The intervention under study in the 'Bottom-up Innovation project', the Heuristic Method (HM), is a participatory action approach applied at the organizational level. HM consists of two 12-month phases: (i) a phase of needs assessment, and (ii) an implementation phase.

In the first phase intervention activities to increase happy and healthy working are developed in conjunction with relevant stakeholders (i.e. staff and teachers) under supervision of an intervention facilitator, hereafter referred to as 'HM facilitator'. The HM facilitator is an expert in organizational change processes, and he or she uses the management's and employees' knowledge, skills and perceptions to thoroughly determine what factors hinder and facilitate "healthy and happy working" in the organization. A participatory group of employees (including a staff member) is formed, they assist the HM facilitator and serve as ambassadors of the project. All employees with an interest in the topic of health at work can apply to serve as a participatory group member, and they are then appointed on a first come first serve basis. Tasks for this group are executed within working hours and time spent is compensated. The HM facilitator, assisted by the participatory group, leads three steps to complete the first phase, the needs assessment, by: (i) approximately ten one-hour interviews with typical optimistic and typical critical teachers and staff selected by the participatory group; (ii) a digital questionnaire for all employees; and (iii) group sessions with all teachers within the intervention department, chaired by members of the participatory group. The result of each step in the intervention determines the content of the following step. Reports of each step are written by the HM facilitator. The participatory group approves each report before it is discussed with the management team. After discussion with the management team it is sent and presented to all employees in the intervention group by the HM facilitator. The third and last report, named "advisory report", is the HM facilitator's advice to the management team on how to proceed in the next phase.

In the second phase, the implementation phase, the intervention activities are implemented by the middle management team under supervision of the director and supported by senior management (hereafter referred to as 'the implementers'). HM prescribes that the implementers expand the HM facilitator's advisory report with an implementation plan, comprising at least a timeframe, a budget and an allocation of roles (e.g. the role of the participatory group), named "action plan". Assistance by the HM facilitator can be provided if the implementers have the means to temporarily hire such help.

Program theory

The program theory is our interpretation of how the intervention would work if implemented as planned. We assume that by involving an HM facilitator and by involving employees in participatory groups a thorough needs assessment can be conducted and based on that appropriate solutions for the improvement of the working environment can be developed and implemented. In particular we assume that the three steps of the needs assessment phase (i.e. interviews, questionnaire, group sessions) lead to identify which factors hinder and facilitate “healthy and happy working” in the organization. We further assume that the advisory report based on the comprehensive needs assessment will lead the management to develop an appropriate action plan that addresses the identified problems, describes how and in which time these should be solved, provides the necessary resources and conduct the implementation of these solutions. We also assume that the quick wins, which are part of the action plan, are implemented quickly. Based on these changes we assume that “healthy and happy working” in the organization will improve resulting in the end into improvements in vitality and need for recovery.

Requirements

In order to successfully implement this intervention, several requirements need to be met. These requirements are the conditions under which we assume the intervention to work. If the requirements are met, the chance of successful implementation of the intervention increases, and if implementation is successful the chance of finding the hypothesized health effects increases (i.e. increase in vitality, decrease in need for recovery). This latter part of the study is investigated in an effect evaluation and reported on in another article. All operationalizations of the requirements for successful implementation are described in Table 1 in chronological order of implementation, they are phrased as questions for comprehensibility. The table also states which data source was used to assess whether the requirement was met. Every requirement is assigned to one of the factors in the Nielsen and Randall framework.

Table 1 — Requirements for successful implementation, their operationalization and data source

Intervention phase 1: needs assessment	Operationalization			Source			
	Process component	Requirement based on program theory		Quan.	Logs	Qual.	
No.				Q1	Q2	I1	I2
1	Initiation ^a	Were senior and middle management committed to the intervention at the start?	<ul style="list-style-type: none"> Reasons for middle and senior management to participate 			X	X
2	Communication ^a	Was the intervention project communicated to the employees at the start?	<ul style="list-style-type: none"> The manner in which the project was communicated to the participants 			X	
3	Participation ^a	Was a participatory group formed? Did the employees feel involved in the intervention?	<ul style="list-style-type: none"> Composition of group and procedure was in line with protocol The majority of the participants scored above the cut-off point on the ‘employee involvement’ scale at T1 	X		X	X
4	Fidelity ^a	Was intervention phase 1 delivered by HM facilitator according to protocol?	<ul style="list-style-type: none"> The extent to which the HM facilitator complied with the needs assessment protocol, according to the facilitator and researcher 			X	
5	Reach ^a	Was intervention phase 1 received by majority of the employees?	<ul style="list-style-type: none"> Attendance of employees in each step of the needs assessment according to objective attendance rates. The rate expresses the number of those who actually participated in each step out of those who were asked to participate in each step. The majority of the participants scored above the cut-off point on the ‘exposure’ scale at T1 	X		X	
6	Communication ^a	Were results of each step in phase 1 communicated to employees by HM facilitator?	<ul style="list-style-type: none"> Percentage of participants who reported to have received a report of intervention step 1 in interviews Percentage of participants who reported to have received a report of intervention step 2 questionnaires Percentage of participants who reported to have received a report of intervention step 3 group sessions 	X			
7	Satisfaction ^a	Were the employees satisfied with intervention phase 1?	<ul style="list-style-type: none"> Satisfaction of all employees with (elements of the) needs assessment (i.e. interviews, questionnaire, group session, advisory report, overall) was moderate (≥6-7.4) or high (≥7.5)^a 	X		X	
8	Middle management support ^b	Was managerial support present at T1 according to management and employees?	<ul style="list-style-type: none"> The majority of the participants scored above the cut-off point on the ‘line manager attitudes and actions’ scale at T1 The managers demonstrated their support of the advisory report 	X		X	X

9	Readiness for change ^c	Was the majority of the employees at T1 ready for the change?	• The majority of the participants scored above the cut-off point on the 'readiness for change' scale at T1	Operationalization			X	Qual.		
				Q1	Q2	Logs		I1	I2	
Intervention phase 2: implementation										
Process component										
Requirement based on program theory										
10	Middle management support ^d	Was an action plan formulated by middle managers based on the advisory report? Were quick wins formulated?	• Middle managers visibly supported the project by designing an action plan including quick wins • The majority of the employees scored above the cut-off point on the 'line manager attitudes and actions' scale at T2	X						X
11	Participation ^d	Did employees participate in formulating an action plan?	• The extent to which employees felt responsible for the action plan (i.e. ownership) • The majority of the employees scored above the cut-off point on the 'employee involvement' scale at T2 • Formal representatives had a role • Middle managers encouraged active participation by employees	X						X
12	Targeting ^d	Did the action plan target the right problems in the workplace?	• The action plan was applicable to the problems of the workplace • Satisfaction with content action plan was moderate (6.0-7.4) or high (≥ 7.5)	X						X
13	Senior management support ^d	Did senior management support the action plan?	• Senior managers supported the project throughout • Senior managers allocated the necessary resources				X			X
14	Communication ^d	Was the action plan communicated to the employees? Were small successes celebrated?	• Employees were informed about the action plan and the progress towards its goals • Small successes were celebrated	X			X			
15	Delivery ^d	Was the action plan implemented by middle managers?	• Perceived implementation of the action plan, including quick wins, according to the implementers							X
16	Exposure ^d	Were the employees exposed to implementation of the action plan?	• Perceived implementation of the action plan, including quick wins, according to the employees • The majority of the participants scored above the cut-off point on the 'exposure to intended intervention' scale at T2	X						X
17	Culture ^b	Did the organizational culture facilitate the implementation of the action plan?	• Inherent characteristics of the organizational culture that facilitated or impeded the implementation of the action plan							X
18	Conditions ^b	Did the organization have the capacity to implement the action plan?	• The organizational characteristics that affected the (implementation) of the action plan • The organization had the capacity and skills to implement the action plan							X
19	Events ^b	Did no events interfere with the implementation of the action plan?	• Events that occurred and influenced the content or the execution of the action plan				X			X
20	Readiness for change ^c	Was the majority of the employees at T2 ready for the change?	• The majority of the participants scored above the cut-off point on the 'readiness for change' scale at T2 • The extent to which employees were ready for change and how this influenced the execution of the action plan	X						X
21	Satisfaction ^a	Were the employees satisfied with intervention phase 2?	• Satisfaction with intervention phase 2 was moderate (6.0-7.4) or high (≥ 7.5) ^d	X						X
22	Perception ^c	Did implementers and employees perceive the intervention as positive?	• The perception of the action plan was positive • Identify the common grounds and changes in the perception of the action plan							X

Note. Quan.: quantitative data, i.e. T1 and T2 questionnaires. Qual.: qualitative data, i.e. logs and T1 and T2 interviews. Q1: questionnaire at T1. Q2: questionnaire at T2. Logs: continuous records by logbooks. I1: interviews at T1. I2: group interviews at T2. ^a 'intervention' theme of process evaluation framework, ^b 'context' theme of process evaluation framework, ^c 'mental models' theme of process evaluation framework. ^d Satisfaction was rated on a 1-10 scale (1=very unsatisfied; 10=very satisfied) and the average satisfaction rate could be classified as poor (<6), moderate (≥ 6 and <7.5), or good (≥ 7.5).

The process evaluation framework

The framework by Nielsen and Randall [17] for the evaluation of organizational level occupational health interventions was applied. The framework describes three themes of process components that may influence intervention outcomes: intervention, context, and mental models. Table 1 lists all operationalizations, in the chronological order of implementing the intervention.

Intervention

The ‘intervention’ theme assesses the level of exposure to the intervention by describing (i) the intervention design and implementation, and (ii) the implementation strategy.

The first cluster, intervention design and implementation, was measured by the process components *initiation, targeting, reach, satisfaction and fidelity*. *Initiation* was operationalized as the sum of reasons for initiating the intervention (Table 1). This is considered to be an important aspect, since these reasons are likely to influence the initial commitment of all stakeholders [37]. *Targeting* was operationalized as the applicability (i.e. tailoring) of the intervention to the workplace’s problems (Table 1). This is of importance because every organization is different and therefore requires unique solutions [38] and intervention activities that do not target the right problems are unlikely to bring about the hypothesized effects. The component *reach* was operationalized as the attendance of employees in each step of the needs assessment according to objective and subjective measures (Table 1). Together with *satisfaction* with the intervention and accordance with the protocol (i.e. *fidelity*), *reach* demonstrates whether discrepancies exist between the planned intervention and its implementation [26]. This is of importance because an intervention activity cannot be effective if it is not implemented.

The second cluster of process components concerns the roles and behaviors of key stakeholders, also known as the implementation strategy. It comprises: *participation, support* of senior and middle management, and *communication*. The component *participation* was operationalized as the extent to which employees actually participated in decision-making (Table 1), and is widely recognized as a precondition for intervention success [39]. Failing to involve employees might lead to a lack of support for intervention activities, dissatisfaction and not targeting the right problems. *Support of senior and middle managers* was operationalized as their role throughout the project, including the allocation of necessary resources and possessing relevant skills (Table 1). Senior management support has a direct effect on the actual participation in the intervention [40]. Middle managers are often responsible for implementing the intervention and they are thus also in the position to obstruct or facilitate the change [41]. Finally, the component *communication* was operationalized as the type and quality of the communication about the intervention (Table 1). Successful communication is a way to commit employees to the project by keeping them informed [42] and enabling them to understand the intentions of the implementers [43].

Context

The ‘intervention context’ theme comprises the organization’s *culture, conditions* (i.e. the omnibus context; [44]) and *events* (i.e. the discrete context; [44]) and supposedly moderates or mediates the link between exposure to an intervention and the outcomes. Measuring these three context elements is of importance because they may either facilitate or hinder the implementation of an intervention [17].

The component *culture* was operationalized as those inherent characteristics of a group that facilitate or impede implementation (Table 1). *Conditions* are defined as the capacity of the organization to implement the actions (Table 1). *Events* are occurrences that influenced the content or execution of the actions (Table 1).

Participant’s mental models

The theme ‘participant’s mental models of the job and intervention’, comprises *readiness for change* and (changes in) *perceptions*. This theme concerns all appraisals and perceptions of key stakeholders and how these may drive their behaviors. Comparable to the context, mental models supposedly moderate or mediate the link between exposure to an intervention and the outcomes [17].

Readiness for change was operationalized as the extent to which participants are ready for the change the intervention implies (Table 1). It has been argued that interventions can only be effective if participants perceive that problems are present, should be solved, and could be solved by the intervention [42].

It was assessed whether *perceptions* of the intervention differed among groups of participants (Table 1). In an intervention context, participants may develop similar ‘models’ to interpret and respond to their work context, including the intervention [45]. If participant’s mental models are not similar this might obstruct implementation, for example because individual have different agendas or see the implemented changes differently.

Procedure of applying the framework

Nielsen and Randall [17] proposed a set of questions to help the operationalization of process components into measurable constructs. In three one-hour consensus meetings, four authors (RS, JvG, NW, KOH) adjusted every question to the current study and reached consensus on the data collection tool, the timing of measurement, and the implementer perspective (i.e. middle and senior management,) or participant perspective (i.e. participatory group, teaching and non-teaching staff; Table 1). Every process component in the framework is assigned to one or more requirements for successful implementation, as is described in the program theory section.

Data Collection

Data were collected by means of questionnaires, interviews and a logbook, from the implementer perspective (i.e. middle and senior management) or participant perspective (i.e. participatory group, teaching and non-teaching staff). Process questionnaires, comprising questions on the process components listed in Table 1, were sent out to all participants at 12 months after the start of the study (T1, i.e. in between intervention phase 1 and 2), and at 24 months after the start of the study (T2). The start of the study is defined by the time the baseline measurements for the outcome evaluations were conducted. Interviews were conducted at T1 and T2. The researcher's logbook was kept up to date throughout the duration of the study.

Questionnaire (measures)

The questionnaires were sent out digitally to all participants and were to be filled out within six weeks. Strategies to increase the response rate were: (i) a maximum of three reminders, (ii) the provision of an incentive (i.e. a book voucher) for those who filled out the complete questionnaire, and (iii) the option to fill out the questionnaire either by telephone communication (school A) or in hard copy (school B), as suggested by the respective participatory groups.

Measures in the questionnaires were a combination of a validated measure and tailor-made, explorative measures. The validated measure is the Intervention Process Measure [46] of which four scales were used: (a) line manager attitudes and actions (e.g. "My immediate manager has done a lot to involve employees throughout the process", 7 items), (b) exposure to components of the intended intervention (e.g. "In this project we openly discuss which traditions or procedures we wish to change and which we wish to keep", 5 items), (c) employee involvement (e.g. "I had the opportunity to give my views about the change before it was implemented", 3 items), and (d) employee readiness for change (e.g. "I look forward to the changes brought about by the intervention project", 4 items). Answers were given on a five-point Likert scale from 1 = totally disagree to 5 = totally agree. The scales were all reliable, respective Cronbach's alphas were: 0.89, 0.80, 0.81, 0.79. In the absence of a clear external criterion, cut-off scores were established by a rule, that is: not more than one item missing per scale and at least two thirds of the items (rounded upwards) has a minimum score of 4. The sum of scores is then divided by the number of items and the cut-off point is determined as the score greater than and not equal to the mean. Respective cut-off points were: (a) 3.14, (b) 2.8, (c) 3, and (d) 3.25. Results are reported as the percentage of participants that scored above the cut-off point. The tailor-made measures were a combination of descriptive yes/no-questions (e.g. "Are the quick wins implemented?") and 1 to 10 ratings with higher ratings indicating higher satisfaction (e.g. "Can you assess the content of the action plan on a scale of 1-10 (where 1 is poor and 10 is excellent)?"). For every question there was an option to elaborate on the response given. Both the validated and explorative measures were tested for comprehensibility by five teachers and one manager at T1 and T2.

Interviews

At T1, 17 individual interviews were held (school A, N=11; school B, N=6). At T2, 16 group interviews were held: in school A, 22 participants took part in in total 8 interviews, and in school B, 25 participants took part in in total 8 interviews. All stakeholders were interviewed at T1 and T2: teachers and non-teaching staff, the participatory group, middle and senior management, the two HM facilitators. At T1 interviewees at the participant level were selected based on their high attendance (participation in three or two steps) or low attendance in the needs assessment (participation in one or none steps), to grasp both the perspective of highly exposed participants and low exposed participants.

Interviews were held face-to-face. Seven interviews (six at T1 for school B, one at T2 for school A) were conducted by telephone because of time constraints. All interviews were scheduled for an hour, except for the interview with the participatory group (1.5 hour) and senior management (0.5 hour).

Logbook

In the logbook, the sequence of planned and unplanned events was listed alongside impressions of the principal researcher. In order to do so, the principal researcher was present as an observer at the meetings of the participatory groups in phase 1. For phase 2 the principal researcher based impressions on reading progress reports by the HM facilitator (school A) or holding periodical (telephone) interviews with the school principal (school B).

Data Analysis

Three sources of mixed methods (i.e. questionnaires, interviews, logbooks) were collected over the course of 24 months to study the implementation components from multiple perspectives, also referred to as methodological triangulation. By comparing and contrasting the perspectives on the implementation components it is possible to arrive at a deeper, wider and more valid understanding of each component than by using only one data source [47]. Two levels of analysis were identified: the implementer perspective and the employee perspective. The implementer perspective comprised two units of analysis (i.e. middle and senior management) and the employee perspective comprised two units of analysis (i.e. the participatory group, and teaching and non-teaching staff). Analyses of the implementation components were first conducted within each level and then between levels to identify similarities and discrepancies between the perspectives. Since most elements of the model were covered by the T1 and T2 interviews, we started our analysis with these sources.

Questionnaires were analyzed using the statistical software package SPSS 22 for the closed-ended questions, using descriptive statistics (i.e. mean, range).

Participants had to have filled out T0 to be included in the analyses. Qualitative, open-ended questions were coded manually in the same manner as the interviews.

All interviews were audiotaped and the tapes were transcribed. The first (RS) and second author (NW) separately analyzed the first four transcripts for each school (i.e. transcripts for interviews with team, director, management, and HM facilitator). During a consensus meeting the separate analyses were compared and common themes were identified, in line with the principles of thematic content analysis [48]. This quality procedure served two purposes: i) to ensure consistent and robust coding following the process evaluation framework, and ii) to ensure that every emerging theme was directly supported by data from the interviews or monitoring. After the researchers had agreed on a classification of themes, further analyses were conducted by the first author (RS). Textual segments were marked with codes indicating the process component it was related. The extracted segments were digitally tracked in Microsoft Excel.

The digital and hard copy logbook data were grouped per school to form a chronological list of events, including the impressions of the principal researcher (RS).

Results

The evaluation of each requirement for successful implementation is described below for school A and B. We distinguish the management perspective (senior and middle management) and the participant perspective (teachers and non-teaching staff, and the participatory group).

Intervention phase 1: needs assessment

1. Were senior and middle management committed to the intervention at the start? (initiation)

Senior and middle management in both schools decided to participate in the study to solve a given problem in a specific department. In school A this was high sickness absence rates combined with an ageing work population. In school B this was signs of diminishing happiness at work and an ageing work population. Since the problems were meaningful to the management, we consider the commitment at the start to be high (Table 2).

Table 2 — Summary of quantitative and qualitative results per process component for school A and B

Process component	Requirement	School A		School B		
		Quan.	Qual.	Quan.	Qual.	
<i>Intervention phase 1: needs assessment</i>						
1	Initiation	Were senior and middle management committed to the intervention at the start?	-	Yes	-	Yes
2	Communication	Was the intervention project communicated to the employees?	-	Yes	-	Yes
3	Participation	Was a participatory group formed?	-	Yes	-	Partly
		Did the employees feel involved in the intervention?	Yes	-	No	-
4	Fidelity	Was intervention phase 1 delivered by HM-facilitator according to protocol?	-	Yes	-	Yes
5	Reach	Was intervention phase 1 received by majority of the employees?	Yes	Partly	No	Partly
6	Communication	Were results of each step in phase 1 communicated to employees by HM-facilitator?	Yes	Yes	Yes	Yes
7	Satisfaction	Were the employees satisfied with intervention phase 1?	Yes	Partly	Yes	Partly
8	Managerial support	Was managerial support present at T1 according to management and employees?	Yes	Yes	Yes	No
9	Readiness for change	Was the majority of the employees at T1 ready for the change?	Yes	-	No	-
<i>Intervention phase 2: implementation</i>						
10	Middle management support	Was an action plan formulated by middle managers based on the advisory report?	Yes	Yes	Yes	No
		Were quick wins formulated?	Yes	Yes	No	Partly
11	Participation	Did the employees participate in formulating an action plan?	Yes	No	Partly	No
12	Targeting	Did the action plan target the right problems in the workplace?	Yes	Partly	No	No
13	Senior management support	Did senior management support the action plan?	-	Partly	-	Partly
14	Communication	Was the action plan communicated to the employees?	Yes	Yes	No	Partly
15	Delivery	Was the action plan implemented by middle managers?	-	Yes	-	Partly
16	Exposure	Were the employees exposed to implementation of the action plan?	Partly	Partly	Partly	Partly
17	Culture	Did the organizational culture facilitate the implementation of the action plan?	-	No	-	No
18	Conditions	Did the organization have the capacity to implement the action plan?	-	Partly	-	No
19	Events	Did no events interfere with the implementation of the action plan?	-	No	-	No
20	Readiness for change	Was the majority of the employees at T2 ready for the change?	No	No	Yes	No
21	Satisfaction	Were the employees satisfied with intervention phase 2?	No	-	No	-
22	Perception	Did implementers and employees perceive the intervention as positive?	-	Partly	-	No

Note. Quan: quantitative data, i.e. questionnaire at T1 and T2 ; Qual: qualitative data, i.e. continuous records by logs, interviews at T1 and group interviews at T2. (-): this aspect was not assessed quantitatively/qualitatively.

2. Was the intervention project communicated to the employees at the start? (communication)

In both schools, the HM facilitator and the participatory group organized a kick-off meeting for all involved, wherein the intervention and accompanying research was explained. The majority of all involved were present at the kick-off, according to logs. All had the possibility to ask questions, and digital slides with information on the project were emailed to all employees of the intervention group afterwards. We conclude that communication about the intervention project was successful at both schools at the start (Table 2).

3. Was a participatory group formed? And did the employees feel involved in the intervention? (participation)

A participatory group of employees and staff was formed consisting of six and eleven members for School A and school B, respectively. Qualitative data showed a deviation from the intervention protocol in the composition of the participatory group: in school B it comprised five more members than foreseen, since the management wanted all teams to be represented. This deviation hindered the intervention process since the participatory group at school B was too large to function effectively. Quantitative data showed that in school A, a majority (71.9%) felt involved at this point in the intervention process, while only about one third of employees in school B (34.4%) felt involved (Table 3). We conclude that participation was sufficient in school A at this point in time, but not in school B (Table 2).

4. Was intervention phase 1 delivered by HM facilitator according to protocol? (fidelity)

The three major intervention steps (interviews, questionnaire, group sessions) were executed as planned in both schools. Two changes in the execution of these steps were noted in both schools which actually improved the tailoring of the intervention. First, the composition of the step 2 questionnaire: questions were tailored to the specific problems in collaboration with the participatory group, instead of using standardized, validated modules. Second, with regard to the chairing of the group sessions: the participatory groups were supposed to chair the sessions, but no guideline for chairing the meetings was available. Therefore a guideline for chairing the meetings was drafted by the HM-facilitator and researcher, to ensure that information was given to all participants and in the same way (e.g. information on the aim and duration of the session, anonymity of the data gathered). The researcher observed 11 sessions and concluded that the guideline was used as planned. These deviations tailored the intervention to the intervention context and thus facilitated the intervention process. We conclude that the first phase of the intervention was delivered as planned in the protocol at both schools (Table 2).

5. Was intervention phase 1 received by the majority of the employees? (reach)

The majority in school A and B did participate in interviews and group sessions, but not in the questionnaire (Table 3). More than the majority scored above the cut-off point on the 'exposure' scale in school A (68.8%), while this was the

case for less than a third for employees in school B (29.7%; Table 3). We conclude that reach in this first phase of the intervention (needs assessment) was only partly successful (Table 2).

Table 3 — Summary of quantitative scores for intervention phase 1 (needs assessment) per school

	School A	School B
3. Participation		
% of employees who feel involved in the intervention	71.9% (n=23/32)	34.4% (n=22/64)
5. Reach (% yes)		
Participation in interviews ^a	91.7% (n=11/12)	100% (n=12/12)
Participation in questionnaire ^a	39.3% (n=59/150)	47.8% (n=77/161)
Participation in group sessions ^b	73.3% (n=110/150)	54.0% (n=71/161)
% of participants who feel exposed to the intervention	68.8% (n=22/32)	29.7% (n=28/64)
6. Communication (% yes)		
Received report on interviews?	53.1% (n=17/32)	57.1% (n=40/70)
Received report on questionnaire?	53.1% (n=17/32)	68.6% (n=48/70)
Received advisory report?	93.8% (n=30/32)	65.7% (n=46/70)
7. Satisfaction^c (SD)^d		
Overall	6.5 (1.19) (n=32)	5.9 (1.58) (n=63)
Interviews	7.5 (.57) (n=4)	8.0 (0.89) (n=11)
Questionnaire	7.1 (1.14) (n=26)	7.3 (0.99) (n=55)
Group sessions	6.7 (1.13) (n=26)	6.8 (1.58) (n=38)
Advisory report correct summary of opinions/wishes/needs? (%yes)	64.5% (n=20/31)	65.4% (n=34/52)
8. Middle management support		
% employees who feel supported	68.8% (n=22/32)	53.1% (n=34/64)
9. Readiness for change		
% employees who feel ready for the change	81.3% (n=26/32)	54.7% (n=35/64)

Note. Variables are whenever possible denoted as percentages (cases/n). The n differs per variable due to the operationalization of the variable (e.g. satisfaction with report interviews only asked if participant reported to have received the report) or due to drop out during the process of filling out the questionnaire. ^aParticipation rates in interviews and questionnaire are based on logbook notes. ^bParticipation rate in group sessions is based on objective attendance lists. ^cSatisfaction was rated on a 1-10 scale (1=very unsatisfied; 10=very satisfied) and the average satisfaction rate was classified as poor (<6), moderate (≥6 and <7.5), or high (≥7.5). ^dSD is standard deviation.

6. Were results of each step in phase 1 communicated to employees by HM facilitator? (communication)

The majority of participants did recall to have received the output of each step in the needs assessment, with highest receiving scores in school A for the advisory report (93.8%) and the report on the group sessions in school B (68.6%; Table 3). Communication halfway through the project was thus considered successful (Table 2).

7. Were the employees satisfied with intervention phase 1? (satisfaction)

All satisfaction scores are shown in Table 3. In school A, the participants were overall moderately satisfied with the implementation of the intervention's first phase (mean 6.5), whereas this was not the case in school B (mean 5.9).

Satisfaction scores for interviews were however high in both schools (7.5 in school A, 8.0 in school B). Qualitative analyses showed that in the interviews, the profound questioning by the facilitator was valued. Participants in both schools were moderately satisfied with the questionnaire (7.1 in school A, 7.3 in school B). Qualitative analysis of open ended questions in the questionnaire and T1 interviews showed that it was especially valued that 'the right themes' were addressed. Again in both schools, participants were moderately satisfied with the group sessions (6.7 in school A, 6.8 in school B). More specifically, most of the participants felt that everyone could freely give his or her opinion (88.9%) and that the chair listened to them (98.6%; results not shown in table). The majority reported a feeling of taking part in potential changes by participating in the group sessions (79.2%) and felt responsible for the outcome of the group session (88.9%; results not shown in table). However, qualitative analysis of open ended questions in the questionnaire showed that some regretted that no solutions were found to the identified problems right away.

In both schools, a majority of participants (64.5% in school A, 65.4% in school B) perceived the HM facilitator's advisory report based on interviews, questionnaire and group sessions as a correct summary of their opinions, needs, and wishes (Table 4 describes the recommendations of the advisory report per school and translation into action plan). However, almost one in three participants who had received the report stated 'I do not know/No opinion' to this question at the T1 questionnaire. This could be considered either a sign of dissatisfaction with the report, or a sign of failed 'reach'.

We conclude that although satisfaction scores for interviews, questionnaire and group sessions are moderate to high, there are some signs of dissatisfaction (with regard to evaluating the advisory report in both schools and overall satisfaction in school B) which might have been a hindrance factor for some participants at this point in the implementation process (Table 2).

Table 4 — Results of the needs assessment and translation into action plan

	Main content of advisory report delivered by facilitator	Main content of action plan ^a constructed by management team
School A	(i) professionalize the teams; (ii) professionalize the management; (iii) improve the administrative support and facilities.	The director, assisted by an HM consultant, translated the recommendations into an action plan with three goals, six changes and a set of quick wins. Goals: i) unambiguous management control; ii) competence and professionalism in the teams, and iii) adequate facilities Changes: (i) compliance to the workload policy, (ii) structured performance reviews; (iii) a continuous dialogue on the organization of the educational programs; (iv) a leading team activities plan; (v) weekly work meetings; and (vi) personalized competence development plans. Quick wins: create adequate facilities by creating a staff room at both locations; place extra walls in some classrooms; place beamers in all class rooms; improve the service by the facilitation services office.
School B	(i) create adequate and effective management control by installing a management team that is approachable, coaching, and leading; (ii) make teams the central executive units by developing a team program; (iii) eliminate cumbersome administrative procedures.	The directors of the management team decided to integrate the facilitator's recommendations in the annual agreements (i.e. a management contract) she made with the Executive Board, instead of writing a separate action plan. A coach was attracted to support teams in a previously initiated change towards becoming self-managing. Goals were formulated in four headlines: i) strategy; ii) education; iii) personnel; iv) organization; and v) business operations. The most important change per headline was: i) alliances with partners in the region are closed; ii) the curriculum of two educations are reconstructed into units of learning; iii) performance review policies are implemented; iv) teams function as self-managing units; and v) a multi-annual housing plan is developed. No quick wins were formulated.

^a Action plan was termed 'Management Contract' in school B. ^b Heading 12 of the results section (targeting) describes that middle managers (other than the two directors) and participants did not see how the advisory report was translated into the Management Contract.

8. Was managerial support present at T1 according to management and employees? (middle management support)

In both schools, the majority of the participants perceived managerial support with regard to the intervention (68.8% in school A, 53.1% in school B) as was demonstrated by quantitative analyses (Table 3).

In school A, the managers were shocked by the problem definition in the advisory report, leading to a state of apathy which slowed down initial actions. As will be outlined below, support was manifested as soon as the shock was descended.

The directors in school B were not satisfied with the advisory report, which they considered to be too focused on the leading role of the management, whereas the school tried to make the transition towards self-managing teams.

In sum, managerial support was present in school A at T1 and only partly present in school B (Table 2).

9. Was the majority of the employees at T1 ready for the change? (readiness for change)

Quantitative analyses showed that the majority of the participants in both schools was indeed ready for the change (Table 2). Scores were higher in school A (81.3%) than in school B (54.7%; Table 3).

Intervention phase 2: implementation

10. Was an action plan formulated by middle managers based on the advisory report? And were quick wins formulated? (middle management support)

With the help of an HM facilitator, the management team of school A developed an action plan that included quick wins (Table 4). In school B, the directors of the management team decided to integrate the recommendations in the advisory report with the annual agreements made with the Executive Board, instead of writing a separate action plan. The annual agreements were named 'Management Contract' (Table 4) and did not comprise a timeframe, a budget and an allocation of roles. The directors did identify several quick wins, however these were not stated in the Management Contract.

A majority of employees perceived that managerial support for the action plan or its equivalent was present at T2 (57.1% in school A, 51.1% in school B, Table 5).

In conclusion: in school A an action plan including quick wins was formulated, whereas no action plan and no quick wins were formulated in school B. In both schools, the majority of employees perceived that managerial support for the action plan was present (Table 2).

Table 5 — Summary of quantitative scores for intervention phase 2 (implementation phase) per school

	School A	School B
10. Middle management support (% employees that perceives managerial support for action plan ^a)	57.1% (n=32/56)	51.1% (n=23/45)
11. Participation (% yes)		
I feel responsible for the implementation of the action plan ^a	73.2% (n=41/56)	76.1% (n=35/46)
I feel responsible for the content of the action plan ^a	71.4% (n=40/56)	76.1% (n=35/46)
<i>Employee involvement</i> (% employees who feel involved in action plan)	55.4% (n=31/56)	37.0% (n=17/46)
12. Targeting (SD)^b		
Satisfaction with content of the action plan ^a	6.5 (1.31) (n=61)	5.4 (1.96) (n=49)
14. Communication (% yes)		
<i>Are you informed about the progress in the action plan^a?</i>	86.7% (n=52/60)	38.3% (n=18/47)
<i>How were you informed (more answers possible)?</i>		
* Via information meetings organized by the management team	75% (n=39/52)	61.1% (n=11/18)
* Via work meetings	44.2% (n=23/52)	44.4% (n=8/18)
* Via the news letter	40.4% (n=21/52)	38.9% (n=7/18)
* In another way	7.7% (n=4/52)	16.7% (n=3/18)
16. Exposure (% yes)		
<i>Have you been informed on existence of the action plan^a?</i>	100% (n=61/61)	87.8% (n=43/49)
<i>Are you familiar with the goals of the action plan^a?</i>		
* unambiguous management control teams	83.6% (n=51/61)	-
* competence and professionalism in the teams	90.2% (n=55/61)	-
* adequate facilities	90.2% (n=55/61)	-
* make teams the central executive units	-	89.8% (n=44/49)
* adequate and effective management control	-	67.3% (n=33/49)
<i>Have you seen changes towards these goals:</i>		
* unambiguous management control teams	43.3% (n=26/60)	-
* competence and professionalism in the teams	43.3% (n=26/60)	-
* adequate facilities (quick wins)	83.6% (n=51/61)	-
* make teams the central executive units	-	59.6% (n=28/47)
* adequate and effective management control	-	25.5% (n=12/47)
<i>Are the quick wins^c celebrated as a success?</i>	67.2% (n=41/61)	-
<i>Did you notice change in:</i>		
* a dialogue on the organization of education	78.3% (n=47/60)	-
* performance reviews	76.7% (n=46/60)	-
* team activities plan	58.3% (n=35/60)	-
* work meetings	56.7% (n=34/60)	-
* workload policy	45% (n=27/60)	-
* competence development plan	28.3% (n=17/60)	-
* teams develop a team- and educational program conform the regulations	-	66.0% (n=31/47)
* a better and larger management team	-	57.4% (n=27/47)
* management sets guiding standards	-	31.9% (n=15/47)
* management is permanently accessible	-	36.2% (n=17/47)

*management coaches teachers in educational leadership	-	19.1% (n=9/47)
*management eliminates 'cumbersome' administrative procedures	-	8.5% (n=4/47)
% of participants that feels exposed to the intervention	56.9% (n=33/58)	39.1% (n=18/46)
20. Readiness for change (% employees who feel ready for the change)	46.4% (n=26/56)	54.3% (n=25/46)
21. Satisfaction (SD)^b with implementation of the action plan/advisory report	5.7 (1.4) (n=58)	4.4 (1.7) (n=46)

^a Since no action plan was constructed in school B, the question concerned the advisory report. ^b SD is standard deviation. Being content with action plan was rated on a 1-10 scale (1=very unsatisfied; 10=very satisfied) and the average satisfaction rate could be classified as poor (<6), moderate (≥6 and <7.5), or high (≥7.5). ^c The quick wins are the 'adequate facilities' in school A.

11. Did employees participate in formulating an action plan? (participation)

In both schools, the majority of participants felt responsible for development and execution of the action plan (Table 5), suggesting that ownership of the intervention was high.

In school A, quantitative data showed that the majority of employees felt involved in the action plan (55.0%; Table 5). However, qualitative data revealed that employees on the lowest level of the schools hierarchy (e.g. administrative staff), saw the fewest possibilities to participate, suggesting that managers did not sufficiently encourage employees at all levels to participate. Furthermore, the contact with formal employee representatives (i.e. Works Council) in formulating the action plan was poor. Lastly, halfway through the intervention period the composition of the participatory group changed and it functioned as a feedback group for the director rather than a group that joined decision making.

In school B, employees perceived that there was no actual participation of employees in the second phase of the intervention (Table 5). The qualitative data showed that the implementers agreed with this perception of employees, because they did not undertake efforts to involve employees in formulating the Management Contract. Contact between implementers and formal employee representatives (i.e. Works Council) was scarce.

We can conclude that both implementers and employees showed high ownership of the intervention, but procedures to ensure participation were insufficient (school A) or lacking (school B) and therefore participation in formulating the action plan was low in both schools (Table 2).

12. Did the action plan target the right problems in the workplace? (targeting)

Qualitative analysis showed that, according to the implementers in school A, the action plan suited the problems in the workplace. The middle management team reported that the choice for the activities in the action plan was obvious, since it were all things that had to be done already to meet requirements from the school, Inspectorate or Ministry of Education. The participants partially agreed with this view. Teachers and the participatory group were disappointed that they could not participate in the translation of the advisory report into

an action plan. To them, the link between their problems and the actions taken was unclear. Quantitative analysis showed that satisfaction with the content of the action plan was moderate in school A (6.5; Table 5).

In school B, qualitative data showed that the participants were unable, even with help from members of the Works Council, to discover how the advisory report was translated into the Management Contract. Quantitative analysis showed that satisfaction with the content of the action plan was poor in school B (5.4; Table 5).

In sum, the implementers in school A and B thought action plans targeted the actual problems in the workplace, whereas employees in school A only partly agreed to that and employees in school B did not agree to that (Table 2).

13. Did senior management support the action plan? (senior management support)

In school A, the senior management (i.e. Executive Board) financially supported the project throughout. However, most teachers perceived the Executive Board as obstructing certain desired innovations (e.g. wireless Internet), whilst the participatory group noted that the Board provided constant support. This latter view was shared by all implementers.

In school B, senior management supported the project at first, but decided not to financially support the second phase of the intervention by hiring an HM facilitator again. The Board rather advised to make use of the services of an internal advisor. The Executive Board agreed with the management team that the advisory report's recommendations were aligned with the Management Contract. And, in retrospect, the Executive Board doubted the decision to take part in the intervention project at all, because the intervention interfered with an ongoing transition towards self-managing teams.

Thus, in both schools senior management was partly supportive (Table 2).

14. Was the action plan communicated to the employees? (communication)

In school A, the majority of the participants was informed about the progress towards the action plan's goals (86.7%), especially via information meetings organized by the management team (75%; Table 5). Qualitative data showed that the implementers thought their communication strategy was coherent with their implementation strategy, whereas the participants thought communication of actions or results was foremost (and for unclear reasons) initiated too late. Small successes in the implementation of actions were celebrated (e.g. cake on the opening day of the staff room).

In school B, a minority of the participants was informed about the progress towards the advisory report's goals (38.3%; Table 5). The participants named several examples of malfunctioning communication due to indecisiveness by the management or due to untimely planning. The management team acknowledged that communication was a concern and that reflections on decisions or planning were lacking.

We conclude that communication of the action plan and its progress was well-organized and effective at the school that formulated the action plan (school A), but less well and effective at the school that did not formulate an action plan (school B; Table 2).

15. Was the action plan implemented by middle managers? (delivery)

In school A, the implementers noted that several changes had been made. They wanted to make the changes consecutively and not parallel, so they started with three of the six intervention activities of the action plan (i.e. 'a dialogue on the organization of education', 'performance reviews' and 'team activities plan'). The other three intervention activities were due later. Some middle managers pointed to the importance of proper and timely implementation, especially of quick wins. For example, at one of the two school locations the staffroom was placed amidst the class rooms and nicely decorated, whereas this was not the case at the other location. This led to corresponding differences in satisfaction with the quick wins per location.

In school B, some of the middle managers perceived little or no changes as a result of the intervention. They found possible progress hard to determine, because of the alignment of the recommendations in the advisory report with the management contract. Senior management and directors reported beginning, yet unstable positive changes towards the goals in the management contract. Furthermore, the identified quick wins were not followed-up in the management contract.

In sum, the implementers in school A perceived the action plan as implemented, whereas this was not fully the case in school B (Table 2).

16. Were the employees exposed to implementation of the action plan? (exposure)

The majority of the participants in school A found that the quick wins were implemented (83.6%, Table 5). However, execution of these quick wins was considered not timely and satisfaction with the way quick wins were implemented differed at the two school locations. Still, two third of the participants (67%) thought the quick wins were celebrated as a success (Table 5). In school B were no quick wins formulated.

In school A, all participants reported to have been informed about the existence of the action plan (Table 5). The majority of the participants was familiar with the other two goals of the plan (83.6%; 90.2%) and some changes towards these goals were noted (43.3%; 43.3%; Table 5). The majority (56.9%) did feel exposed to the intervention in general. Changes were observed by 28.3%-78.3% of the employees, depending on the intervention activity in the action plan (Table 5). The three intervention activities that were delivered according to the implementers indeed had the highest scores ('a dialogue on the organization of education', 'performance reviews' and 'team activities plan'). Qualitative data showed however that participants judged the changes in general to be minor, some of the improvements were considered already commonplace

before the intervention or not really an improvement after all.

In school B, a majority of the participants (87.8%) reported to have been informed about the existence of the advisory report (Table 5). Most of them (67.3%-89.8%) were familiar with the two goals of the advisory report and some changes were noted towards these goals (25.5% and 59.6%; Table 5). A minority (39.1%) did feel exposed to the intervention in general. Percentages of perceived implementation of the recommendations in the advisory report were generally lower than in school A: improvements were observed by 8.5%-66.0% of the participants, depending on the intervention activity (Table 5). Qualitative data also showed that the participants perceived little or no changes as a result of the intervention. They also reported that possible progress was hard to determine, because of the alignment of the recommendations in the advisory report with the management contract.

Overall, employees in both schools were aware of the action plan (or management contract) and its content. In school A more intervention activities were perceived as implemented than in school B, but the actual changes were perceived as minor in both schools (Table 2).

17. Did the organizational culture facilitate the implementation of the action plan? (culture)

In school A, implementers and employees reported that the intervention department's culture was characterized by a distant relation between management and 'shop floor', which hindered implementation because of lacking mutual trust.

In school B, both the implementers and participants pointed to the 'collective history' of this school as a reason for lacking implementation. According to the implementers, teachers demonstrated either some sort of learned helplessness, or high levels of independency, which hindered the implementation process. Some of the employees felt that over the years formerly friendly hierarchical relations had developed into business relations, constructing 'a culture of fear', and this belief also hindered the implementation process.

In short, the culture was a hindrance in implementing the intervention in both schools (Table 2).

18. Did the organization have the capacity to implement the action plan? (conditions)

In school A, an HM facilitator was temporarily hired to facilitate the second phase of the intervention. The implementers considered the intervention time and energy consuming, whereas the investment was only perceived as 'returned' if progress was notable for all involved. The participants perceived the structured, target-driven way of implementing the intervention activities in the second phase not suited to the school structure that valued autonomy of the individual teacher.

In school B senior management insisted that the middle managers received coaching in implementing the management contract, because without the coaching, they had too little confidence that the management contract would be realized. The management team changed during the implementation phase: only two of the seven members stayed in position, including only one of the two directors. The two directors were responsible for executing the management contract and this hindered the implementation process.

In sum, school A had more capacities to implement the action plan than school B (Table 2).

19. Did no events interfere with the implementation of the action plan? (events)

The Ministry of Education announced a structural change for the secondary vocational education and training sector, implying an intensification of classroom-bound lessons from 850 hours to 1000 hours per year. In School A, an intensification trajectory was developed to redesign the curriculum (in school B, no preparations were made to make this change). At the same time in school A, fewer students enrolled and therefore inflow of tuition fees decreased. Temporary contracts had to be ended. In parallel, the Executive Board decided that the seven week summer holiday had to be reduced to six weeks. All events hindered the implementation of intervention activities, foremost because not every activity could be maintained in these changing circumstances.

In school B, two curriculums and their 34 respective teachers were positioned in another department within the greater college. Secondly, teams were reorganized and their composition changed. Thirdly, only two of the seven earlier members of the management team were still there at the end of the implementation phase. Fourthly, a third party advisor, other than the consultant appointed by the Board, came and went. Fifthly, structural changes in the Dutch healthcare sector were planned, which influenced the composition of the educational curriculum. All events hindered implementation, foremost because they interfered with the planned changes in the management contract.

Thus, at both schools unexpected events negatively interfered with the implementation process over the course of 24 months (Table 2).

20. Was the majority of the employees at T2 ready for the change? (readiness for change)

Qualitative analysis of school A showed that, the participants were disappointed in the lack of timely progress towards the action plan's goals, after the positively perceived first phase of the intervention. During the implementation phase the organization functioned top-down instead of bottom-up, and the actions in the action plan were all considered management instruments, leading to skepticism on the shop floor. The quantitative analysis confirmed this, since a minority scored above the cut-off point on the readiness for change scale (46.4%, Table 5).

In school B, qualitative data showed that the implementers and the participants really wanted positive changes, but both did not know how to contribute to this change process and the intervention did not support this sufficiently. The middle and senior managers reported beginning, yet unstable positive changes in the way managers and employees interacted in general and about the Management Contract. This seemed to be reflected in the quantitative analysis of readiness for change in participants: a majority was ready for change at T2 (54.3%, Table 5).

In sum, readiness for change was insufficient at T2 in school A due to disappointment in the result of the intervention activities, whereas in school B the majority seemed to be ready for the change (Table 2).

21. Were the employees satisfied with intervention phase 2? (satisfaction)

In both schools, overall satisfaction with the implementation of the action plan was poor (Table 5). The low satisfaction scores are considered the result of the sum of program requirements that were not met (Table 2).

22. Did implementers and employees perceive the intervention as positive? (perception)

In school A, the implementers were shocked by the problem definition in the advisory report, leading to a state of apathy which slowed down initial actions. Implementers noted that the needs assessment phase led to participants' high expectancies of quick changes, whereas the action plan had to be drafted and actions implemented, which was time-consuming. And exactly this time-consuming process proved the to the participants that the organizations' problems were still not solved.

In retrospect, the implementers in school B expressed their doubts about the whole intervention project, since the added value compared to all ongoing measures was unclear to them. The participants were disappointed in the intervention project since no progress was observed.

In summary, the intervention was not fully perceived as positive in both schools (Table 2).

Discussion

Our objective was to systematically evaluate the quality of the implementation process of a primary intervention for workplace mental health, by using a theoretical framework. The research question was whether the use of this process evaluation framework helps us to understand why the implementation was or was not successful.

Main findings

We have demonstrated that the Nielsen and Randall process evaluation framework indeed helps us to understand why the implementation was not successful and before discussing the implications, we summarize the results (Table 2). In school A, all intervention requirements for successful implementation of intervention phase 1 were met. However, it seems that not all employees were reached (no.5) and were only partly satisfied (no.7). In school B, two intervention requirements for successful implementation were not met: employees did not feel involved in the intervention (no. 3) and only a minority was ready for the change (no. 9). Furthermore, as in school A, reach (no. 5) and satisfaction (no. 7) were only partly met and in addition, managerial support (no. 8) was confirmed in the quantitative data but not in the qualitative in school B.

In the second phase in both schools most requirements were not met, except for middle management support (no. 10), communication (no. 14) and delivery (no. 15) in school A. Some requirements were partly met, that is senior manager support (no. 13) and exposure (no. 16) for both schools. Mixed findings, wherein results from quantitative and qualitative analysis pointed in a different direction (yes/no), were found for the requirements participation (no. 11) in school A, and middle manager support (no. 10) and readiness for change (no. 20) in school B.

Findings in context of the literature

The results resemble the findings reported in the literature on the implementation of participatory action approaches. That is, it has been convincingly demonstrated that partial implementation can have detrimental effects on commitment to the intervention [49]. Partial implementation might be due to a range of factors that all were present to some extent in the current study: a changing organizational context, low ownership of stakeholders, and flaws in the intervention design [49].

Managing the perceptions of the intervention process seems of utmost importance for successful implementation of the intervention. In our study, participants who felt not involved in the intervention did not perceive changes, as has been demonstrated before [50]. The notion that the perception of the change may be even more important than actual exposure to the change [33, 34] was underlined by this finding. The even more explicit statement that “more harm

can be done by disappointing expectations than by not conducting an intervention” [51] was also confirmed by the results of this process evaluation.

Furthermore, the organizational culture and conditions were suboptimal for implementing an intervention in both schools. The participating schools chose the intervention groups instead of random allocating them to an intervention or control condition, and trust between managers and employees in the appointed intervention groups turned out to be fragile at both schools. The earliest participatory action research studies in occupational health already showed that involvement in the intervention only led to increased participation in decision making when cooperative relations were present (e.g. [32]). The intervention did not succeed entirely to overcome the suboptimal preconditions for implementation. Organizational culture and the capacity to implement the intervention (conditions) are requirements for successful implementation that can be assessed before the start of the intervention, so to decide whether the intervention project should be initiated at all.

Implications for research

To our knowledge, this is one of the first organizational level participatory interventions evaluated by using the Nielsen and Randall framework [17]. In our experience the framework is a comprehensive tool for designing a process evaluation and accompanying data collection. With this framework, we captured context factors and perceptions that are known to be of importance in this type of intervention. For example, because of the “participant’s mental models” dimension we were able to demonstrate that the partial implementation was due to disappointment in the results of phase 2, which probably reduced readiness for change and definitely increased negative perceptions of the intervention and the organization.

Process evaluation and frameworks for process evaluation are in a dynamic development process. We think process evaluation should be based on 1) a combination of quantitative and qualitative data, and 2) the most appropriate model which is available at the time when the evaluation is done. First, adding qualitative measures increases the chance of grasping the story behind the numbers and we demonstrated that this story might point in a different direction for middle management support (10; school B), participation (11; school A), and readiness for change (20; school B). Second, future process evaluations might use other frameworks building on experiences with existing frameworks. We highlight three recent frameworks that might be helpful, in addition to the one applied and the ones already mentioned in the introduction. The framework by Fridrich and colleagues [52], which suggests to view the entire intervention process as a continuous change and thus suggests to blur the line between process and outcome evaluation by distinguishing between proximate, intermediate and distal outcomes of interventions. Or the model by Von Thiele Schwarz and colleagues [53], that integrates intervention design and intervention evaluation, because the collected data is used to adapt

intervention activities. Or otherwise the generic model by Damschröder and colleagues [54], that was based on a combination of existing frameworks and contains several potentially relevant domains for the study of implementation processes (i.e. the intervention, the outer setting, the inner setting, individual characteristics, the process).

In the current evaluation, the initiation and needs assessment phase plays a large role, and even though these are part of the Nielsen and Randall framework we had to develop a number of evaluation aspects for the initiation and needs assessment phase, since they were not listed in the model (e.g. satisfaction). And while the framework is useful for designing the process evaluation and data collection, it is less useful in describing the results. Foremost because the framework does not provide a strategy to relate the fragmented process components to each other, it is unclear which process components are most important. This leaves room for the individual researcher to interpret or weigh the components at her discretion, and that can be a risk for the replicability of process evaluations and the generalizability of their conclusions. There are too few detailed process evaluations that would make testing these process components possible, so we encourage researchers to conduct more detailed process evaluations. We recommend to develop a program theory that lists all requirements for successful implementation of the intervention. Then use the framework to operationalize all requirements for successful implementation, assess whether requirements are met by comparing the actual implementation to the requirements in the program theory [55].

There is also a downside to the comprehensiveness of the framework. Extensive data collection needs to be carried out to cover all headings in the framework, especially if one is interested in perceptions of different stakeholders and changes over time. This can be difficult and is time-consuming. Moreover, it can be challenging to present the huge amount of data in an attractive but concise manner. However, it seems important to continue conducting detailed process evaluation studies in order to advance our knowledge about what is needed to make organizational level and participatory workplace interventions work. Less demanding methods for doing detailed process evaluations need to be developed, but we can only do that if we know more about process evaluations, and this study contributes to that knowledge base.

Implications for practice

The HM intervention was developed by a Dutch consultancy firm and had been applied over a hundred times in the last decade mainly in public organizations, prior to evaluation within the current trial. We, in the role of independent researchers, expected the intervention to do well in a trial after such extensive piloting and adjusting. We also expected positive results because the intervention encompassed a participatory action approach, which provides an implementation strategy [39, 56, 57]. However, we encountered two flaws in the intervention design. Firstly, the protocol did not support the transition from HM facilitator as a primary implementer in phase 1 to the management team

in phase 2. Neither did the intervention protocol provide sufficient guidance on what to do if the management did not take account of the advisory report. As a result, the intervention process stagnated at first due to resistance to the advisory report among the management teams in both schools. The intervention ought to be revised at this point, for example by specifying the intervention protocol and managing the schools' expectations of the intervention better and from the beginning. Secondly, the intervention ought to facilitate fast implementation of quick wins, so to fulfill expectations and make optimal use of the readiness for change resulting from phase 1. There seems to be a 'window of opportunity' wherein actual changes will be perceived. If this window is missed, it will be hard to successfully implement the intervention.

Strengths and limitations

A strength of the current study is the use of a theoretical framework, which is not yet commonplace according to a review of process evaluations of stress management interventions [25]. The long term follow-up is also an apparent strength (e.g. [58]), which makes the findings worthwhile. The mixed methods design dealt effectively with both recall bias and common method bias, since both objective and subjective measures were combined.

A limitation of the current study is the rather low response rate we encountered at first follow-up, despite all efforts to increase the response. This is a common problem in intervention research (with intensive process evaluation)[59]. Since the first phase of the intervention consisted of a comprehensive needs assessment, which drew on a number of research methods (i.e. interviews, questionnaire, group sessions), participants supposedly became "research-tired" when receiving the T1 and later on T2 questionnaire. Furthermore, 34 employees in school B were relocated outside the intervention department shortly prior to T1, which might have hampered their motivation to fill out the questionnaire. The low response to the questionnaires limits the generalizability of the results somewhat. However, since different data sources were triangulated while studying the same phenomenon (i.e. different sources, different type, different measurement moments), we were able to cross verify our findings, which still makes them worthwhile.

Finally, a shortcoming in the evaluation of organizational level occupational health interventions is that process and effect data are often not combined. This comprehensive process evaluation did not integrate data either. However, we can formulate more specific hypotheses about the intervention effects than we could have done without this process evaluation: implementation of the intervention's first phase was more successful than implementation of the intervention's second phase in both schools. We thus expect to find an effect of the intervention's first phase on occupational self-efficacy, but no effect or even a negative effect on organizational efficacy, job characteristics and health outcomes as a result of the second phase.

Conclusion

The Nielsen and Randall process evaluation framework helped us to understand why the implementation process of an intervention was not successful and it is therefore considered of added value for the evaluation of implementation processes in participatory organizational level interventions. However, the framework requires collecting a large amount of qualitative and quantitative data and extensive data analysis. Less demanding methods for doing detailed process evaluations need to be developed. This can only be done if we know what are the most important process components and this study contributes to that knowledge base.

Abbreviations

HM: Heuristic Method, the trademark name of the intervention under study; T0: Baseline measurement; T1: First follow-up measurement; T2: Second follow-up measurement.

References

- Johnson S, Cooper C, Cartwright S, Donald I, Taylor P, Millet C: **The experience of work-related stress across occupations.** *J Manage Psychol* 2005, **20**(2):178-187.
- Montgomery C, Rupp AA: **A meta-analysis for exploring the diverse causes and effects of stress in teachers.** *Canadian Journal of Education* 2005, **28**(3):458-486.
- Klassen R, Wilson E, Siu AFY, Hannok W, Wong MW, Wongsri N, Sonthisap P, Pibulchol C, Buranachaitavee Y, Jansem A: **Preservice teachers' work stress, self-efficacy, and occupational commitment in four countries.** *European Journal of Psychology of Education* 2013, **28**(4):1289-1309.
- van Zwieten MHJ, de Vroome EMM, Mol MEM, Mars GMJ, Koppes LLJ, van denBossche SNJ: **Nationale Enquête Arbeidsomstandigheden 2013: Methodologie en globale resultaten.** [Netherlands Working Conditions Survey 2013: Methodology and overall results]. Hoofddorp: TNO; 2014.
- Franco C, Mañas I, Cangas AJ, Moreno E, Gallego J: **Reducing teachers' psychological distress through a mindfulness training program.** *Spanish Journal of Psychology* 2010, **13**(2):655-666.
- Gold E, Smith A, Hopper I, Herne D, Tansey G, Hurland C: **Mindfulness-based stress reduction (MBSR) for primary school teachers.** *J Child Fam Stud* 2010, **19**(2):184-189.
- Poulin PA, Mackenzie CS, Soloway G, Karayolas E: **Mindfulness training as an evidenced-based approach to reducing stress and promoting well-being among human services professionals.** *International Journal of Health Promotion and Education* 2008, **46**(2):72-80.
- Siu OL, Cooper CL, Phillips DR: **Intervention studies on enhancing work well-being, reducing burnout, and improving recovery experiences among Hong Kong health care workers and teachers.** *International Journal of Stress Management* 2014, **21**(1):69-84.
- Unterbrink T, Pfeifer R, Krippel L, Zimmermann L, Rose U, Joos A, Hartmann A, Wirsching M, Bauer J: **Burnout and effort-reward imbalance improvement for teachers by a manual-based group program.** *Int Arch Occup Environ Health* 2012, **85**(6):667-674.
- Zońmierczyk-Zreda D: **An intervention to reduce work-related burnout in teachers.** *International journal of occupational safety and ergonomics : JOSE* 2005, **11**(4):423-430.
- Ross SW, Romer N, Horner RH: **Teacher well-being and the implementation of school-wide positive behavior interventions and supports.** *Journal of Positive Behavior Interventions* 2012, **14**(2):118-128.
- Van der Klink JJJ, Blonk RWB, Schene AH, Van Dijk FJH: **The benefits of interventions for work-related stress.** *Am J Public Health* 2001, **91**(2):270-276.
- Richardson KM, Rothstein HR: **Effects of Occupational Stress Management Intervention Programs: A Meta-Analysis.** *J Occup Health Psychol* 2008, **13**(1):69-93.
- Egan M, Bamba C, Petticrew M, Whitehead M: **Reviewing evidence on complex social interventions: appraising implementation in systematic reviews of the health effects of organisational-level workplace interventions.** *J Epidemiol Community Health* 2009, **63**(1):4-11.
- Murta SG, Sanderson K, Oldenburg B: **Process evaluation in occupational stress management programs: a systematic review.** *Am J Health Promot* 2007, **21**(4):248-254.
- Semmer NK: **Job stress imbalance improvement and the organization of work.** *Scandinavian Journal of Work, Environment and Health* 2006, **32**(6):515-527.
- Nielsen K, Randall R: **Opening the black box: Presenting a model for evaluating organizational-level interventions.** *European Journal of Work and Organizational Psychology* 2013, **22**(5):601-617.
- Goldenhar LM, LaMontagne AD, Katz T, Heaney C, Landsbergis P: **The intervention research process in occupational safety and health: an overview from the National Occupational Research Agenda Intervention Effectiveness Research team.** *J Occup Environ Med* 2001, **43**(7):616-622.
- Hawe P, Shiell A, Riley T: **Complex interventions: How "out of control" can a randomised controlled trial be?** *Br Med J* 2004, **328**(7455):1561-1563.

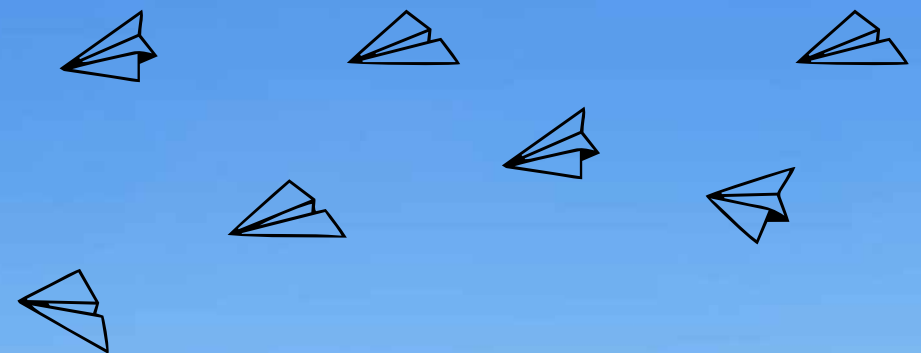
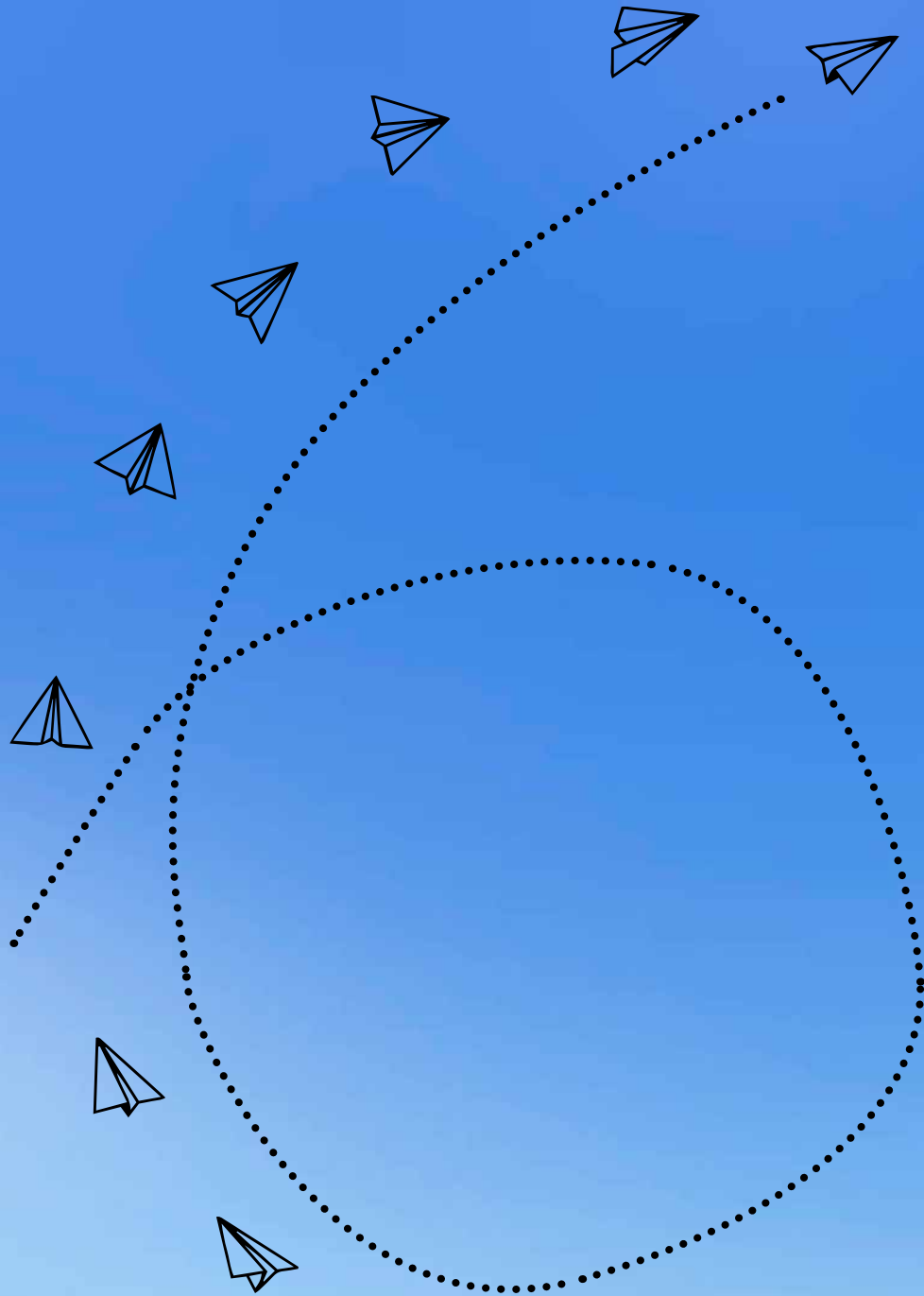
20. Oakley A, Strange V, Bonell C, Allen E, Stephenson J: **Process evaluation in randomised controlled trials of complex interventions.** *Br Med J* 2006, **332**(7538):413-416.
21. Armstrong R, Waters E, Moore L, Riggs E, Cuervo LG, Lumbiganon P, Hawe P: **Improving the reporting of public health intervention research: advancing TREND and CONSORT.** *J Public Health (Oxf)* 2008, **30**(1):103-109.
22. Cooper C, Dewe P, O'Driscoll M: **Organizational interventions.** In *Organizational stress: A review and critique of theory, research and applications.* Edited by Cooper C, Dewe P, O'Driscoll M. CA: Thousand Oaks; 2001:187-232.
23. Grol R, Grimshaw J: **From best evidence to best practice: effective implementation of change in patients' care.** *Lancet* 2003, **362**(9391):1225-1230.
24. Nielsen KN: **Review Article: How can we make organizational interventions work? Employees and line managers as actively crafting interventions.** *Human Relations* 2013, **66**(8):1029-1050.
25. Havermans BM, Schelvis RMC, Boot CR, Brouwers EP, Anema JR, van der Beek AJ: **Process variables in organizational stress management intervention evaluation research: a systematic review.** *Scand J Work Environ Health* 2016, **42**(5):371-381.
26. Steckler A, Linnan L: *Process evaluation. For public health intervention and research:* San Francisco: Jossey-Bass; 2002.
27. Nielsen K, Abildgaard JS: **Organizational interventions: A research-based framework for the evaluation of both process and effects.** *Work Stress* 2013, **27**(3):278-297.
28. Biron C, Karanika-Murray M: **Process evaluation for organizational stress and well-being interventions: Implications for theory, method, and practice.** *International Journal of Stress Management* 2014, **21**(1):85-111.
29. Pawson R, Tilley N: *Realistic Evaluation:* London: SAGE Publications Inc; 1997.
30. Coffeng JK, Hendriksen IJM, Van Mechelen W, Boot CRL: **Process evaluation of a worksite social and physical environmental intervention.** *Journal of Occupational and Environmental Medicine* 2013, **55**(12):1409-1420.
31. Van Berkel J, Boot CRL, Proper KI, Bongers PM, van der Beek AJ: **Process evaluation of a workplace health promotion intervention aimed at improving work engagement and energy balance.** *Journal of Occupational and Environmental Medicine* 2013, **55**(1):19-26.
32. Heaney CA, Israel BA, Schurman SJ, Baker EA, House JS, Hugentobler M: **Industrial relations, worksite stress reduction, and employee well-being: A participatory action research investigation.** *J Organ Behav* 1993, **14**(5):495-510.
33. Nielsen K, Randall R, Albertsen K: **Participants' appraisals of process issues and the effects of stress management interventions.** *J Organ Behav* 2007, **28**(6):793-810.
34. Hasson H, Brisson C, Guérin S, Gilbert-Ouimet M, Baril-Gingras G, Vézina M, Bourbonnais R: **An organizational-level occupational health intervention: Employee perceptions of exposure to changes, and psychosocial outcomes.** *Work Stress* 2014, **28**(2):179-197.
35. Schelvis RMC, Oude Hengel KM, Wiezer NM, Blatter BM, van Genabeek JA, Bohlmeijer ET, van der Beek AJ: **Design of the Bottom-up Innovation project--a participatory, primary preventive, organizational level intervention on work-related stress and well-being for workers in Dutch vocational education.** *BMC Public Health* 2013, **13**:760-2458-13-760.
36. Fredslund H, Strandgaard J: **Methods for process evaluation of work environment interventions.** In *6th Annual Conference of the European Academy of Occupational Health Psychology: Proceedings. 24-26 November 2004* Edited by Houdmont J, McIntyre S. PUBLISMAI Publishers Oporto, Portugal; 2005:109-117.
37. Hurrell JJ, Murphy LR: **Occupational stress intervention.** *Am J Ind Med* 1996, **29**(4):338-341.
38. Nielsen K, Randall R, Holten A-, González ER: **Conducting organizational-level occupational health interventions: What works?** *Work Stress* 2010, **24**(3):234-259.
39. Randall R, Cox T, Griffiths A: **Participants' accounts of a stress management intervention.** *Human Relations* 2007, **60**(8):1181-1208.
40. Nielsen K, Randall R: **Managers' Active Support when Implementing Teams: The Impact on Employee Well-Being.** *Applied Psychology: Health and Well-Being* 2009, **1**(3):374-390.
41. Nytrø K, Saksvik PØ, Mikelsen A, Bohle P, Quinlan M: **An appraisal of key factors in the implementation of occupational stress interventions.** *Work & Stress* 2000, **14**(3):213-225.
42. Weick KE, Sutcliffe KM, Obstfeld D: **Organizing and the process of sensemaking.** *Organization science* 2005, **16**(4):409-421.
43. Johns G: **The essential impact of context on organizational behavior.** *Academy of management review* 2006, **31**(2):386-408.
44. Mathieu JE, Heffner TS, Goodwin GF, Cannon-Bowers JA, Salas E: **Scaling the quality of teammates' mental models: Equifinality and normative comparisons.** *J Organ Behav* 2005, **26**(1):37-56.
45. Randall R, Nielsen K, Tvedt SD: **The development of five scales to measure employees' appraisals of organizational-level stress management interventions.** *Work Stress* 2009, **23**(1):1-23.
46. Greene JC, Benjamin L, Goodyear L: **The merits of mixing methods in evaluation.** *Evaluation* 2001, **7**(1):25-44.
47. Green J, Thorogood N: *Qualitative methods for health research:* Sage; 2013.
48. Biron C, Gatrell C, Cooper CL: **Autopsy of a failure: Evaluating process and contextual issues in an organizational-level work stress intervention.** *International Journal of Stress Management* 2010, **17**(2):135-158.
49. Landsbergis PA, Vivona-Vaughan E: **Evaluation of an occupational stress intervention in a public agency.** *J Organ Behav* 1995, **16**(1):29-48.
50. Aust B, Rugulies R, Finken A, Jensen C: **When workplace interventions lead to negative effects: Learning from failures.** *Scand J Public Health* 2010, **38**(SUPPL. 3):106-119.
51. Fridrich A, Jenny GJ, Bauer GF: **The context, process, and outcome evaluation model for organisational health interventions.** *BioMed research international* 2015: 1-12.
52. Thiele Schwarz U, Lundmark R, Hasson H: **The dynamic integrated evaluation model (DIEM): achieving sustainability in organizational intervention through a participatory evaluation approach.** *Stress Health* 2016, **32**(4):285-293.
53. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC: **Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science.** *Implementation science* 2009, **4**(1):50.
54. Kompier M, Aust B: **Organizational stress management interventions: Is it the singer not the song?** *Scand J Work Environ Health* 2016, **42**(5):355-358.
55. Westgaard RH, Winkel J: **Occupational musculoskeletal and mental health: Significance of rationalization and opportunities to create sustainable production systems - A systematic review.** *Appl Ergon* 2011, **42**(2):261-296.
56. Dollard MF, Le Blanc PM, Cotton SJ: **Participatory action research as work stress interventions.** In *The individual in the changing working life.* Edited by Naswall K, Hellgren J, Sverke M. Oxford: Cambridge University Press; 2008: 353-379.
57. Hugentobler MK, Israel BA, Schurman SJ: **An action research approach to workplace health: integrating methods.** *Health Educ Q* 1992, **19**(1):55-76.
58. McVicar A, Munn-Giddings C, Seeborn P: **Workplace stress interventions using participatory action research designs.** *International Journal of Workplace Health Management* 2013, **6**(1):18-37.

An evaluation perspective



Evaluation of occupational health interventions using a randomized controlled trial: challenges and alternative research designs

Roosmarijn M.C. Schelvis
Karen M. Oude Hengel
Alex Burdorf
Birgitte M. Blatter
Jorien E. Strijk
Allard J. van der Beek



Abstract

Occupational health researchers regularly conduct evaluative intervention research for which a randomized controlled trial (RCT) may not be the most appropriate design (eg, effects of policy measures, organizational interventions on work schedules). This article demonstrates the appropriateness of alternative designs for the evaluation of occupational health interventions, which permit causal inferences, formulated along two study design approaches: experimental (stepped-wedge) and observational (propensity scores, instrumental variables, multiple baseline design, interrupted time series, difference-in-difference, and regression discontinuity). For each design, the unique characteristics are presented including the advantages and disadvantages compared to the RCT, illustrated by empirical examples in occupational health. This overview shows that several appropriate alternatives for the RCT design are feasible *and* available, which may provide sufficiently strong evidence to guide decisions on implementation of interventions in workplaces. Researchers are encouraged to continue exploring these designs and thus contribute to evidence-based occupational health.

Introduction

The randomized controlled trial (RCT) is considered the gold standard in evaluative medical research as causal inferences about the therapy under study can be drawn. The first RCT was reported in a 1948 issue of the *British Medical Journal* (BMJ) and involved the experimental treatment of pulmonary tuberculosis [1]. In this trial, a particular group of English tuberculosis patients from different care facilities, comparable in the symptoms of the disease and age were included. The included patients were assigned to either a combined medicine and bed-rest therapy, or bed-rest therapy alone, based on a statistical series of random sampled numbers. Neither the patients nor the doctors involved knew the condition the patient was assigned to, later to be named a “double blind” procedure. Therapy progress was reported on forms particularly designed for this trial. Due to this design, the researchers were able to demonstrate the added value of the combined treatment over the bed-rest treatment, but only in the first three months after onset of the disease. Thereafter a deterioration emerged, probably due to resistance to the medicine under study. Many researchers have followed this example ever since. The beauty of the randomization procedure is that chance (probably) ensures that known and unknown prognostic factors are balanced over the treatment conditions and thus do not interfere with the treatment–outcome relationship. Therefore, conclusive statements about the effectiveness of the therapy can be made.

In occupational health research, a typical RCT aims, for instance, to reduce productivity loss at work (ie, a primary outcome) for a randomly chosen group of employees with medically verified upper-extremity disorder (ie, specific characteristics) via an ergonomic assessment at the worksite and a physician contacting each employee’s supervisor to discuss potential accommodations at work (ie, a multicomponent intervention). The effectiveness of the intervention is evaluated by the change in primary outcome from pre- to post-test in the intervention group relative to the change in this outcome in the reference group that did not receive the intervention [2]. However, occupational health researchers are increasingly addressing questions regarding the outcomes of complex interventions. A complex intervention can consist of (i) multiple components, (ii) multiple providers and thus multiple levels, (iii) multiple locations, and/or (iv) multiple (varying) outcomes. The components, providers, locations and outcomes are interdependent and therefore the intervention can be difficult to standardize or administer uniformly [3–5]. Furthermore, the context is often complex and thus nearly impossible to control entirely [6]. Conducting an RCT on a complex intervention within an occupational health context is thus not always the most feasible option [7, 8].

The British Medical Research Council (MRC) recently published an updated guide that underlines the need for innovative evaluation methods [9]. Although the MRC considers individual randomization in trials as the most robust design to prevent allocation bias, it is more and more acknowledged that common evaluation methods are not always practical or ethical for com-

plex interventions [9]. The RCT sometimes even offers too little information to draw meaningful conclusions for science or practice. More specifically, an RCT allows conclusions on the effectiveness of the intervention for a selected sample of individuals. Researchers have argued that because of complexity in the intervention and context, the required conditions that are needed for an efficacy trial will never occur [10]. Even if an efficacy trial has been performed with success, then it still is “highly unlikely that interventions that do well in efficacy studies will do well in effectiveness studies, or in real-world applications” [10, p1262].

In order to further develop the evidence base in occupational health there is a clear need for alternatives to the RCT. These alternatives can be formulated along two lines: experimental (most often RCT variants) and observational studies [11]. Some (experimental) alternatives have been applied already in the occupational setting. The most commonly applied RCT variant is the cluster RCT, in which groups of individuals rather than individuals are randomized [12]. Cluster RCT typically involve two levels, the cluster (eg, department) and their individual members (eg, worker), although trials of more than two levels (eg, company, department, and worker) also exist [12]. Cluster RCT have several advantages over individual RCT in organizational interventions, namely (i) increased logistic feasibility in delivering the intervention, (ii) analysis and evaluation is conducted at the same level as the intervention is applied to (ie, the group), and (iii) contamination is avoided, which might occur when unblinded interventions are administered to some individuals but not to others in the same setting (eg, department, team, occupational physician) [13]. Another commonly applied variant is the controlled trial wherein a selected group of individuals or clusters receiving the intervention is compared to a reference group that is matched on known prognostic factors (eg, age) [14]. This design can be preferable to an individual or cluster RCT for practical or ethical reasons in an occupational setting. Apart from randomization, the controlled trial shares all characteristics with an RCT, but lacks the advantage of balanced unknown prognostic factors in both conditions.

However, for these alternative RCT designs, challenges remain that impede drawing causal inferences [15]. The cluster RCT, for example, needs much larger numbers of participants within an experimental setting, which is often problematic in terms of feasibility and costs. The controlled trial suffers from the non-random allocation to groups, which may introduce known and unknown factors to be unbalanced between both groups. This article presents an overview of other experimental and observational study designs for occupational health interventions, starting with an overview of practical challenges in conducting an RCT, the methodological consequences of these challenges, and an empirical example. Thereafter, the key features of each design are described, including the advantages and disadvantages, and how the challenges are minimized by applying this design.

Challenges in applying RCT to evaluate complex interventions

When conducting an RCT in the occupational setting, researchers faces challenges concerning the methodology (ie, randomization and control group), the intervention, and the context. Empirical examples for each challenge are given in Table 1.

Methodology: randomization

Randomization of participants to the experimental condition (intervention group) or usual care/placebo condition (control group) eliminates allocation bias and internal validity threats, maximizing the probability that (un)known confounding variables will be evenly distributed over groups [16].

Challenge 1. Only few clusters at the organizational level are available to evaluate the intervention.

Many workplace interventions are implemented at the group level (eg, company, facility, department, and team). The randomization procedure is then applied at the group or – in methodological terms – cluster level. However, recruiting enough clusters within a specific context is often difficult. If too few clusters are included, controlling by chance for all factors and conditions that might differ between groups is impeded. Consequently, there might be an unequal distribution of baseline characteristics between groups which introduces bias to the study [eg 17].

Challenge 2. The organization objects to random assignment of persons or departments.

In practice, acknowledgment of a problem which is unique to a certain department (eg, high sickness absence, lagging work performance) can be a strong driver for organizations to participate in intervention research. Targeting this department with an intervention is at least in their interest and at the most a precondition to participate. Thereby, companies obstruct randomization and potential bias is thus introduced. If the organization wants to decide on the allocation of the employees to the intervention and control group, the two most important resulting biases are confounding (ie, error due to a third variable that influences the exposure–outcome relation) [16] and selection bias (ie, error due to systematic differences in characteristics such as motivation, between intervention and control group) [16], which are difficult to overcome [18].

Methodology: control group

The effect of an intervention is measured as the difference on a certain outcome between the intervention group and the control group [18]. A control

Table 1 — Challenges, their examples, and consequences in occupational health intervention studies

Clustered main challenge	Challenge	Example ^a	Consequence
Randomization	Only few clusters at organizational level are available to evaluate the intervention	A case management approach for workers on prolonged sick leave was evaluated in one hospital compared to a neighboring hospital (74).	Potential confounding due to differences between groups resulting from unreliable randomization
Control group	The organization objects to random allocation to an intervention or control group	Department supervisors allocated participants to a prevention program, which could have reflected their subjective interest in the program, the time and workload to complete the program, and their expectations for enhancing workplace mental health (75).	Potential confounding due to differences between groups resulting from selection bias
Control group	The organization wants to target all employees with an intervention	An intervention on behavioral techniques was implemented in two departments within a hospital. However, the content of the intervention was different between the two departments (76).	Difficult to differentiate effects of the intervention from changes due to other causes resulting from the lack of a control group
Intervention	The organization or researcher wants to adjust the intervention protocol	The goal of the intervention was to increase the range of healthy foods on offer in worksite cafeterias in two supermarkets in The Netherlands so as to discourage eating snacks. There were conflicting interests between the intervention's goal and management's targets: the greatest profits came from selling snacks (77).	Difficult to differentiate which effects were caused by which element of the intervention resulting from changes in the intervention protocol
Context	The organization is subject to internal change	An intervention to improve sustainable employability was conducted across multiple companies and worksites. During the intervention period, workers from several worksites were discharged as result of the economic recession (78).	Potential confounding due to differences between groups resulting from selective loss to follow-up
Context	The organization is subject to external changes	In a construction company in The Netherlands, an intervention is introduced to decrease inhalation of particulate matter at the worksite. At the same time, the Labor Inspectorate decides to enforce regulations and plans to visit all worksites in the Netherlands (79).	Difficult to differentiate effects of the intervention from those due to unintended co-interventions resulting from external changes

^aThe empirical examples were selected on relevance and appropriateness for the occupational setting

group is needed to distinguish between change in outcome over time due to the planned intervention, or changes over time due to unmeasured or unknown factors (eg, a policy measure).

Challenge 3. The organization wants to target all employees with an intervention.

Organizations are often willing to participate in an intervention study if an acknowledged problem is to be solved (eg, high prevalence of low-back pain). Hence, the employer is motivated to demonstrate that (s)he takes the problem seriously and therefore demands that everyone should be able to participate. The employer considers it unethical to offer the intervention to a selected group only, while every employee has a potentially elevated risk. As a consequence, studies within the occupational health setting sometimes have to be performed without a control group, complicating the distinction between effect of the intervention and autonomous change over time.

Intervention

When following the guidelines of conducting an RCT, a predefined protocol for implementing and evaluating interventions is preferred in order to reach high internal validity [16]. High fidelity to the protocol is furthermore important in order to understand key intervention processes and functions, and thus enable answers to the question of why the intervention is or is not effective.

Challenge 4. The organization or the researcher wants to adjust the intervention protocol.

Either the organization or the researchers may want to adjust the intervention protocol to fit the specific context per cluster, thereby violating the standardization principle. For instance, the order of intervention components may be altered or the intervention components may be tailored to a specific group of workers or to specific occupational health problems. If adjustments are made within clusters, it becomes difficult to establish which intervention components or what implementation processes contributed to the effectiveness or lack thereof of the intervention, a situation sometimes referred to as a 'black box' [19, 20].

Context

For most occupational health interventions a double-blind-placebo trial is nearly impossible: complex interventions are dependent on the context in which they are applied. Moreover, besides the intentional adjustments described under challenge four, the intervention provider, the participants or the context may unintentionally influence the delivery and content of the intervention and thereby the outcomes (ie, information bias).

Challenge 5. The organization is subject to internal change.

Many worksites and departments are subject to continuous change [21]. For

example, within the participating department not only the intervention under study, but also a co-intervention is delivered. In hospitals, lifting devices may be introduced to reduce mechanical load among nurses, whereas simultaneously hospital beds are replaced by high-low beds that also reduce nurses' mechanical load. Implementing the intervention under fully controlled conditions is thereby impeded. A second example is a change in staffing: employees and managers change jobs or retire, new employees are hired, and teams are moved to other areas or downsized. Consequently, high loss to follow-up can be expected and a decreased study sample complicates reliable conclusions regarding intervention effects.

Challenge 6. The organization is subject to external change.

Even when the intervention is performed under controlled conditions within the company, external changes might interfere with the intervention [21]. For instance, increased enforcement of regulations by the Labour Inspectorate on the main outcome of the intervention might take place simultaneously (eg, exposure to dust containing quartz). Or, a nationwide campaign on work stress is implemented during the same period as a local stress management intervention, which motivates the control group to implement stress prevention measures as well. As a consequence of these so-called co-interventions, it becomes more difficult to distinguish autonomous change from effect, even if a control group is present.

In sum, difficulties with regard to methodology, intervention, and context may hamper the evaluation of complex occupational health interventions by means of an RCT. However, we fully agree with Kristensen [18], who stated that “there may be many good reasons for not performing a RCT in an occupational setting. But there are no good reasons for ignoring the problems created by not applying such a design.”

Alternatives for evaluating complex occupational health interventions

Several alternative experimental designs and designs using observational data are potentially interesting for the evaluation of complex interventions [14, 22]. The core team of authors discussed a list of potential alternatives for the occupational health setting and those most relevant and applicable to the occupational health setting were selected for this article. In contribution to the current debate on alternatives to randomization in the evaluation of public health interventions [9], the selection of alternatives is described based on theoretical literature and empirical examples (tables 2 and 3).

Table 2 — Overview of evaluative designs, their characteristics, and data requirements. [RCT=randomized controlled trial]

Design type	Allocation of intervention	Confounding	Data requirements	Measurements
Experimental				
RCT	Randomization at individual level	Known and unknown prognostic factors are balanced	Longitudinal data	Before and at least once after intervention
Cluster RCT	Randomization at group level	Known and unknown prognostic factors may be unbalanced over clusters	Longitudinal data	Before and at least once after intervention
Stepped-wedge randomized trial	Randomization of intervention to all individuals or groups sequentially over time	Known and unknown prognostic factors may be unbalanced over clusters and over time	Longitudinal data	Repeated before and after
Observational				
Propensity score method	Likelihood to have been offered the intervention	Matching, stratification or adjustment for known prognostic factors	Longitudinal data	Before and at least once after intervention
Instrumental variable method	Exposure to 'the instrument' predicts actual intervention received	No influence of unknown prognostic factors	Longitudinal data	Before and at least once after intervention
Multiple baseline design	Intervention to all individuals or groups sequentially over time	Adjustment for known prognostic factors	Longitudinal data	Repeated before and after
Interrupted time series	Intervention to all individuals at particular moment in time	Adjustment for known prognostic factors	Cross-sectional data	Multiple repeats over time (eg, routinely collected data) before and after
Differences-in-differences	Intervention to selected individuals at particular moment in time	Adjustment for known prognostic factors	Cross-sectional data	Multiple repeats over time (eg, routinely collected data) before and after
Regression discontinuity	Intervention to individuals at particular moment in time	No influence of unknown prognostic factors	Cross-sectional data	Multiple repeats over time (eg, routinely collected data) before and after

Table 3 — Alternative research designs, advantages and disadvantages, and the challenges overcome. [RCT=randomized controlled trial.]

Alternatives	Advantages	Disadvantages	Solution to challenges ^a
Stepped-wedge design	<ul style="list-style-type: none"> All participants and clusters receive the intervention Randomization of order in receiving the intervention, thus preventing selection bias Improves feasibility for practical, ethical and/or financial reasons Changes in protocol are possible before the next step 	<ul style="list-style-type: none"> More measurements and a longer time period High burden on participants and researchers Complexity of statistical analyses Decrease in statistical power and requires larger sample sizes 	3, 4
Propensity scores	<ul style="list-style-type: none"> Solution if randomization has unethical consequences Advantageous when effects are expected in the long term, which makes an RCT costly 	<ul style="list-style-type: none"> Very large sample sizes are needed If the propensity score is estimated incorrectly or the covariates are measured imperfectly, bias is introduced 	2
The method of instrumental variables	<ul style="list-style-type: none"> Solution if an RCT is not possible for practical reasons Instrumental variables can also correct for unmeasured confounders 	<ul style="list-style-type: none"> Rarely used in research practice because of strong assumptions Weak instruments lead to large standard errors 	2
Multiple baseline design	<ul style="list-style-type: none"> Fewer cohorts are required as cohorts act as their own controls Analyses can be done with routinely collected data, cohort data or (historic) reference groups 	<ul style="list-style-type: none"> Including fewer cohorts is only possible if the effect sizes are large and if the intervention results in rapidly observable changes in the outcome variable The autocorrelation can lead to inaccurate estimates of the intervention effect Sufficient baseline stability is needed The timing of the intervention introduction and duration in each cohort ought to be known before starting 	1, 3, 5, 6
Interrupted time series design	<ul style="list-style-type: none"> Suitable if establishing control groups is difficult Analyses can be based on routinely collected data, cohort data or external (historic) data or reference groups 	<ul style="list-style-type: none"> Underpowered studies because of the number of measurements before and after the intervention and the time lags of measurements One needs to be able to determine specifically at what time point the intervention started and ended 	2
Differences-in-differences	<ul style="list-style-type: none"> An elegant way to study 'naturally occurring' internal or external changes 	<ul style="list-style-type: none"> Sophisticated analyses required Does not account for invariant factors or macro trends that interfere with the outcome Under- or overestimation is a risk in individual level interventions 	5, 6
Regression discontinuity	<ul style="list-style-type: none"> Establish causal effects based on observational data Feasibility of this design can be improved by using routine clinical or administrative data 	<ul style="list-style-type: none"> Assumption that individuals around the threshold are similar is debatable The assignment variable possibly changes over time Requires larger samples sizes than an RCT 	1, 2, 3

^a Challenge 1: Only few clusters exist to cluster the intervention at organizational level; Challenge 2: The organization objects to random assignment of persons or departments; Challenge 3: The organization wants to target all employees with an intervention; Challenge 4: The organization or the researcher wants to adjust the intervention protocol; Challenge 5: The organization is subject to internal change; Challenge 6: The organization is subject to external change

Alternative design in experimental research

Stepped-wedge randomized trial

The stepped wedge randomized trial is a modification of the individual or cluster RCT in which an intervention is sequentially rolled-out to all participants over consecutive time periods [23]. The order in which the individuals or clusters receive the intervention is randomized, so that at the end of the entire time period all participants have received the intervention, thereby counteracting challenge 3 (*the organization wants to target all employees with an intervention*) [24]. The stepped-wedge design is particularly suitable if it is considered unethical to withhold the intervention from participants in a control group [25]. Additionally, the stepped wedge design allows for improvement of the intervention based on lessons learned in every subsequent step (which makes it very suitable for effectiveness trials in practice) and thereby eliminates challenge 4 (*the organization or the researcher wants to adjust the intervention protocol*). Due to the within and between cluster comparisons at each measurement time across all time periods, this design allows for a variety of conclusions: both short-and long-term effects, fade out effects, and the natural course of the condition under study [26].

For the evaluation of a care program for staff members in dementia special care units, a stepped-wedge design was used [27]. The care program consisted of tools and procedures to guide staff members through the detection, analysis, treatment and evaluation of residents' challenging behavior. After allocating seventeen units randomly to five groups, every four months a new group started with the intervention (24 months in total). Burnout, job satisfaction, and job demands were self-assessed before the start, midway and after the implementation process. The results of the multilevel analyses of 380 staff members showed a significant positive effect for job satisfaction [β 0.93, 95% confidence interval (95% CI) 0.48–1.38], whereas no statistically significant effects were found for burnout and job demands.

Although the stepped-wedge design helps to minimize or overcome two important challenges, it introduces new challenges in itself. These challenges are firstly that larger sample sizes might be required for some outcomes since, with the increased number of groups to compare, the design may have less statistical power than the regular (cluster) RCT [28, 29]. Secondly, the data collection in each time period can put a high burden on participants and researchers, which might hamper the feasibility of the study [29]. The design is most feasible if data can be (partly) routinely collected at the appropriate time intervals in a reliable and valid way [28]. Thirdly, statistical analysis is complex because both a random coefficient for cluster and a fixed effect for time need to be taken into account [23].

Alternative designs in observational studies

In observational studies, assignment to the experimental condition is not under the researchers' control. The intervention and control group may differ in (observed) covariates, which could lead to biased estimates of intervention effects. Hereafter we describe alternative evaluation designs specifically developed to evaluate interventions with observational data while dealing with potential bias.

Propensity scores

The propensity scores method is a statistical matching technique that can be applied to control for confounding in evaluation studies with observational data [30, 31]. The first step is to estimate propensity scores for all individuals, defined as the conditional probability of (a particular) exposure to the intervention given a number of confounding variables [32]. The propensity score can be estimated with logistic regression analysis, modeling the exposure as dependent variable and the potential confounders as independent variables [33]. Because some individuals with similar propensity scores are exposed to the intervention, whereas others with a similar score are not, the method assumes that actual exposure to the intervention within these individuals mimics randomization [34, 35], thereby counteracting challenge 2 (*the organization objects to random assignment of persons or departments to the intervention or control group*). Then, the intervention effect will be estimated using the propensity score through matching of individuals, stratification or regression adjustment [33, 36].

In a Finnish study, the propensity score was calculated for 24 000 persons in a cohort of public sector employees in municipalities and hospitals so they could be assigned to a multidisciplinary, vocational rehabilitation intervention to improve work ability [37]. The propensity score was calculated using logistic regression analysis with 25 variables, including demographics (eg, gender), work characteristics (eg, work schedules), health risk indicators (eg, psychological distress), and health risk behaviors (eg, smoking status) [38]. Once the propensity score was estimated, 859 employees who participated in the intervention were matched by propensity score with 2426 controls, thereby excluding all other, unmatched employees in the entire cohort. The intervention showed adverse effects on perceived work ability and no beneficial effects on work disability: the risk of suboptimal work ability was somewhat higher after short- and long-term follow-up for participants than for controls (prevalence ratio 1.23 and 1.18, respectively) [37], while an earlier study showed that incident long-term work disability was about the same for participants and controls (hazard ratio 0.98) [38].

Some conditions need to be fulfilled before propensity scores can be considered as an alternative. The method assumes that all important prognostic variables are included and the model can be built perfectly [33, 34]. If the propensity score is estimated or the covariates measured imperfectly, this bias may affect the estimated intervention effect [33]. One way to cope with this problem is to construct different sets of propensity scores to test its robustness [39–41].

Instrumental variables method

The method of an instrumental variable is well known in the field of economics and applied to explore causal relationships between the intervention and an outcome in longitudinal studies [42]. The method relies on finding a valid prediction variable, named "the instrument", that meets three assumptions: it (i) predicts the actual intervention received, (ii) is not directly related to the outcome, except by the direct effect of the intervention, and (iii) is not related to the outcome by any other measured or unmeasured path [42–44]. Elovainio and colleagues recently investigated the association of job demands and job strain with perceived stress, psychological distress and sleeping problems among elderly care workers [45]. Staffing level (ie, the ratio of the total number of nursing staff to the number of residents in the elderly care wards) appeared to be a strong instrument for both job demands and job strain, and instrumental regression analyses showed statistically significant associations with perceived stress and psychological distress. Self-reported job demands and job strain revealed the same results. An advantage of this method is that it provides a way to obtain a potentially unbiased estimate of treatment effect, even in the presence of strong unmeasured confounding [44]. Since instrumental variables predict compliance to an intervention (or actual exposure) but have, by definition, no direct, independent effect on the outcome, the method of instrumental variables can reach the same effect as randomization [44] and thereby counteracts challenge 2 (*the organization objects to random assignment of persons or departments to the intervention or control group*).

As an example of this method, Behncke [46] investigated the effects of retirement on various health outcomes (eg, self-assessed health, chronic condition, and biological measures). Of the 1439 individuals at baseline, 192 subjects retired during the two year follow-up. Behncke assumed that reaching the state pension age affected the retirement decision, but was not directly related to health outcomes. The analyses showed that state pension age was a good predictor for retirement and thus a strong instrument. The results of the instrumental analyses showed that retirement significantly increased the risk of being diagnosed with a chronic condition.

Choosing the correct instrument for the analysis is a crucial factor in this design. Weak instruments (ie, a low correlation between the instrument variable and intervention or exposure variable) lead to large standard errors resulting in imprecise and biased results when the sample size is small [47]. Therefore, this method is particularly useful for large samples and in case of moderate confounding.

Multiple baseline design

In a multiple baseline design the same intervention is implemented at different time points across groups with pre- and post-measurements [48, 49]. Outcome variables are measured in all groups at baseline. Then, one or more groups receive the intervention while others remain in the control condition. After sufficient time has passed for the intervention to affect the outcome, outcomes are again measured in all groups and the intervention is introduced in the next one or more groups [48–50]. This procedure minimizes challenge 3 (*the organization wants to target all employees with an intervention*). By sequentially

introducing the intervention to groups, patterns of unexpected internal or external events can be studied; counteracting challenge 5 and 6 (*the organization is subject to internal/external change*). Compared to the RCT, fewer groups of participants are required in the multiple baseline design, since the group also acts as its own control [49]; counteracting challenge 1 (*only few clusters exist to cluster the intervention at organizational level*). The design can be considered the non-randomized observational equivalent of the stepped-wedge design.

The evaluation of a behavioral contingency feedback intervention to increase attendance among 64 certified nursing assistants at three hospitals was conducted by applying a reversal (ie, ABA) multiple baseline design [51]. The nine-week intervention was introduced across three groups at 16, 19, and 21 weeks after baseline measurement. All groups returned to the baseline situation (ie, A) after receiving the intervention (ie, B). The study ended with a final measurement after 39 weeks. The hospitals provided the research team with the working schedules of the participants and their sickness absence records. The repeated measures analysis of variance showed that the total number of absent days per week decreased in the intervention period (mean 0.13, standard deviation (SD) 0.17) compared to baseline (mean 0.24, SD 0.19) and increased again after returning to the baseline situation (mean 0.24, SD 0.20).

The main statistical challenge in using the multiple baseline design is the high autocorrelation of repeated measurements over time, which can lead to imprecise estimates of the intervention effect [49]. Autocorrelation can be removed by Auto-Regressive Integrated Moving Average or Independent Time Series Analysis modeling [49]. Another challenge is achieving sufficient baseline stability, which includes enough data points for precise estimates [52]. Third, the duration of the study should be sufficiently long to monitor external variations without interference of other influences, such as seasonal effects [49]. Routinely collected data are an efficient means to establish a stable baseline over an extensive time period and this may even reduce data collection costs [49].

Interrupted time series design

In the interrupted time series design, a series of measurements is performed before and after implementation of the intervention at population level in order to detect whether the intervention has a significantly greater effect than the underlying secular trend, such as an economic, market or demographic trend (eg, the change in average body height of a population over time) [53]. Whether the intervention had a significantly larger effect than any underlying trend is estimated by comparing the trend in the outcome after the intervention to the trend in the pre-intervention period [54, 55]. Since randomization is not a prerequisite in this design, challenge 2 (*The organization objects to random assignment of persons or departments*) does not apply. The design is particularly relevant when using routinely collected data, such as workers' medical examinations, income insurance data, or workers' compensation data [26].

Farina and colleagues investigated the impact of national legislation on minimum safety and health requirements in 1999 on injuries at construction sites

[56]. Total and serious injury rates in the construction sector were calculated from 1994–2005, based on an integrated database (ie, Work History Italian Panel Salute). By applying segmented regression models that take into account secular trends and correct for any autocorrelation between the single observations, the results showed that the injury rates (per 10 000 weeks worked) decreased by 0.21 (95% CI -0.41– -0.01) per year more after the intervention than in the period before.

The main methodological concerns in applying the interrupted time series design for interventions are determining both the number of measurements before and after the intervention and the necessary time lags between measurements (eg, monthly or yearly data of sickness absence) to detect autocorrelations or secular trends [26, 57]. Being able to determine specifically at what time point the intervention started is a precondition for applying the interrupted time series design [58].

Differences-in-differences

Differences-in-differences methods are common practice in economics to evaluate and interpret the effect of an inevitable change (eg, policy measure). In this design, observational data are used to compare the change in the outcome of a certain group that is subjected to an intervention at a specific time point to a change in the outcome in a group that is not exposed to this intervention [59]. The method relies on finding a naturally occurring control group that mimics the properties of the intervention group and is therefore expected to follow the same time trend on the outcome as the intervention group would have in absence of the intervention [60]. This design does not necessarily require measurements for the same individuals in each group over time, since repeated cross-sectional surveys can also be used [61]. The intervention effect is calculated by subtracting the average change over time in the outcome variable in the control group from the average change in the intervention group. The design is thus an elegant way to study the internal or external changes that were named challenges earlier (*challenge 5 and 6*).

The differences-in-differences approach was applied to study the impact of a quality improvement intervention on reducing work disability, disability days, and disability and medical costs [62]. The intervention firstly provided financial incentives to 512 health providers for faster adoption of occupational health best practices, and secondly focused on improvement of care coordination and disability management at patient level. A control group of 2297 providers with the same characteristics as the intervention group was constructed. Two cross sections of data were made, which included 33 910 workers' compensation claims in the baseline period (15 408 and 18 502 for the intervention and control groups, respectively) and 71 696 (31 520 and 40 176 in the intervention and control groups, respectively) claims during the follow-up period. Patients of the providers in the intervention group were significantly less likely to be off work after one year, leading to a reduction in disability days, and lower disability and medical costs.

As with the multiple baseline design and the interrupted time series design, the main methodological concern in this approach is the autocorrelation of the outcome [63]. To deal with this issue, Bertrand and colleagues recommended conducting quite sophisticated analyses, such as bootstrap techniques, when the number of groups is sufficiently large [63]. Also, the differences-in-differences approach does not account for invariant factors and macro trends in one or both groups that might interfere with the outcome. Lastly, at the individual level, the impact of an intervention can be under- or overestimated due to unobserved, temporary and individual-specific events [60].

Regression discontinuity

The regression discontinuity design has been well established in economics over the last two decades, but not often applied in epidemiological studies. This design exploits a threshold or “cut-off” in a continuous variable used to assign treatment or intervention, and implies that individual whose assignment values lies “just above” or “just below” this threshold belong to the same population [64, 65] and thus can be compared to each other. The causal effects can be estimated by comparing the outcome between the two groups [66], assuming that subjects are not able to manipulate the threshold value. Hence, challenge 1, 2, and 3 concerning randomization and control group are minimized.

The causal effect of extending unemployment benefit duration on unemployment duration and post-unemployment outcomes was estimated in a regression discontinuity design [67]. A sharp discontinuity for age could be used, since the maximum duration of unemployment benefits increases from 12 to 18 months at the age of 45. Age was considered the threshold value, ie, the assignment variable. The study population consisted of 3432 men (44–46 years) and 3784 women (43.5–46.5 years) who were unemployed in the period from 2001–2003. By including a dummy for being exposed (ie, being >45 years old), the exit rates from employment and unemployment in the group aged >45 years were compared to the exit rates from those in the control group. The hazard rates showed that a shorter duration of unemployment benefit was associated with a higher probability of entering paid employment.

The regression discontinuity design is only appropriate when treatment is applied to a strictly defined rule, linked to a continuously measured variable (such as duration of unemployment benefit in the example above) [66]. The assumption that individuals around the threshold are similar is often debatable [64]. Other important factors to consider when applying this design are the possibility of change over time in the assignment variable and the unequal distribution of missing data between the two groups. Applying this design requires larger sample sizes than an RCT to achieve sufficient statistical power [68]. The feasibility of this design can be improved by using routine clinical or administrative data [66].

Discussion

This article demonstrated the appropriateness of research designs other than the RCT for the evaluation of occupational health interventions. Studies where in these research designs have been applied successfully showed that the most fundamental research question in intervention research could be answered, ie, did change actually occur as a result of the intervention? The designs were either experimental in nature (ie, stepped wedge) or observational (ie, propensity scores, instrumental variables, multiple baseline design, interrupted time series, difference-in-difference, and regression discontinuity).

Some of the alternative designs (eg, multiple baseline design) require using more complex statistical models that may contain a relatively large number of parameters in order to account for heterogeneity across clusters. In these cases, larger sample sizes might be needed than would be the case for individually based RCT. Furthermore, in any intervention evaluation, it seems worthwhile to determine systematically how implementation influenced the results by conducting a process evaluation. A well-known implementation model for public health and community-based interventions is the RE-AIM framework, which assesses reach, efficacy, adoption, implementation, and maintenance [69]. Nielsen and Randall’s implementation model might be more helpful for organizational-level occupational health interventions since it additionally takes into account the mental models (ie, readiness for change and perception) of those involved [70].

Even though several researchers have acknowledged that conducting an RCT on a complex intervention within an occupational health context is not always preferable, the described alternative designs are not yet widely adopted in occupational health. This could be explained by unfamiliarity of researchers with the alternatives and their advantages and disadvantages compared to the RCT, or researchers feeling pressured to apply an RCT to maximize the possibility for publication. Hopefully, this article serves as a nudge for colleagues to consider alternative research designs for the evaluation of interventions. This article also aimed to provide the necessary information to decide on selecting the most appropriate design to answer the research question, with the highest level of internal and external validity possible and the lowest costs. Designs using observational data, for instance, are particularly useful for organizational interventions or policy measures with availability of sufficient administrative data allowing for a timely evaluation of the impact of such interventions. Observational designs may be especially applicable to research in dynamic work contexts characterized by eg, high turnover, organizational restructuring, or internal mobility. While the RCT is based on a fixed cohort whereby individuals are enrolled at the same time (ie, the start of the study) and followed up for a similar period, this may be difficult when conducting an RCT in organizations with a high annual turnover of personnel. Some alternative designs are based on dynamic cohorts whereby individuals can enter and leave the cohort at different times, eg, the designs based on repeated cross-sectional data (see table 2). This may be an additional advantage to consider an observational design over an RCT.

The societal trend of big data deserves to be mentioned at this point. Some have proclaimed the current period, with its digitized patient records in large databases, to be an “open information era” as a result of public institution’s and government’s increased transparency [71]. Research can benefit from the readily accessible data this “era” yields by combining large amounts of information gathered for different purposes via different devices or media (ie, big data, so called for its variety, volume, and velocity) [72]. In doing so, we can discover correlations that would not be discovered in carefully constructed evaluations, which are typically set out to test causal relations. Big data are thus especially of interest for the described alternative research designs drawing on routinely collected data.

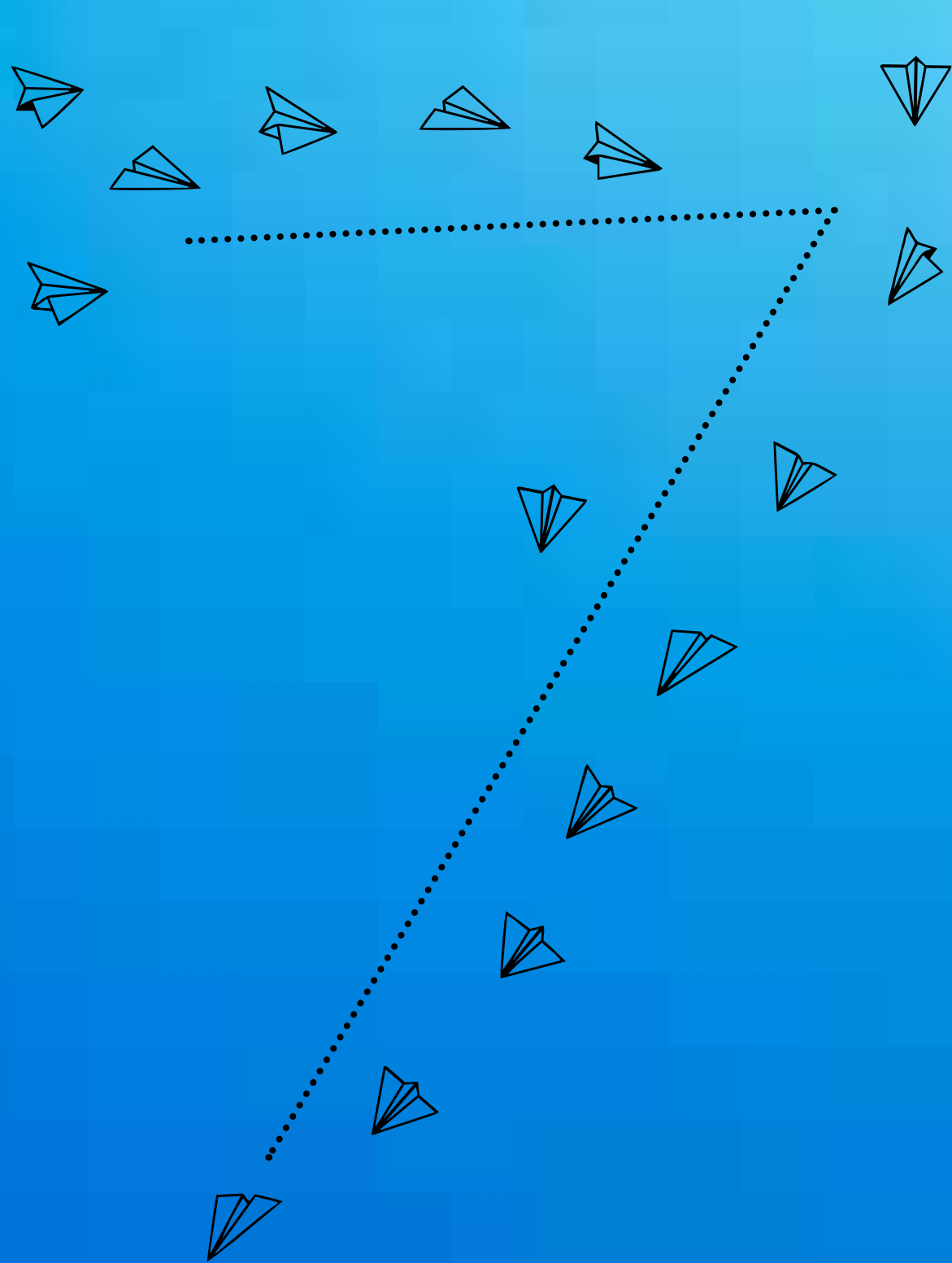
The research designs described in this article are appropriate to evaluate the effect of an intervention that is noticeable within a period of months to several years. However, the time lag between the intervention and the consequences for health can take many more years (eg, the effect of an intervention to reduce occupational exposure to dust on diseases such as silicosis and COPD). In these situations, other designs need to be considered, such as health impact assessment (HIA), which simulates the development of illness over time, based on the combined estimate of three models on: stage of the disease, the effect of exposure on stages of disease, and population characteristics. Meijster and colleagues combined a multi-stage model of respiratory problems, exposure to flour dust and allergens, and career length and influx of new workers, to estimate respiratory health outcomes of workers in the bakery sector [73]. The probability on transitioning to the next stage of disease, per unit of exposure, per year was calculated, so that incidence could be determined. The combined model demonstrated how respiratory problems develop over time and how exposure and population characteristics contributed, eg, a mean latency period of 10.3 years (95% CI 8.3–12.3) for developing respiratory symptoms in bakers was predicted [73]. Even though the RCT is still preferred as design for interventions targeted at individual level, this article provides an overview of appropriate alternatives when a group level intervention is applied, or if methodological or feasibility issues are encountered in an individual RCT that obscure the intervention-outcome relationship. The choice of the most appropriate design will be guided by the specific research question, complexity of the intervention, data available, context, and costs. Moreover, researchers conducting systematic reviews should not neglect evidence from studies applying alternative research designs. They should broaden their inclusion criteria towards observational studies with appropriate designs. When these alternative designs are applied more often, further research is necessary on the development and implementation of a guideline to improve the quality of reporting non-randomized controlled trials. We highly recommend to adopt and further explore the possibilities of both experimental alternatives and alternatives based on observational data for the evaluation of occupational health interventions.

References

- Crofton JW, Clegg JW, Mitchison DA, Hurford JV, Douglas Smith BJ, Snell WE, Forbes GB, Holt HD, Brent RH, Houghton LE, Nassau E, Ross ID, Purdie I, Santon Gilmour W, Reeve AM, McLeod JW, Nash FA, Shoulman R, Alston JM, Mohun A, Thomas ME, West LR, Tytler WH: **STREPTOMYCIN treatment of pulmonary tuberculosis. A Medical Research Council Investigation.** *Brit Med J* 1948, **2**(4582):769-782.
- Martimo KP, Shiri R, Miranda H, Ketola R, Varonen H, Viikari-Juntura E: **Effectiveness of an ergonomic intervention on the productivity of workers with upper-extremity disorders - A randomized controlled trial.** *Scand J Work Environ Health* 2010, **36**(1):25-33.
- Wolff N: **Using randomized controlled trials to evaluate socially complex services: problems, challenges and recommendations.** *J Ment Health Policy Econ* 2000, **3**(2):97-109.
- Blackwood B: **Methodological issues in evaluating complex healthcare interventions.** *J Adv Nurs* 2006, **54**(5):612-622.
- Hawe P, Shiell A, Riley T: **Important considerations for standardizing complex interventions.** *J Adv Nurs* 2008, **62**(2):267.
- Glasgow RE, Klesges LM, Dziewaltowski DA, Bull SS, Estabrooks P: **The future of health behavior change research: what is needed to improve translation of research into health promotion practice?** *Ann Behav Med* 2004, **27**(1):3-12.
- Campbell NC, Murray E, Darbyshire J, Emery J, Farmer A, Griffiths F, Guthrie B, Lester H, Wilson P, Kinmonth AL: **Designing and evaluating complex interventions to improve health care.** *Brit Med J* 2007, **334**(7591):455-459.
- Shepperd S, Lewin S, Straus S, Clarke M, Eccles MP, Fitzpatrick R, Wong G, Sheikh A: **Can we systematically review studies that evaluate complex interventions?** *PLoS Med*. 2009, **6**(8):doi: 10.1371/journal.pmed.1000086.
- Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M: **Developing and evaluating complex interventions: The new Medical Research Council guidance.** *Int J Nurs Stud* 2013, **50**(5):587-592.
- Glasgow RE, Lichtenstein E, Marcus AC: **Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition.** *Am J Public Health* 2003, **93**(8):1261-1267.
- Ioannidis JPA, Haidich A-, Pappa M, Pantazis N, Kokori SI, Tektonidou MG, Contopoulos-Ioannidis DG, Lau J: **Comparison of evidence of treatment effects in randomized and nonrandomized studies.** *JAMA-J Am Med Assoc* 2001, **286**(7):821-830.
- Campbell MK, Piaggio G, Elbourne DR, Altman DG: **Consort 2010 statement: Extension to cluster randomized trials.** *Brit Med J Online* 2012, **345**(7881).
- Puffer S, Torgerson DJ, Watson J: **Cluster randomized controlled trials.** *J Eval Clin Pract* 2005, **11**(5):479-483.
- Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, Oxman AD, Moher D: **Improving the reporting of pragmatic trials: An extension of the CONSORT statement.** *Brit Med J* 2008, **337**(7680):1223-1226.
- Concato J, Shah N, Horwitz RJ: **Randomized, controlled trials, observational studies, and the hierarchy of research designs.** *New Engl J Med* 2000, **342**(25):1887-1892.
- Last JM, Spasoff RA, Harris SS: **A Dictionary of Epidemiology.** 2000.
- Oude Hengel KM, Blatter BM, Joling CI, van der Beek AJ, Bongers PM: **Effectiveness of an intervention at construction worksites on work engagement, social support, physical workload, and need for recovery: results from a cluster randomized controlled trial.** *BMC Public Health* 2012, **12**.
- Kristensen TS: **Intervention studies in occupational epidemiology.** *Occup Environ Med* 2005, **62**(3):205-210.
- Nielsen K, Taris TW, Cox T: **The future of organizational interventions: Addressing the challenges of today's organizations.** *Work Stress* 2010, **24**(3):219-233.
- Bonell C, Fletcher A, Morton M, Lorenc T, Moore L: **Realist randomised controlled trials: A new approach to evaluating complex public health interventions.** *Soc Sci Med* 2012, **75**(12):2299-2306.

21. Griffiths A: **Organizational interventions. Facing the limits of the natural science paradigm.** *Scand J Work Environ Health* 1999, **25**(6):589-596.
22. West SG, Duan N, Pequegnat W, Gaist P, Des Jarlais DC, Holtgrave D, Szapocznik J, Fishbein M, Rapkin B, Clatts M, Mullen PD: **Alternatives to the randomized controlled trial.** *Am J Public Health* 2008, **98**(8):1359-1366.
23. Hussey MA, Hughes JP: **Design and analysis of stepped wedge cluster randomized trials.** *Contemp Clin Trials*. 2007, **28**(2):182-191.
24. Brown CA, Lilford RJ: **The stepped wedge trial design: A systematic review.** *BMC Med Res Methodol* 2006, **6**:doi: 10.1186/1471-2288-6-54.
25. Mdege ND, Man M-, Taylor CA, Torgerson DJ: **Systematic review of stepped wedge cluster randomized trials shows that design is particularly used to evaluate interventions during routine implementation.** *J Clin Epidemiol* 2011, **64**(9):936-948.
26. Sanson-Fisher RW, D'Este CA, Carey ML, Noble N, Paul CL: **Evaluation of systems-oriented public health interventions: Alternative research designs.** *Annu Rev Publ Health* 2014, **35**:9-27.
27. Zwijsen SA, Gerritsen DL, Eefsting JA, Smalbrugge M, Hertogh CM, Pot AM: **Coming to grips with challenging behaviour: a cluster randomised controlled trial on the effects of a new care programme for challenging behaviour on burnout, job satisfaction and job demands of care staff on dementia special care units.** *Int J Nurs Stud* 2015, **52**(1):68-74.
28. Kotz D, Spigt M, Arts ICW, Crutzen R, Viechtbauer W: **Researchers should convince policy makers to perform a classic cluster randomized controlled trial instead of a stepped wedge design when an intervention is rolled out.** *J Clin Epidemiol* 2012, **65**(12):1255-1256.
29. Kotz D, Spigt M, Arts ICW, Crutzen R, Viechtbauer W: **Use of the stepped wedge design cannot be recommended: A critical appraisal and comparison with the classic cluster randomized controlled trial design.** *J Clin Epidemiol* 2012, **65**(12):1249-1252.
30. D'Agostino RB, Jr: **Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group.** *Stat Med* 1998, **17**(19):2265-2281.
31. Shadish WR, Luellen JK, Clark MH: **Propensity Scores and Quasi-Experiments: A Testimony to the Practical Side of Lee Sechrest.** In *Strengthening research methodology: Psychological measurement and evaluation.* Edited by Bootzin RR, McKnight PE. Washington, DC, US: American Psychological Association; 2006:143-157.
32. Rosenbaum PR, Rubin DB: **The central role of the propensity score in observational studies for causal effects.** *Biometrika* 1983, **70**(1):41-55.
33. Luellen JK, Shadish WR, Clark MH: **Propensity scores: an introduction and experimental test.** *Evaluation Rev* 2005, **29**(6):530-558.
34. Austin PC: **An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies.** *Multivar Behav Res* 2011, **46**(3):399-424.
35. Cousens S, Hargreaves J, Bonell C, Armstrong B, Thomas J, Kirkwood BR, Hayes R: **Alternatives to randomisation in the evaluation of public-health interventions: statistical analysis and causal inference.** *J Epidemiol Commun H* 2011, **65**(7):576-581.
36. Schafer JL, Kang J: **Average causal effects from nonrandomized studies: a practical guide and simulated example.** *Psychol Methods* 2008, **13**(4):279-313.
37. Saltychev M, Laimi K, Oksanen T, Pentti J, Kivimäki M, Vahtera J: **Does perceived work ability improve after a multidisciplinary preventive program in a population with no severe medical problems? The Finnish Public Sector Study.** *Scand J Work Environ Health* 2013, **39**(1):57-65.
38. Saltychev M, Laimi K, El-Metwally A, Oksanen T, Pentti J, Virtanen M, Kivimäki M, Vahtera J: **Effectiveness of multidisciplinary primary prevention in decreasing the risk of work disability in a low-risk population.** *Scand J Work Env Hea* 2012, **38**(1):27-37.
39. Austin PC: **Double propensity-score adjustment: A solution to design bias or bias due to incomplete matching.** *Stat Methods Med Res* 2014.:doi:0.1177/0962280214543508.
40. Steiner PM, Cook TD, Shadish WR: **On the Importance of Reliable Covariate Measurement in Selection Bias Adjustments Using Propensity Scores.** *J Educ Behav Stat* 2011, **36**(2):213-236.
41. Steiner PM, Cook TD, Shadish WR, Clark MH: **The importance of covariate selection in controlling for selection bias in observational studies.** *Psychol Methods* 2010, **15**(3):250-267.
42. Newhouse JP, McClellan M: **Econometrics in outcomes research: the use of instrumental variables.** *Annu Rev Publ Health* 1998, **19**:17-34.
43. Bennett DA: **An introduction to instrumental variables analysis: part 1.** *Neuroepidemiology* 2010, **35**(3):237-240.
44. Rassen JA, Brookhart MA, Glynn RJ, Mittelman MA, Schneeweiss S: **Instrumental variables I: instrumental variables exploit natural variation in nonexperimental data to estimate causal relationships.** *J Clin Epidemiol* 2009, **62**(12):1226-1232.
45. Elovainio M, Heponiemi T, Kuusio H, Jokela M, Aalto A, Pekkarinen L, Noro A, Finne-Soveri H, Kivimäki M, Sinervo T: **Job demands and job strain as risk factors for employee wellbeing in elderly care: an instrumental-variables analysis.** *Eur J Publ Health* 2014, **1**:1-6.
46. Behncke S: **Does retirement trigger ill health?** *Health Econ* 2012, **21**(3):282-300.
47. Martens EP, Pestman WR, de Boer A, Belitser SV, Klungel OH: **Instrumental variables: application and limitations.** *Epidemiology* 2006, **17**(3):260-267.
48. Moeyaert M, Ugille M, Ferron JM, Beretvas SN, Van den Noortgate W: **Modeling external events in the three-level analysis of multiple-baseline across-participants designs: a simulation study.** *Behav Res Methods* 2013, **45**(2):547-559.
49. Hawkins NG, Sanson-Fisher RW, Shakeshaft A, D'Este C, Green LW: **The multiple baseline design for evaluating population-based research.** *Am J Prev Med* 2007, **33**(2):162-168.
50. Rhoda DA, Murray DM, Andridge RR, Pennell ML, Hade EM: **Studies with staggered starts: multiple baseline designs and group-randomized trials.** *Am J Public Health* 2011, **101**(11):2164-2169.
51. Camden M, Ludwig T: **Absenteeism in Health Care: Using Interlocking Behavioral Contingency Feedback to Increase Attendance With Certified Nursing Assistants.** *Journal of Organizational Behavior Management* 2013, **33**(3):165-184.
52. Cuvo AJ: **Multiple-baseline design in instructional research: pitfalls of measurement and procedural advantages.** *Am J Ment Def* 1979, **84**(3):219-228.
53. Biglan A, Ary D, Wagenaar AC: **The value of interrupted time-series experiments for community intervention research.** *Prevention Science* 2000, **1**(1):31-49.
54. Fretheim A, Soumerai SB, Zhang F, Oxman AD, Ross-Degnan D: **Interrupted time-series analysis yielded an effect estimate concordant with the cluster-randomized controlled trial result.** *J Clin Epidemiol* 2013, **66**(8):883-887.
55. Zhang F, Wagner AK, Soumerai SB, Ross-Degnan D: **Methods for estimating confidence intervals in interrupted time series analyses of health interventions.** *J Clin Epidemiol* 2009, **62**(2):143-148.
56. Farina E, Bena A, Pasqualini O, Costa G: **Are regulations effective in reducing construction injuries? An analysis of the Italian context.** *Occup Environ Med* 2013, **70**(9):611-616.
57. Ramsay CR, Matowe L, Grilli R, Grimshaw JM, Thomas RE: **Interrupted time series designs in health technology assessment: Lessons from two systematic reviews of behavior change strategies.** *Int J Technol Assess* 2003, **19**(4):613-623.
58. Matowe LK, Leister CA, Crivera C, Korth-Bradley JM: **Interrupted time series analysis in clinical research.** *Ann Pharmacother* 2003, **37**(7-8):1110-1116.
59. Angrist J, Krueger A: **Empirical strategies in Labor Economics.** In *Handbook in Labor Economics.* Edited by Ashenfelter O, Card D. Elsevier; 2000:1277-1366.
60. Blundell R, Costa Dias M: **Alternative Approaches to Evaluation in Empirical Microeconomics.** *J Hum Resour* 2009, **44**:565-640.
61. Blundell R, MaCurdy T: **Labor Supply.** In *Handbook of Labor Economics.* Edited by Ashenfelter O, Card D. Elsevier; 2000:1559-1695.
62. Wickizer TM, Franklin G, Fulton-Kehoe D, Gluck J, Mootz R, Smith-Weller T, Plaeger-Brockway R: **Improving quality, preventing disability and reducing costs in workers' compensation healthcare: a population-based intervention study.** *Med Care* 2011, **49**(12):1105-1111.
63. Bertrand B, Duflo E, Mullainathan S: **How much should we trust difference-in-difference estimates?** *The Quarterly Journal of Economics* 2004, **119**(1):249-275.

64. O'Keefe AG, Geneletti S, Baio G, Sharples LD, Nazareth I, Petersen I: **Regression discontinuity designs: an approach to the evaluation of treatment efficacy in primary care using observational data.** *Brit Med J* 2014, **349**:g5293.
65. Moscoe E, Bor J, Barnighausen T: **Regression discontinuity designs are underutilized in medicine, epidemiology, and public health: A review of current and best practice.** *J Clin Epidemiol* 2015, **68**(2):122-133.
66. Bor J, Moscoe E, Mutevedzi P, Newell ML, Barnighausen T: **Regression discontinuity designs in epidemiology: causal inference without randomized trials.** *Epidemiology* 2014, **25**(5):729-737.
67. Caliendo M, Tatsiramos K, Uhlendorff A: **Benefit duration, unemployment duration and job match quality: a regression-discontinuity approach.** *J Appl Econom* 2013, **28**:604-627.
68. Pennell ML, Hade EM, Murray DM, Rhoda DA: **Cutoff designs for community-based intervention studies.** *Stat Med* 2011, **30**(15):1865-1882.
69. Glasgow RE, Vogt TM, Boles SM: **Evaluating the public health impact of health promotion interventions: the RE-AIM framework.** *Am J Public Health* 1999, **89**(9):1322-1327.
70. Nielsen K, Randall R: **Opening the black box: Presenting a model for evaluating organizational-level interventions.** *European Journal of Work and Organizational Psychology* 2013, **22**(5):601-617.
71. Groves P, Kayyali B, Knott D, Van Kuijen S: **The 'big data' revolution in healthcare: Accelerating value and innovation.** January 2013.
72. Mooney SJ, Westreich DJ, El-Sayed AM: **Commentary: epidemiology in the era of big data.** *Epidemiology* 2015, **26**(3):390-394.
73. Meijster T, Tielemans E, Heederik D: **Effect of an intervention aimed at reducing the risk of allergic respiratory disease in bakers: Change in flour dust and fungal alpha-amylase levels.** *Occup Environ Med* 2009, **66**(8):543-549.
74. Smedley J, Harris EC, Cox V, Ntani G, Coggon D: **Evaluation of a case management service to reduce sickness absence.** *Occup Med* 2013, **63**(2):89-95.
75. Kobayashi Y, Kaneyoshi A, Yokota A, Kawakami N: **Effects of a worker participatory program for improving work environments on job stressors and mental health among workers: a controlled trial.** *J Occup Health* 2008, **50**(6):455-470.
76. Frykman M, Hasson H, Muntlin Athlin A, Von Thiele Schwarz U: **Functions of behavior change interventions when implementing multi-professional teamwork at an emergency department: A comparative case study.** *BMC Health Services Research* 2014, **14**(1).
77. Steenhuis I, van Assema P, Reubsaet A, Kok G: **Process evaluation of two environmental nutrition programmes and an educational nutrition programme conducted at supermarkets and worksite cafeterias in the Netherlands.** *J Hum Nutr Diet* 2004, **17**(2):107-115.
78. Oude Hengel KM, Blatter BM, Van Der Molen HF, Joling CI, Proper KI, Bongers PM, Van Der Beek AJ: **Meeting the challenges of implementing an intervention to promote work ability and health-related quality of life at construction worksites: A process evaluation.** *J Occup Environ Med* 2011, **53**(12):1483-1491.
79. van Deurssen EH, Pronk A, Meijster T, Tielemans E, Heederik D, Oude Hengel KM: **Process evaluation of an intervention program to reduce occupational quartz exposure among Dutch construction workers.** *J Occup Environ Med* 2015, **57**(4):428-435.



Process variables in organizational stress management intervention evaluation research: a systematic review

Bo M. Havermans
Roosmarijn M.C. Schelvis
Cécile R.L. Boot
Evelien P.M. Brouwers
Johannes R. Anema
Allard J. van der Beek

Abstract

Objectives

This systematic review aimed to explore which process variables are used in stress management intervention (SMI) evaluation research.

Methods

A systematic review was conducted using seven electronic databases. Studies were included if they reported on an SMI aimed at primary or secondary stress prevention, were directed at paid employees, and reported process data. Two independent researchers checked all records and selected the articles for inclusion. Nielsen and Randall's model for process evaluation was used to cluster the process variables. The three main clusters were context, intervention, and mental models.

Results

In the 44 articles included, 47 process variables were found, clustered into three main categories: context (two variables), intervention (31 variables), and mental models (14 variables). Half of the articles contained no reference to process evaluation literature. The collection of process evaluation data mostly took place after the intervention and at the level of the employee.

Conclusions

The findings suggest that there is great heterogeneity in methods and process variables used in process evaluations of SMI. This, together with the lack of use of a standardized framework for evaluation, hinders the advancement of process evaluation theory development.

Introduction

Work stress is a problem for individuals, organizations, and society at large. It poses a threat to workers' well-being by increasing mental and physical health risks [1, 2]. Work stress also contributes substantially to sickness absence [3]. This is a costly problem for organizations and society in general. The annual price tag of work stress to society amounts to €20 billion in the European Union alone [4]. In order to combat this problem, organizations deploy stress management interventions (SMI). Scientific evaluations of these SMI can support organizations in making an informed choice about the most effective and appropriate intervention and may also help to test theories upon which the interventions are based.

The most widely used approach to SMI evaluation is characterized by a (quasi-) experimental research design that focuses on outcomes at the level of the worker (eg, stress, burnout) [5–7]. According to Kristensen [8], whether or not the intervention has had the desired effect on the targeted outcome is only one of three important questions to ask when evaluating an intervention. To interpret the effect, one should firstly assess if the intervention was carried out as intended and then assess if the intervention brought about the intended (change in) exposure or behavior. This way, a distinction can be made between program versus theory success in effect interpretation.

A way to gather information about the success or failure of an intervention program is to look at intervention implementation. This can be done by studying process variables [9–12]. There are different ways of investigating the implementation process. Steckler and Linnan [9], for instance, propose a focus on intervention delivery and participation. Fleuren et al [13] assert that components, such as the sociopolitical context, and characteristics of the organization, participant (skills, knowledge, and perceived support), and the intervention itself (complexity, relative advantage) are also important for implementation. Finally, Nielsen and Randall [11] suggest that mental models (pertaining to constructs such as readiness for change) should be added to existing process evaluation frameworks.

Despite increasing support for the incorporation of process factors into the evaluation of SMI in the last 10–15 years, there is still limited consensus on which process variables should be assessed. In addition to frameworks that offer suggestions for the use of certain process variables, insight into current practice could also support future process evaluations of SMI. More overviews that stress the importance of process measures in organizational-level intervention evaluations do exist. Egan and colleagues [14], for example, provide a review of implementation appraisal of complex social interventions, concluding that implementation and context are crucial for impact assessment of these interventions. To the authors' best knowledge, a decade ago, Murta et al [15] have provided the only review describing which process variables are used in SMI evaluation research. In accordance with the aforementioned different perspectives on process evaluation, they observed great heterogeneity in variables and designs researchers use for SMI process evaluation. Murta and col-

leagues made this observation using a restricted selection of process variables [15]. A limitation of this restricted selection is that publications reporting other process variables could have been neglected. Building on their research, a broader approach can leave room for more current frameworks to be recognized in process evaluation practice. The aim of this review was to explore which process variables are used in SMI evaluation research.

Methods

A systematic literature review was performed to investigate which process variables are reported in SMI evaluation research. Components from the PRISMA statement [16] were used in reporting this systematic review.

Search and study selection

Studies were eligible for inclusion if they (i) reported on an SMI directed at paid workers aged ≥ 18 years, (ii) reported a process evaluation of the intervention (at least one process variable assessed), (iii) were published in a peer-reviewed journal (conference abstracts, books and design protocols were excluded), and (iv) were written in English or Dutch. An SMI was defined as an organizational intervention focusing on individual or organizational changes, targeted to prevent or reduce stress in employees at the primary or secondary prevention level. A process variable was defined as any measure included in the evaluation study that is hypothesized to be associated with the process of SMI implementation.

Together with a library search specialist, the following databases were searched from inception to October–December 2014: PubMed, PsychINFO (October 8, 2014), ISI/Web of Science, Embase (October 24, 2014), Proquest (December 3, 2014), EconLit (December 5, 2014), and Ebsco/Cinahl (December 11, 2014). For every database, the search was adapted to the appropriate terminology specific to that database, using synonyms and closely related words (for the complete search, see the Appendix, www.sjweh.fi/data_repository.php). If a process evaluation was mentioned in design protocols then the first author searched the electronic literature databases and contacted the authors to identify additional studies.

The first author removed the duplicates from the records identified. Then, the first and the second author independently screened titles and abstracts of all remaining records, selecting articles for fulltext inspection, using the aforementioned eligibility criteria. If at least one of the two authors had selected a record for fulltext inspection then it was retrieved. Subsequently, both authors independently assessed the remaining selection of articles for inclusion. There was an independent consensus for in- and exclusion of fulltext articles of 72%. Remaining discrepancies were resolved with face-to-face deliberation. When this did not lead to consensus, one of the co-authors was consulted to make a final decision.

Data extraction

A template was constructed, containing a list of data to be extracted from the included articles. This template was used independently by the first and second author by applying it to two, randomly selected articles. Then, they compared the data they had extracted and modified and further specified the template towards consensus. Random selection and coding of studies by both the first and second author was repeated for 20% of all articles after which a clear coding format was obtained.

The template used for data extraction contained three main component categories: (i) study and intervention characteristics, (ii) process evaluation methods, and (iii) process variables. Intervention characteristics were adapted from Murta and colleagues [15]. The process evaluation methods components were adapted from Wierenga and colleagues [17]. The specific components are listed in table 1. Process variables were coded using a list of concepts derived from process evaluation literature [9, 11, 13, 15, 17]. During coding, the researchers used the list of concepts as a frame of reference, as a starting point. When necessary, the researchers diverged from this list so as not to exclude variables that were not on the list but were used as process variables. Data were collected at the level of the employee (micro level), the level of the supervisor, manager, or department (meso level), and at the level of the CEO, organization, or sector (macro level). For every process variable, it was assessed how many articles reported data collected at the micro, meso, or macro level.

Analyses

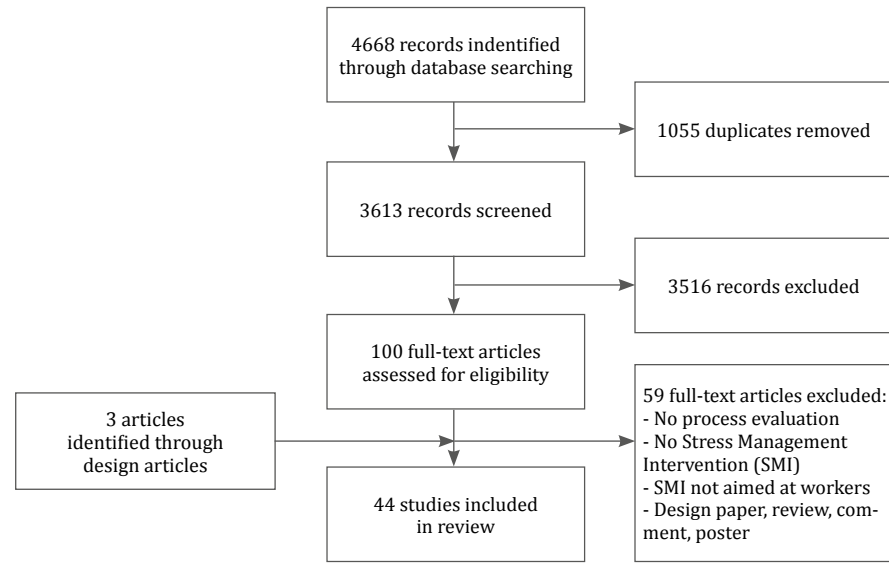
The Nielsen and Randall model for process evaluation [11] was used to cluster the process variables retrieved because the model was developed especially for organizational-level interventions and it provides the opportunity to take a broad perspective to process evaluation. Using this model, mediating and moderating factors of implementation can be detected. It was deliberated under which cluster a process variable should go, first with two of the co-author and then with all authors of this review. After each deliberation, the arrangement was adjusted according to the feedback received. The main clusters were context, intervention, and mental models. The content of the three clusters is in accordance with the three central themes of the Nielsen and Randall model for process evaluation [11]. Context pertains to situational aspects that affect organizational behavior and functional relationships between variables, and contains hindering and facilitating factors. The intervention cluster refers to aspects of intervention design and implementation that determine the maximum levels of intervention exposure that can be reached, and contains the sub-clusters initiation, intervention activities, implementation, and implementation strategies. The mental models cluster refers to underlying psychological aspects that may help explain stakeholders' behavior in and reaction to the intervention. The mental models cluster contains the sub-clusters readiness for change, perceptions, and changes in mental models.

Results

Study selection

The database search identified 4668 records. After duplicates were removed, an initial screening of titles and abstracts of the 3613 remaining records produced 100 potentially relevant publications. After screening of the 100 retrieved fulltext articles, 59 were excluded. The three main reasons for exclusion were: no process evaluation, no SMI, no results presented (design or protocol paper). Additionally, three eligible articles were identified through design papers. Finally, 44 articles met the selection criteria, and were included in this review (Figure 1). Two studies were reported in more than one of the included articles, so 42 studies are represented in this review.

Figure 1 — Flow Chart of Study Selection Process



Study and intervention characteristics

Table 1 presents the most important characteristics of the studies in the context of this review. Of the 42 reported studies, 27 were conducted in Europe (64%), 10 in North America (24%), 4 in Australia (10%), and 1 in Asia (2%). Most studies were conducted in the healthcare sector (45%), followed by education (13%). Of the 42 studies, 9 (21%) were conducted in a mixed set of organizations in more than one sector. More than half of the interventions (55%) had a participatory format. A participatory approach is characterized by cooperation of different stakeholders (eg, employees, managers, intervention providers) in the assessment, targeting, and prevention of work stress. Intervention duration ranged from 1–312 weeks, with most intervention durations (64%) not exceeding one year.

Table 1 — Studies reporting process variables [CR=company records; DC=data collection; N=number of intervention participants; ON=observation notes; PA=participatory approach; PE=reference made to process evaluation literature; Pre/during/post=before/during/after the intervention; Ref=reference to literature]

Author	Country	Sector	N	PA ^a	PE	DC moment	Level of DC ^b	DC ^c	Method of DC
Albertsen et al, 2014 (35)	Denmark	Healthcare	840	No	No	Pre, post	Micro, meso	Qualitative	Focus groups, interviews
Aust et al, 2010 (26)	Denmark	Healthcare	399	Yes	No	During, post	Micro, meso	Combined	ON, questionnaire, reports
Biron et al, 2010 (29)	UK	Utilities	205	Yes	Yes	Pre, during, post	Micro, meso	Combined	Interviews, ON
Bourbonnais et al, 2006 (27)	Canada	Healthcare	492	Yes	No	During	Micro, meso	Qualitative	Interviews, ON
Bruneau et al, 2004 (36)	UK	Healthcare	18	No	No	During	Micro	Quantitative	Questionnaire
Bunce et al, 1996 (37)	UK	Healthcare	118	No	Yes	Post	Micro	Quantitative	Questionnaire
Burton et al, 2010 (38)	Australia	Education	18	No	No	During, post	Micro	Combined	ON, questionnaire, CR
Coffeng et al, 2013 (18)	NL	Financial	306	Yes	Yes	Post	Micro, meso	Combined	ON, questionnaire, CR
Cohen-Katz et al, 2005 (39)	USA	Healthcare	25	No	No	During, post	Micro, macro	Combined	Focus groups, interviews, ON, questionnaire, CR
Falck et al, 1984 (40)	USA	Education	32	No	Yes	During	Micro, meso, macro	Combined	Conference, ON, questionnaire
Gümüşen et al, 2009 (23)	Turkey	Healthcare	72	No	No	Post	Micro	Qualitative	Interviews
Hasson et al, 2014 (20)	Sweden	Mixed	- ^d	No	Yes	During	Meso, macro	Qualitative	Interviews
Hasson et al, 2014 (41)	Canada	Financial	1714	No	Yes	Post	Micro	Quantitative	Questionnaire
Heaney et al, 1993 (42)	USA	Manu-facturing	176	Yes	No	During, post	Micro, meso, macro	Combined	Interviews, ON, questionnaire
Helms-Lorenz et al, 2013 (43)	NL	Education	192	Yes	No	Post	Micro	Quantitative	Questionnaire
Ipsen et al, 2014 (19)	Denmark	Mixed	129	Yes	Yes	During, post	Micro, meso, macro	Combined	Interviews, ON, questionnaire, workshop
Jeffcoat et al, 2012 (22)	USA	Education	121	No	No	During, post	Micro	Quantitative	Questionnaire, CR
Jenny et al, 2014 (34)	Switzer-land	Mixed	3532	Yes	Yes	During, post	Micro, meso	Combined	Interviews, questionnaire, CR
Kawai et al, 2010 (44)	Japan	Mixed	168	No	Yes	During	Micro	Quantitative	Questionnaire
Keller et al, 2012 (45)	USA	Healthcare	60	No	No	Pre, during, post	Micro	Quantitative	Questionnaire
Landsbergis et al, 1995 (46)	USA	Healthcare	63	Yes	No	Post	Micro	Qualitative	Interviews

Author(s) and Year(s)	Country	Police	30	No	No	Post	Micro	Qualitative	Interviews
Leonard et al, 1999 (47)	Australia	Police	30	No	No	Post	Micro	Qualitative	Interviews
Mellor et al, 2011 (48)	UK	Mixed	- ^d	Yes	No	During, post	Micro, meso, macro	Combined	Interviews, CR
Mellor et al, 2013 (49)	UK	Mixed	- ^d	Yes	No	During	Meso	Combined	Interviews, CR
Millar et al, 2008 (50)	Australia	Utilities	28	No	Yes	Post	Micro, meso	Combined	Questionnaire, records
Nielsen et al, 2006 (51)	Denmark	Healthcare	144	Yes	Yes	During, post	Micro, meso	Combined	Focus groups, group meetings, interviews, ON, reports, questionnaire
Nielsen et al, 2007 (52)	Denmark	Mixed	538	Yes	Yes	During, post	Micro, meso	Quantitative	Questionnaire
Nielsen et al, 2012 (53)	Denmark	Healthcare	583	No	Yes	Post	Micro	Quantitative	Questionnaire
Pettersson et al, 1998 (54)	Sweden	Healthcare	3506	Yes	No	- ^d	Micro, meso	Quantitative	Questionnaire
Randall et al, 2007 (24)	UK	Healthcare	- ^f	Yes	Yes	Post	Micro	Qualitative	Interviews
Randall et al, 2009 (31)	Denmark	Healthcare	551	No	Yes	Post	Micro, meso	Qualitative	Focus groups, interviews
Renaud et al, 2008 (21)	Canada	Financial	656	No	No	Post	Micro, meso	Combined	Interviews, questionnaire
Reynolds et al, 1993 (55)	UK	Healthcare	92	No	Yes	During, post	Micro	Quantitative	Questionnaire
Reynolds et al, 1993 (56)	UK	Healthcare	92	No	Yes	During, post	Micro	Quantitative	Questionnaire
Saksvik et al, 2002 (25)	Norway	Mixed	685	No	Yes	Pre, post	Micro, meso	Qualitative	Interviews, ON, reports
Schwerman et al, 2012 (57)	USA	Healthcare	3930	No	No	Post	- ^d	Quantitative	Questionnaire
Sorensen et al, 2014 (28)	Denmark	Mixed	163	Yes	Yes	During, post	Micro, meso, macro	Combined	Interviews, ON, questionnaire, CR
Swain et al, 2014 (58)	NZ	Healthcare	56	No	No	Post	Micro	Qualitative	Questionnaire
Van Bogaert et al, 2014 (59)	Belgium	Education	170	Yes	No	Post	Micro	Quantitative	Questionnaire
Van Wingerden et al, 2013 (60)	NL	Education	50	No	No	During, post	Micro	Qualitative	Interviews
Weigl et al, 2012 (32)	Germany	Healthcare	17	Yes	No	Post	Micro	Qualitative	Interviews
Weigl et al, 2013 (61)	Germany	Healthcare	19	Yes	Yes	Post	Micro	Qualitative	Interviews
Westlander et al, 1995 (62)	Sweden	Telecom	300	Yes	Yes	During, post	Micro, meso, macro	Combined	Group meetings, interviews, questionnaire, CR
Van Berkel et al, 2013 (63)	NL	Research	257	No	Yes	Pre, during, post	Micro, meso	Combined	Interviews, questionnaire

^a Cooperation of different stakeholders in the assessment, targeting, and prevention of work stress.

^b Micro=employee, Meso=supervisor/manager/department, Macro=CEO/organization/sector.

^c Quantitative=questionnaires/company records (CR); qualitative=conference/focus groups/group meetings/interviews/observation notes (ON)/reports/workshop; Com-

bined=Combination of quantitative and qualitative methods.

^d Not specified.

Half (50%) of the 44 articles did not contain any reference to process evaluation literature in the introduction or methods section. In 20 articles, process evaluation data were collected at more than one moment. Collection of process evaluation data mostly took place post (84%) or during the intervention (55%). In five cases, process evaluation data were collected pre-intervention (11%). In most articles (93%), process evaluation data were collected at the micro level. In 22 and 8 articles, process evaluation data were collected at the meso and macro level, respectively. All articles that reported only quantitative data for process evaluation used a questionnaire for the process. In the articles that reported a qualitative or a combined approach, (group) interviews were mostly used for process evaluation.

Process variables

Table 2 shows all 47 process variables that were retrieved. Some of the most striking findings are discussed below. The context cluster contained 2 process variables, the intervention cluster 31, and the mental models cluster 14. The intervention sub-cluster initiation contained 3 process variables, intervention activities 8, implementation 8, and implementation strategy 12. The mental models sub-cluster readiness for change contained 4 process variables, perceptions 7, and changes in mental models 3. For every process variable, the general level of data collection is reported.

Attitudes and perceptions of intervention users were reported most frequently (30 articles), followed by support (N=24), context (N=19), information/communication (N=15), and effectiveness beliefs (N=14).

Clusters

Context

Both a cluster and process variable, context was the third most reported variable. Coffeng and colleagues [18], for instance, reported a reorganization at the beginning of the intervention period as an example of context. Another contextual factor they reported was the fact that three months before the intervention project started, another intervention to improve the work environment had been piloted. Of 19 articles reporting context, 18 reported data that were collected at least at the micro level. The other process variable in this cluster was barriers and facilitators, which was reported six times. Ipsen et al [19] gave an example of both: making the wrong changes slows the process (barrier) and the intervention constitutes a collective process (facilitator).

Intervention

The first sub-cluster of the intervention cluster was initiation, which contains the process variables motivation, initiation of the intervention, and ownership. Motivation was reported in 7 articles, all of which collected data at least at the meso level. In the second sub-cluster, intervention activities, responsibility

Table 2 — Process variables reported in the included articles

Clusters and concepts	Concept descriptions ^a	Studies		Level of data collection ^b			References
		(N)	(N)	Micro	Meso	Macro	
<i>Context</i>							
Context	Contextual factors affecting the implementation of intervention	19	18	13	6		(18–20, 23–25, 27–29, 34, 35, 37–39, 46, 48, 51, 62, 63)
Barriers/facilitators	Factors that hinder or help implementation of intervention	6	6	6	2		(18, 19, 25, 48, 51, 63)
<i>Intervention</i>							
Initiation		(N)	Micro	Meso	Macro	References	
Motivation	Motivation to use intervention	7	6	7	2		(20, 25, 26, 29, 34, 40, 63)
Initiation	Activities related to the initiation of intervention	3	2	3	3		(19, 20, 48)
Ownership	Extent to which intervention stakeholders consider themselves drivers of implementation	2	2	2	1		(19, 29)
Activities							
Responsibility	Responsibility of intervention stakeholders for implementation	10	8	10	3		(19, 20, 25, 28, 29, 34, 35, 49, 51)
Integration	Extent to which intervention is integrated into daily work processes	6	4	5	3		(19, 20, 24, 28, 34, 49)
Appropriateness	Extent to which stakeholders consider intervention appropriate for the organization	4	4	3	3		(19, 23, 28, 48)
Monitoring	Monitoring progress of intervention implementation	3	3	2	1		(19, 34, 61)
Adoption	Extent to which stakeholders adopt intervention initiatives	2	2	2	1		(28, 34)
Maintenance	Extent to which the intervention is used after implementation	2	2	1	-		(34, 59)
Comfort	Extent to which stakeholders feel comfortable with requirements	1	1	1	-		(29)
Tailoring	Extent to which intervention is tailored to user needs	1	1	1	-		(63)
Implementation							
Dose received	Extent to which users actively engage in intervention	13	12	10	4		(18–22, 26, 29, 34, 37, 38, 40, 45, 63)
Participation	Participation in intervention	12	12	8	4		(21, 24, 25, 28, 39, 42, 48, 52–54, 61)
Outcome	Extent to which stakeholders consider implementation of intervention successful	8	7	5	4		(19, 20, 29, 32, 40, 43, 61, 62)
Dose delivered	Amount of intended intervention components delivered to users	8	8	7	3		(18, 19, 24, 26, 28, 50, 51, 62)
Fidelity	Extent to which the intervention is delivered as planned	7	7	7	2		(18, 25, 40, 50, 51, 62, 63)
Reach	Proportion of target population that participates in intervention	5	5	5	-		(18, 27, 29, 34, 63)
Exposure	Extent to which intervention users are exposed to intervention	4	4	2	-		(21, 24, 31, 41)
Recruitment	Procedures used to recruit intervention users	4	4	3	1		(18, 34, 40, 43)
Implementation strategy							
Support	Intervention stakeholders' support for intervention	24	22	15	5		(19–21, 23–29, 32, 34, 35, 40, 46, 49, 51, 55, 56, 59–63)
Information / communication	Information and communication about intervention	15	13	12	7		(19, 20, 24, 27, 28, 34, 40, 42, 44, 48, 49, 52, 61–63)
Involvement	Involvement of stakeholders in intervention and/or intervention activities	11	10	7	3		(24, 25, 28, 31, 42, 48, 49, 54–56, 59)
Resources	Resources (eg, money, time, manpower) available for intervention	7	6	6	2		(19, 25, 34, 48, 49, 60, 63)
Roles	Clarity of stakeholders' roles within intervention	6	5	6	2		(19, 20, 25, 26, 35, 51)
Expertise	Extent to which users have experience with specific intervention activities (eg, risk assessment)	3	3	3	2		(28, 35, 48)
Attractiveness	Extent to which intervention materials are attractive for users	1	1	1	-		(63)
Cooperation	Extent to which intervention stakeholders work together	1	1	-	-		(32)
Independency	Extent to which the project agenda is independent	1	1	-	-		(61)
Coherence	Extent to which intervention elements are related to each other	1	1	1	-		(34)
Responsiveness	Extent to which intervention provider is responsive to stakeholders	1	1	1	-		(63)
Scale	Scale to which intervention is implemented in the organization	1	1	1	1		(28)
<i>Mental models</i>		(N)	Micro	Meso	Macro	References	
Readiness for change							
Awareness of problem	Awareness of challenges related to either work stress or implementation of the intervention	7	7	5	2		(28, 34, 48, 54–56, 63)
Readiness for change	Extent to which intervention stakeholders are ready for change	6	5	6	1		(19, 29, 31, 49, 51, 63)

Project fatigue	Feeling that organization has initiated too many surveys/projects	3	2	3	1	(25, 48, 49)
Intention to act	Intention to participate in intervention program	1	1	-	-	(44)
Perceptions						
Attitudes and perceptions	Attitudes and perceptions of users related to the intervention	30	29	18	7	(19-21, 23, 25-29, 31, 32, 34-41, 46, 48, 51, 52, 54-56, 58, 60, 62, 63)
Satisfaction	Satisfaction with the intervention	6	6	5	-	(18, 21, 26, 35, 47, 63)
Engagement	Engagement of stakeholders in intervention implementation	4	3	4	2	(20, 21, 25, 28)
Trust	Trust between stakeholders involved in intervention	3	3	2	1	(19, 23, 63)
Enjoyment	Enjoyment in intervention	2	2	1	-	(44, 50)
Enthusiasm	Enthusiasm about the intervention	1	1	1	-	(25)
Influence	Influence on the contents of the intervention	1	1	1	-	(52)

^a Stakeholder=person who is directly or indirectly involved in the intervention; User=person who directly participates in the intervention.

^b Micro=employee; Meso=supervisor/manager/department; Macro=CEO/organization/sector.

was reported most and refers to the extent to which different stakeholders are accountable for carrying out intervention actions. Hasson and colleagues [20] found that senior management differed with human resource professionals about who was responsible for involving line managers in the intervention. In the third sub-cluster, implementation, dose received was the most reported process variable, examples of which included self-reported participation in intervention modules [21] and quiz completion of participants across several quizzes during the intervention [22]. Data collection at the micro level was dominant in this sub-cluster. For the sub-cluster implementation strategy, support was the most-reported process variable, examples of which include support from management toward employees to attend intervention sessions [23] and the visibility of senior management's involvement in the intervention [24]. Also in this sub-cluster, 11 articles reported the process variable involvement, which was, for example, described as the extent to which stakeholders took part in the development of a plan of action [25].

Mental models

In mental models, the first sub-cluster was readiness for change, of which awareness of problem/intervention was the most reported process variable. Readiness for change was reported in 6 articles, all measured at the meso level. The second sub-cluster, perceptions, contained the most-reported process variable in all 44 articles: attitudes and perceptions of intervention users, examples of which included criticism of employees towards intervention consultants [26] and the belief of employees that management did not take their needs into account [27]. In the 30 articles reporting attitudes and perceptions of intervention users, almost all reported data were collected at least at the micro level. Engagement was found in articles that primarily reported data collection at the meso level. For engagement, Sorensen and Holman [28] reported differences in working groups in the extent to which they were able to include employees in the implementation process. In the third and last sub-cluster, process variables that represent changes in mental models were included. In this sub-cluster, effectiveness beliefs were reported most and are most often investigated at the micro-level.

Discussion

The aim of this systematic review was to explore which process variables have been used in SMI evaluation research. In the 44 articles, we found 47 process variables, which were divided into three clusters: (i) context contained 2 variables, (ii) intervention contained 31 variables, and (iii) mental models contained 14 variables. There was great variety in the process variables assessed, but the three most-reported were attitudes and perceptions of intervention users (mental models cluster), support (intervention cluster), and context (context cluster). Many process variables were different from those reported by Murta and colleagues [15]. This systematic review revealed that relatively few studies contained theoretical frameworks to guide process evaluations.

Half of the articles did not contain any reference to process evaluation literature in the introduction or the methods section.

Different frameworks for process evaluation are available. Two in particular were present in the studies included in this systematic review, and each provide a different perspective on process evaluation. The first framework, proposed by Linnan and Steckler [9], focuses on implementation. In the findings of the present review, the framework was represented by process variables such as dose delivered (the extent to which the intervention was made available by its providers), dose received (the extent to which the target population actively uses or engages in intervention facilities and activities), and fidelity (the extent to which the intervention was delivered as planned). Evaluating implementation answers the question “Was the intervention carried out as intended?”. However, this is only one of three important questions for intervention evaluation [8]. Focusing solely on implementation in the process evaluation leaves unanswered the question “Did the intervention bring about the intended (change in) exposure/behavior?”. The second framework, a model proposed by Nielsen and Randall [11], takes a broader view. It does not focus on implementation alone but also incorporates concepts such as initiation, implementation strategy, and mental models. By taking this broader view of process evaluation, information could also be gathered about the (change in) exposure or behavior. By adding mental models, for example, an explanation could be found for participants’ motivation to take part in intervention activities or make use of intervention facilities. This was illustrated by Biron et al [29], who reported that managers failed to use a stress risk assessment tool (ie, dose received) because they did not feel that stress was a problem (ie, attitudes and perceptions of intervention users).

A problem with this broader approach is that it might blur the lines between process variable and effect outcome. An attitude or perception that seems to influence intervention participation (and implementation) can be regarded as a process variable. Alternatively, maintaining or changing an attitude or perception can be an intermediate effect of an intervention, in which case it may be more accurately described as an effect outcome. An example of a process variable that could also be an intermediate effect is communication. Communication about the intervention may be important for implementation (the process thereof), but an intervention can also change the way different stakeholders interact, leading to improved communication (intermediate effect). The dilemma that arises is in which part of the intervention evaluation should this information be gathered and reported in the context of the process or the effect evaluation? A way to make this decision is to establish beforehand whether the variable is part of the underlying theory or working mechanism behind the intervention [8]. If this is the case, the variable should be regarded as an intermediate effect and measured as part of the effect evaluation. A systematic way to take intermediate effects into account is to formulate a program theory [10]. A program theory states under which conditions researchers expect proximal changes to occur [30] but seems to be missing in many of the included studies. Program theory evaluation can provide quantitative outcomes, which can be

related to intervention effect outcomes. Quantitative variables can give insight into the extent to which the intervention was used (eg, dose received), whereas qualitative data can provide more in-depth information (eg, barriers and facilitators). Sometimes, researchers might not yet be aware of certain intermediate effects. In that case, a qualitative process evaluation offers room for exploration, catering more to the practical nature of the applied research setting of interventions, in which fewer factors can be controlled than in a laboratory setting. This may explain why half of the articles contained reports of process variables but did not mention the use of any theoretical framework for their measurement.

Strengths and limitations

A strength of this review is our elaborate and thorough selection of studies; two independent researchers searched seven databases and systematically inspected 3613 titles and abstracts. Second, the background information on interventions and methods provided unique insight into specific circumstances in which process variables were assessed. Finally, careful deliberation resulted in a clustering structure tailored to the findings.

Some limitations should be considered when interpreting the results. First, as this is an explorative review, we chose to use broad definitions of process variables and evaluation. Consequently, more generally defined process variables were reported more often than specifically defined ones. One could argue that a study reporting only one broadly defined process variable can hardly be called a process evaluation. However, using broad definitions served the exploratory goal of this review, in which we aimed at inclusion rather than exclusion. Resulting from this, many process variables found were not part of the preliminary design of the study (ie, they were not part of the theoretical framework used for the evaluation of the intervention). A second limitation is the fact that during data extraction, interpretation was sometimes necessary to tease out the process variables. This meant that not every variable could be extracted literally. For example, employee readiness [31] was coded as readiness for change. To enhance the coding format and curb possible observer effects, the first and second author coded 20% of the included articles independently. Coding was completed only after consensus was reached on the first 20%. Despite the relatively large number of process variables found, it is possible that some variables were missed, especially because there was great heterogeneity in the naming of process variables.

Implications for research and practice

Both the heterogeneity of process variables used and lack of the use of a (standard) framework in process evaluation limit the possibility to compare results and build on previous experiences. This hinders the advancement of process evaluation theory development and limits the possibility to advise organizations about what is important for successful implementation of SMI. Future

process evaluations of SMI should be guided by a standardized, comprehensive framework that goes beyond assessing implementation only. The Nielsen and Randall [11] model of process evaluation provides a good starting point. Standardization would also be supported by the systematic use of a program theory, which would obligate researchers to measure if conditions for changes in behavior or exposure were in place and assess if intermediate stages were reached [30].

In most cases, process evaluation data were collected after intervention implementation and at the micro level (ie, at the level of the employee). As argued by Nielsen and Randall [11], retrospective evaluation may not capture changes in the process, and (in non-randomized controlled trial settings) does not provide the opportunity to take corrective action during intervention implementation should gaps emerge. Failing to collect information from stakeholders other than employees (micro level) also means that differences in perspectives among stakeholders might be overlooked. Many studies show, however, that for implementation success, support from other stakeholders is important [19, 29, 32–34]. In future process evaluations, researchers could place more emphasis on the collection of process data at different levels.

It should be noted that even though there have been substantial developments in the research field of process evaluations (for instance, the inception of the new journal *Implementation Science* in 2006), advancements should still be made in relating available process data to effect outcomes. This way, it could be assessed which process variables are central to successful implementation and predictive of intervention success.

Concluding remarks

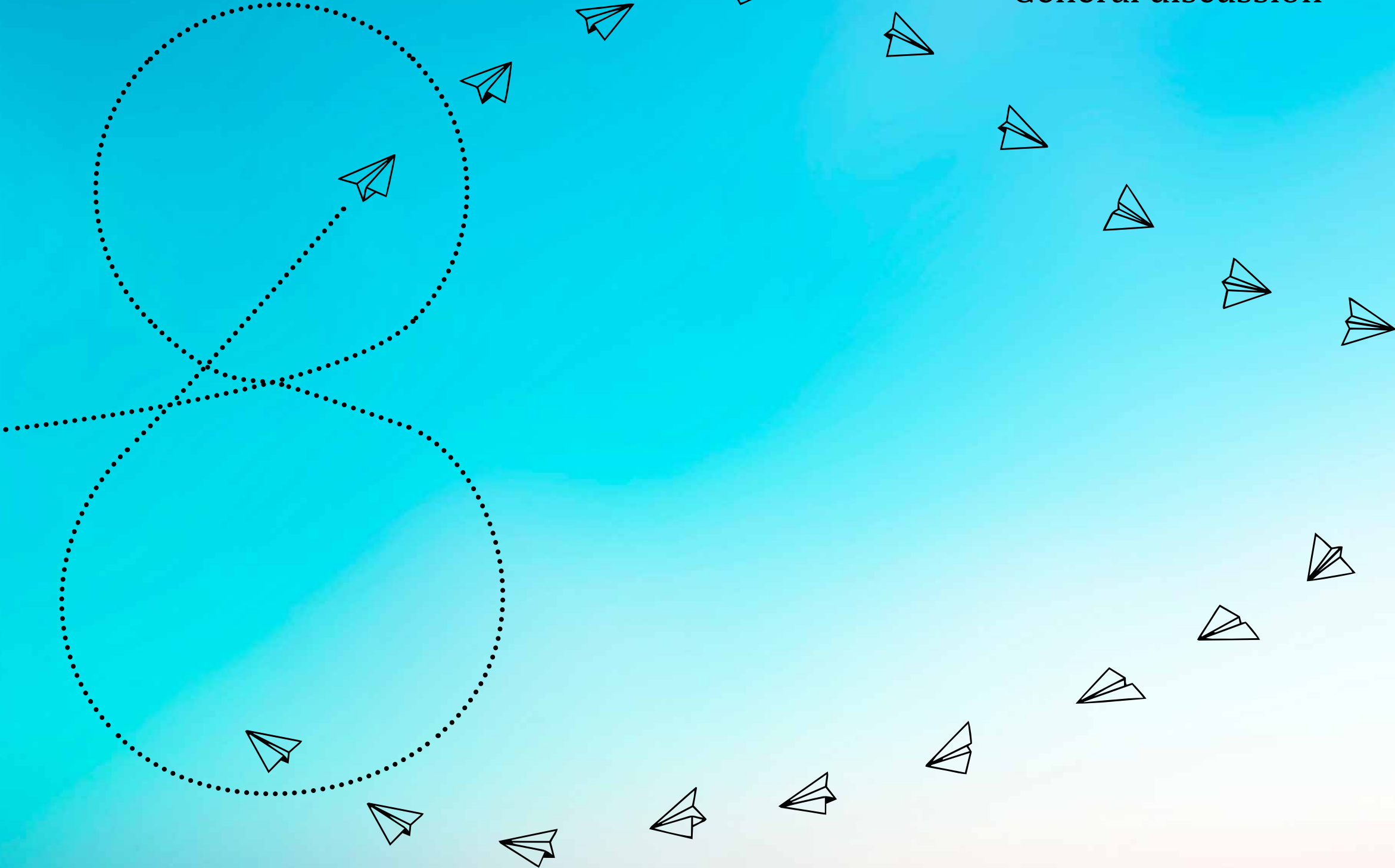
This review complements the process evaluation literature by giving insight into the use of process variables in SMI evaluation research. It revealed that there still is great heterogeneity in the methods and process variables used. It also found that many process variables were used in SMI evaluations other than those reported earlier and that, in many cases, no theoretical framework or program theory was used to guide measurement of process variables. In most cases, process variables were measured at the level of the employee and post intervention. Future process evaluations of SMI could benefit from data collection from different stakeholders (eg, employees, management, CEO) and at different times (before, during, and after the intervention). Also, the use of a theoretical framework could support a broader approach to process evaluation and may lead to a more standardized way of assessing intervention implementation.

References

- Ganster DC, Rosen CC: **Work Stress and Employee Health: A Multidisciplinary Review.** J Manage. 2013;**39**(5):1085-1122.
- Steptoe A, Kivimaki M: **Stress and cardiovascular disease: an update on current knowledge.** Annu Rev Public Health. 2013;**34**:337-354.
- Henderson M, Glozier N, Elliott KH: **Long term sickness absence: is caused by common conditions and needs managing.** Br Med J. 2005;**330**(7495):802-803.
- Hassard J, Teoh K, Cox T, Dewe P, Cosmar M, Gründler R *et al.*: **Calculating the costs of work-related stress and psychosocial risks.**, 1st edn. Luxembourg: European Agency for Safety and Health at Work; 2014.
- van der Klink JJ, Blonk RW, Schene AH, van Dijk FJ: **The benefits of interventions for work-related stress.** Am J Public Health. 2001;**91**(2):270-276.
- Ruotsalainen J, Serra C, Marine A, Verbeek J: **Systematic review of interventions for reducing occupational stress in health care workers.** Scand J Work Environ Health. 2008;169-178.
- Richardson KM, Rothstein HR: **Effects of occupational stress management intervention programs: a meta-analysis.** J Occup Health Psychol. 2008;**13**(1):69-93.
- Kristensen TS: **Intervention studies in occupational epidemiology.** Occup Environ Med. 2005;**62**(3):205-210.
- Steckler AB, Linnan L, Israel BA: **Process evaluation for public health interventions and research**, 1st edn. San Francisco: Jossey-Bass; 2002.
- Nielsen K, Abildgaard JS: **Organizational interventions: A research-based framework for the evaluation of both process and effects.** Work Stress. 2013;**27**(3):278-297.
- Nielsen K, Randall R: **Opening the black box: Presenting a model for evaluating organizational-level interventions.** Eur J Work Organ Psy. 2013;**22**(5):601-617.
- Montano D, Hoven H, Siegrist J: **Effects of organisational-level interventions at work on employees' health: a systematic review.** BMC Pub Health. 2014;**14**:135.
- Fleuren M, Wiefferink K, Paulussen T: **Determinants of innovation within health care organizations.** Int J Qual Health Care. 2004;**16**(2):107-123.
- Egan M, Bambra C, Petticrew M, Whitehead M: **Reviewing evidence on complex social interventions: appraising implementation in systematic reviews of the health effects of organisational-level workplace interventions.** J Epidemiol Community Health. 2009;**63**(1):4-11.
- Murta SG, Sanderson K, Oldenburg B: **Process evaluation in occupational stress management programs: A systematic review.** Am J Health Promot. 2007;**21**(4):248-254.
- Moher D, Liberati A, Tetzlaff J, Altman DG: **Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.** Ann Intern Med. 2009;**151**(4):264-269.
- Wierenga D, Engbers LH, Van Empelen P, Duijts S, Hildebrandt VH, van Mechelen W: **What is actually measured in process evaluations for worksite health promotion programs: a systematic review.** BMC Pub Health. 2013;**13**(1):1190-1205.
- Coffeng JK, Hendriksen IJM, van Mechelen W, Boot CRL: **Process evaluation of a worksite social and physical environmental intervention.** J Occup Environ Med. 2013;**55**(12):1409-1420.
- Ipsen C, Gish L, Poulsen S: **Organizational-level interventions in small and medium-sized enterprises: Enabling and inhibiting factors in the PoWRS program.** Saf Sci. 2015;**71**:264-274.
- Hasson H, Villaume K, Schwarz UV, Palm K: **Managing Implementation Roles of Line Managers, Senior Managers, and Human Resource Professionals in an Occupational Health Intervention.** J Occup Environ Med. 2014;**56**(1):58-65.
- Renaud L, Kishchuk N, Juneau M, Nigam A, Tetreault K, Leblanc M-C: **Implementation and outcomes of a comprehensive worksite health promotion program.** Can J Public Health. 2008;**99**(1):73-77.
- Jeffcoat T, Hayes SC: **A randomized trial of ACT bibliotherapy on the mental health of K-12 teachers and staff.** Behav Res Ther. 2012;**50**(9):571-579.

23. Gunusen NP, Ustun B: **Turkish nurses' perspectives on a programme to reduce burnout.** *Int Nurs Rev.* 2009;**56**(2):237-242.
24. Randall R, Cox T, Griffiths A: **Participants' accounts of a stress management intervention.** *Hum Relat.* 2007;**60**(8):1181-1209.
25. Saksvik PO, Nytro K, Dahl-Jorgensen C, Mikkelsen A: **A process evaluation of individual and organizational occupational stress and health interventions.** *Work Stress.* 2002;**16**(1):37-57.
26. Aust B, Rugulies R, Finken A, Jensen C: **When workplace interventions lead to negative effects: learning from failures.** *Scand J Public Health.* 2010;**38**(3 Suppl):106-119.
27. Bourbonnais R, Brisson C, Vinet A, Vezina M, Lower A: **Development and implementation of a participative intervention to improve the psychosocial work environment and mental health in an acute care hospital.** *Occup Environ Med.* 2006;**63**(5):326-334.
28. Sorensen OH, Holman D: **A participative intervention to improve employee well-being in knowledge work jobs: A mixed- methods evaluation study.** *Work Stress.* 2014;**28**(1):67-86.
29. Biron C, Gatrell C, Cooper CL: **Autopsy of a Failure: Evaluating Process and Contextual Issues in an Organizational-Level Work Stress Intervention.** *Int J Stress Manage.* 2010;**17**(2):135-158.
30. Biron C, Karanika-Murray M: **Process evaluation for organizational stress and well-being interventions: Implications for theory, method, and practice.** *Int J Stress Manage.* 2014;**21**(1):85-111.
31. Randall R, Nielsen K, Tvedt SD: **The development of five scales to measure employees' appraisals of organizational-level stress management interventions.** *Work Stress.* 2009;**23**(1):1-23.
32. Weigl M, Hornung S, Glaser J, Angerer P: **Reduction of hospital physicians' work-flow interruptions: A controlled unit-based intervention study.** *J Healthc Eng.* 2012;**3**(4):605-620.
33. Dahl-Jorgensen C, Saksvik PO: **The impact of two organizational interventions on the health of service sector workers.** *Int J Health Serv.* 2005;**35**(3):529-549.
34. Jenny GJ, Brauchli R, Inauen A, Fullemann D, Fridrich A, Bauer GF: **Process and outcome evaluation of an organizational-level stress management intervention in Switzerland.** *Health Promot Int.* 2015;**30**:573-585.
35. Albertsen K, Garde AH, Nabe-Nielsen K, Hansen AM, Lund H, Hvid H: **Work-life balance among shift workers: results from an intervention study about self-rostering.** *Int Arch Occup Environ Health.* 2014;**87**(3):265-274.
36. Bruneau BM, Ellison GT: **Palliative care stress in a UK community hospital: evaluation of a stress-reduction programme.** *Int J Palliat Nurs.* 2004;**10**(6):296-304.
37. Bunce D, West MA: **Stress management and innovation interventions at work.** *Hum Relat.* 1996;**49**(2):209-232.
38. Burton NW, Pakenham KI, Brown WJ: **Feasibility and effectiveness of psychosocial resilience training: a pilot study of the READY program.** *Psychol Health Med.* 2010;**15**(3):266-277.
39. Cohen-Katz J, Wiley S, Capuano T, Baker DM, Deitrick L, Shapiro S: **The Effects of Mindfulness-based Stress Reduction on Nurse Stress and Burnout: A Qualitative and Quantitative Study, Part III.** *Holist Nurs Pract.* 2005;**19**(2):78-86.
40. Falck VT, Kilcoyne J: **A health promotion program for school personnel.** *J Sch Health.* 1984;**54**(7):239-243.
41. Hasson H, Brisson C, Guerin S, Gilbert-Ouimet M, Baril-Gingras G, Vezina M *et al.*: **An organizational-level occupational health intervention: Employee perceptions of exposure to changes, and psychosocial outcomes.** *Work Stress.* 2014;**28**(2):179-197.
42. Heaney CA, Israel BA, Schurman SJ, Baker EA, House JS, Hugentobler M: **Industrial relations, worksite stress reduction, and employee well-being: A participatory action research investigation.** *J Organ Behav.* 1993;**14**(5):495-510.
43. Helms-Lorenz M, Slof B, van de Grift W: **First year effects of induction arrangements on beginning teachers' psychological processes.** *Eur J Psy Educ.* 2013;**28**(4):1265-1287.
44. Kawai K, Yamazaki Y, Nakayama K: **Process evaluation of a web-based stress management program to promote psychological well-being in a sample of white-collar workers in Japan.** *Ind Health.* 2010;**48**(3):265-274.
45. Keller SR, Engen DJ, Bauer BA, Holmes DR, Rihal CS, Lennon RJ *et al.*: **Feasibility and effectiveness of massage therapy for symptom relief in cardiac catheter laboratory staff: A pilot study.** *Complement Ther Clin Pract.* 2012;**18**(1):4-9.
46. Landsbergis PA, Vivena-Vaughan E: **Evaluation of an occupational stress intervention in a public agency.** *J Organ Behav.* 1995;**16**(1):29-48.
47. Leonard R, Alison L: **Critical incident stress debriefing and its effects on coping strategies and anger in a sample of Australian police officers involved in shooting incidents.** *Work Stress.* 1999;**13**(2):144-161.
48. Mellor N, MacKay C, Packham C, Jones R, Palferman D, Webster S *et al.*: **'Management Standards' and work-related stress in Great Britain: Progress on their implementation.** *Saf Sci.* 2011;**49**(7):1040-1046.
49. Mellor N, Smith P, MacKay C, Palferman D: **The "Management Standards" for stress in large organizations.** *Int J Workplace Health Manage.* 2013;**6**(1):4-17.
50. Milliar P, Lioussis P, Shochet IM, Biggs H, Donald M: **Being on PAR: Outcomes of a pilot trial to improve mental health and wellbeing in the workplace with the promoting adult resilience (PAR) program.** *Behav Change.* 2008;**25**(4):215-228.
51. Nielsen K, Fredslund H, Christensen KB, Albertsen K: **Success or failure? Interpreting and understanding the impact of interventions in four similar worksites.** *Work Stress.* 2006;**20**(3):272-287.
52. Nielsen K, Randall R, Albertsen K: **Participants' appraisals of process issues and the effects of stress management interventions.** *J Organ Behav.* 2007;**28**(6):793-810.
53. Nielsen K, Randall R: **The importance of employee participation and perceptions of changes in procedures in a teamworking intervention.** *Work Stress.* 2012;**26**(2):91-111.
54. Petterson I-L, Arnetz BB: **Psychosocial stressors and well-being in health care workers. The impact of an intervention program.** *Soc Sci Med.* 1998;**47**(11):1763-1772.
55. Reynolds S, Taylor E, Shapiro D: **Session impact and outcome in stress management training.** *J Community Appl Soc Psy.* 1993;**3**(4):325-337.
56. Reynolds S, Taylor E, Shapiro DA: **Session impact in stress management training.** *J Occup Org Psy.* 1993;**66**(2):99-113.
57. Schwerman N, Stellmacher J: **A holistic approach to supporting staff in a pediatric Hospital setting.** *Work.* 2012;**60**(9):385-390.
58. Swain N, Gale C: **A communication skills intervention for community healthcare workers reduces perceived patient aggression: A pretest-posttest study.** *Int J Nurs Stud.* 2014;**51**(9):1241-1245.
59. Van Bogaert P, Van Heusden D, Somers A, Tegenbos M, Wouters K, Van Der Straeten J *et al.*: **The productive ward program (trade-mark): A longitudinal multilevel study of nurse perceived practice environment, burnout, and nurse-reported quality of care and job outcomes.** *J Nurs Admin.* 2014;**44**(9):452-461.
60. van Wingerden J, Derks D, Bakker AB, Dorenbosch L: **Job crafting in schools for special education: A qualitative analysis.** *Gedr Org.* 2013;**26**(1):85-103.
61. Weigl M, Hornung S, Angerer P, Siegrist J, Glaser J: **The effects of improving hospital physicians working conditions on patient care: a prospective, controlled intervention study.** *BMC Health Serv Res.* 2013;**13**:401-408.
62. Westlander G: **Means, Goals, and Outcomes of a Comprehensive Occupational Health Program for Telephone Operators.** *Int J Health Serv.* 1995;**25**(2):313-332.
63. van Berkel J, Boot CRL, Proper KI, Bongers PM, van der Beek AJ: **Process evaluation of a workplace health promotion intervention aimed at improving work engagement and energy balance.** *J Occup Environ Med.* 2013;**55**(1):19-26.

General discussion



General discussion

The main objective of this thesis was to explore ways to decrease work stress in teachers from an individual, organizational and intervention evaluation perspective, and thereby eventually contribute to excellent education. The exploration was conducted along three key questions:

1. How can we decrease work stress in teachers from an individual perspective?
2. How can we decrease work stress in teachers from an organizational perspective?
3. How can we gather the most relevant evidence in intervention studies in the occupational setting, for example to decrease work stress?

In this chapter the main findings are presented. Thereafter these findings are placed in context and (methodological) considerations are described. Lastly, recommendations for future practice and research are discussed.

Main findings

How can we decrease work stress in teachers from an individual perspective?

In order to answer this question a longitudinal study among 549 older teachers (aged 45-64 years) was conducted, wherein the role of a personal resource, mastery, in the decrease of work stress was studied (**chapter 2**). Path analysis was applied to explore the interplay between mastery and job demands on the one hand (i.e. psychological demands and emotional demands) and job resources on the other hand (i.e. autonomy and social support), in influencing work stress-related outcomes (i.e. depression and work engagement). Mastery had a direct negative influence on depression and a direct positive influence on work engagement. Mastery partly countered the negative effect of high psychological job demands on depression, and it revealed a small positive effect on work engagement, whereas it partly explained the longitudinal relations between emotional job demands and both outcomes. Increasing mastery in teachers might be a promising intervention strategy to decrease depression and increase work engagement, although this finding requires replication and actual experimentation.

How can we decrease work stress in teachers from an organizational perspective?

A participatory, primary preventive, organizational level intervention was conducted among the employees of two secondary vocational education schools

(school A and school B), in order to decrease work stress. The intervention had been developed by a consultancy firm and applied over a hundred times in both public and private organizations. However, it was never evaluated scientifically. As described in more detail in **chapter 3**, the intervention comprised a needs assessment phase (consisting of interviews, a questionnaire, and group sessions) and a phase of implementing intervention activities, which were derived from the needs assessment phase (e.g. structured performance reviews, or creating a staff room). A facilitator supervised the needs assessment phase and he was assisted by a participatory group of employees and staff members. In the phase of implementation of intervention activities the facilitator's help was optional, and only one of the two schools (school A) purchased this aid. The primary hypothesis was that the intervention would decrease work stress when need for recovery was decreased and vitality was increased (**chapter 4**). In a controlled trial including 356 employees this hypothesis could not be confirmed. Two small but statistically significant effects in unfavorable direction were found on two of the secondary outcomes (i.e. absorption and organizational efficacy). However, post-hoc per protocol analyses demonstrated that employees who participated in at least two of the three steps of the needs assessment (e.g. questionnaire and group session) had a significantly higher post-intervention score on occupational self-efficacy (i.e. the belief in one's ability at work) than the control group, which is favorable. This result indicates that those employees with a higher participation grade regarding the intervention, might have had a mastery experience, which then positively affected occupational self-efficacy.

The lack of positive effects on the primary outcomes could be partly due to implementation failure, as was described in the process evaluation in **chapter 5**. This evaluation showed that the needs assessment phase was implemented successfully in school A, but not in school B where participation and readiness for change were insufficient. The results for school A and B diverged further in the phase of implementation of intervention activities: in school A several intervention activities were implemented, whereas this was hardly the case in school B. Regardless of actual participation and actual implementation of intervention activities, participants in both schools felt not involved in the choice of intervention activities. Furthermore, in both schools participants generally perceived the intervention's second phase negatively, for example because the intervention activities were not seen as the right solutions for the problems. Three type of context factors further inhibited the implementation. That is, the schools had no capacity (school B) or only partial capacity (school A) to implement the intervention activities, foremost because middle managers were not fully equipped to do so. Secondly, the organizational culture in both schools was characterized by a lack of mutual trust between managers and teachers. Thirdly, several external events interfered with the planned intervention activities (e.g. a national change was announced by the Ministry of Education, implying an intensification of classroom-bound lessons from 850 to 1000 hours per year).

In conclusion, further implementation of the intervention in its current form is not recommended. To decrease work stress in teachers (and other employees in schools) participatory, primary preventive, organizational level interventions should firstly include a more elaborate implementation strategy upfront, including an allocation of roles and responsibilities of all involved (i.e. Executive Board, directors, middle managers, intervention facilitator, participatory group, participants). Furthermore, those responsible for the intervention should communicate more specifically how the intervention activities to be implemented in the intervention's second phase relate to the problems described in the intervention's first phase (i.e. the needs assessment). Also, it should be taken into account that participants' perception of what has happened during the intervention process can differ from what actually happened, underlining the importance of 'appraisal' in intervention processes. Lastly, the attitudes and actions of managers seem to be important for successful implementation, as well as the organizational culture.

How can we gather the most relevant evidence in intervention studies?

In designing the evaluation of the organizational level intervention two methodological problems were encountered, which led to chapter 6 and 7. Firstly, both schools wanted to participate in the study with a specific department, for which they believed the intervention would be most relevant. Although conceivable from the school perspective, this posed a methodological issue because randomization procedures were thereby obstructed. A controlled trial was conducted, making the effect evaluation prone to confounding and selection bias. As a consequence, baseline differences on several outcomes were present between the intervention and control group. Our intervention evaluation study was not the first to suffer from such a challenge and the narrative review (**chapter 6**) described this and other common challenges that occupational health intervention researchers encounter. In this chapter we furthermore explored alternative research designs for the situation wherein a regular Randomized Controlled Trial (RCT) is not feasible. One experimental alternative design (i.e. stepped wedge design) and several observational designs were found and discussed (i.e. propensity scores, instrumental variables, multiple baseline design, interrupted time series, difference-in-difference, and regression discontinuity). Future primary preventive, participatory, organizational level interventions could benefit especially from the stepped wedge randomized design whereby the intervention is randomly applied to all groups but sequentially over time, or the multiple baseline design, which does the same although not randomly. Of course these designs poses their own challenges, such as the increased number of participants and measurement moments that are needed. In order to gather the most relevant evidence in intervention studies, researchers should consider beforehand the specific research question, the complexity of the intervention, the availability of existing data, the context, and the costs before choosing the most appropriate design.

The second methodological problem that was encountered was related to the design of the process evaluation. Because the field of process evaluation is rapidly developing we noticed that many different theoretical frameworks exist, which include many overlapping process variables. In order to understand what is the most relevant evidence with regard to the implementation process, a systematic literature review of 44 process evaluations of organizational level stress management interventions was conducted (**chapter 7**). The review demonstrated that there is indeed great heterogeneity in methods and process variables studied: among the 44 evaluations, 47 different process variables were found. Also a theoretical framework or program theory, which could guide the choice of process variables, was missing in half of the articles. Furthermore, the process evaluation data were mostly collected after the intervention took place, instead of also before and during the intervention. We concluded that a standardized framework including variables with regard to the intervention, participants' mental models, and context factors would improve the theory development in the field of process evaluation for stress management interventions.

The main findings in context

The intervention study did not render the hypothesized results. This finding is placed in context starting with a comparison of the expert-based intervention studied in this thesis, with the evidence-based approach for decreasing work stress. In the sections following the comparison, explanations for the lack of results are presented related to employees' behavior, and leadership and organizational culture. Lastly, the relation with participation is discussed.

Similarities and differences between expert-based versus evidence-based approach

The participatory, organizational level intervention (i.e. the Heuristic Method, HM) studied in the second part of this thesis was developed in practice, by a consultancy firm. HM consists of two phases: a needs assessment phase and a phase of implementing intervention activities. The phases are facilitated by an intervention facilitator. A representative participatory group is formed, thereby ensuring a fit between the organizational context and content and the process of the intervention. Even though developed, applied, and redeveloped in practice, the structure of the intervention is similar to the evidence-based psychosocial risk management approach [1-4]. The psychosocial risk management approach consists of five participatory steps that can be summarized as: 1) Preparation; 2) Screening; 3) Action Planning; 4) Implementation; and 5) Evaluation [4, 5].

A central common factor in HM and the psychosocial risk management approach is participation of the target group. Furthermore, steps 2 and 3 of the psychosocial risk management approach are similar to HM's needs assessment phase, whereas step 4 is similar to the implementation of intervention activities phase.

The most prominent difference between HM and the psychosocial risk management approach is step 1, the preparation phase, which is not a formal element of HM. In the evidence-based approach, this step typically consists of: the formation of a steering group, ensuring senior management support, and communicating about the project [4]. Previous research has also demonstrated that employee and organizational readiness for change are important prerequisites [4], as a lack of employee and organization readiness for change is likely to influence intervention development, implementation, and effectiveness [6]. Readiness for change will more likely occur if participants perceive the current situation as undesirable and change as necessary [7, 8]. With regard to the intervention activities, participants need to believe they can benefit from it, and they need to be motivated to engage in it [7, 8]. The 'preparation' step is ideal for creating readiness for change, in practice however this step is sometimes reduced to installing a steering group, signing a declaration or contract stating higher management's intended support, and sending a department-wide email about the intervention project. In the following paragraphs I will look back on the intervention study to argue that a more elaborate 'preparation' step can help to identify preexisting context conditions, which are indicative of (lacking) readiness for change. Impression from the research logbooks will be used as an illustrative starting point.

Explanation 1: Employees' participation hindered by learned helplessness

Impression 1: The director and his secretary are welcome to attend the first fifteen minutes of a participatory group meeting, to jointly prepare an informative event with all employees wherein the intervention project will be presented. The director proposes an idea and does much of the talking, while the goal of the meeting was to get employees' input for the event. The participatory group gets quieter and quieter. The arrangements are made quickly. When the director and his secretary have left after ten minutes, it turns out that the teachers completely disagree with the directors' proposal for the event. They talk loudly and boldly. To my question 'why didn't you state these objections earlier?' they have no clear answer.

In this impression there are several players: the director (and his secretary), the participatory group, the intervention facilitator, and the researcher. The most important lesson to be learnt from this impression is probably in the interaction between the director and the participatory group. The group did not

give their opinion in presence of the director, even though the whole purpose of him attending their meeting was for the group to give their opinion. Why? Because the director started the conversation with a concrete proposal, the participatory group might have felt little space to contradict his proposal, or they might not have been accustomed to disagreeing with a superior. To me as a researcher it reflected an unequal relationship, resembling the traditional teacher-student relationship: we will nod and say yes when he is watching, however if he turns to the blackboard we will throw paper balls at his head. The characteristics of the principal – teacher relationship have found to be important for teacher as well as student achievement [9, 10].

Another explanation could be that the teachers did not feel speaking up would matter, because they believed the director had already proposed something and that is how it will be executed, regardless of their opinion. This phenomenon could be explained using insights from (expanded) 'learned helplessness' theory [11, 12]. Learned helplessness is human behavior, characterized by refraining from responding to (and thus influence) an event, while it could have been controlled by our own actions [11]. Learned helplessness is also learned behavior, which originates after experiencing (an) uncontrollable event(s), and it can result in passivity, inability to learn that responding can be effective, and stress [11, 12]. How employees react to a felt lack of control varies with individual characteristics and with situations, it depends on an individual's dominant perception, referred to as 'explanatory style' [13]. The explanatory style can be changed by cognitive behavioral techniques, by replacing unconstructive perceptions with more constructive perceptions [11].

I argue that employees in the educational sector experienced many uncontrollable events in the last decades, which has led to a degree of learned helplessness. In the international scientific literature on the organization of education, evidence has been found of several 'uncontrollable' events that have impacted the basis of the teaching profession. For example the increased student-teacher-ratio [14], the integration of students with special needs in regular classes [15], and the increase of accountability measures, which increased the administrative load [15]. This increased administrative burden has not only contributed to an expansion in the core tasks of a teacher, many teachers feel it also questions their professionalism. They perceive accountability measures as a series of mistrust measures, a control mechanism that limits their decision latitude (an important job resource). Australian research demonstrated that uncontrollable events have happened in the educational sector, for example the implementation of several (often even overlapping) school and school system reforms [16]. In the Netherlands, the Parliamentary Commission Dijsselbloem reached the same conclusion in 2008 after extensive research in the educational field, among policy makers, and schools [17]. The Commission found that the school and school system reforms had been planned with political support, but without satisfactory support in the educational sector. Probably because the union representatives who were involved in planning these reforms were closer to politics than to the field [17].

We have observed that if employees experience a degree of learned helplessness due to the experience of several uncontrollable events over time, it is difficult to give them ‘control’ in a participative intervention. In such case it seems not likely that the participative elements of an intervention will be implemented as intended and thus have the intended effects. In future participatory studies in education (and maybe also in other sectors that have experienced many events perceived as uncontrollable), researchers should be aware of the possible influence of learned helplessness on the intervention process. Ideally researchers get an impression of whether it is present in employees (and to what degree) during the first step of the psychosocial risk management approach, so that it can be either taken into account in the planned intervention or ‘unlearned’ before the planned intervention is implemented, by means of cognitive behavioral techniques. Cognitive behavioral techniques can help to structure and adjust employees’ misconceptions and faulty assumptions [18]. This is done by creating specific learning experiences, wherein employees firstly learn to recognize automatic, negative thoughts. Secondly they learn how these thoughts are related to their affect and subsequent behavior. And thirdly, employees learn to gather evidence for and against their automatic thoughts and to replace them with more balanced thoughts. The empirical evidence base for CBT as a helpful tool in changing cognitions in order to increase well-being, is strong [19].

The concept of learned helplessness is related to employees’ perceptions and appraisal, which is described as a second explanation for not finding the expected results.

Explanation 2: Addressing employees’ perception and appraisal was not an intervention component

Impression 2: In conducting the observations for the process evaluation I found that employees told each other (often faulty or misinformed) ‘stories’ about the intervention and the implementation process, such as: “nothing is changing, because the director opposes the change” or “we never see the intervention facilitator anymore because he was fired by the director for speaking up and telling him our wishes”.

During a working day employees continuously engage in sense making processes, they ‘perceive’ their environment and attribute meaning to it, that is they ‘appraise’ their situation [20]. This process continues when an intervention is being conducted within an organization, probably impacting the outcomes under study. Evidence for this sense making process was found by Nielsen and colleagues, when they assessed the role of employee appraisal of the intervention process in eleven intervention projects to reduce work stress in Denmark [21]. The researchers found full mediation of appraisal between exposure to the intervention and outcomes. More specifically, employees’ appraisal of intervention quality and their ability to influence the intervention

mediated between exposure and outcomes (i.e. changes in working conditions, work stress, and job satisfaction). This example shows that employees’ appraisal can have a powerful influence in the intervention process.

In the medical and public health literature appraisal has been discussed mostly in terms of placebo effects. A placebo effect is observed as a patients’ health improvement after medical treatment without known active medical substance [22]. A negative placebo effect (a health deterioration after receiving medical treatment without known active medical substance) is called nocebo effect [22]. The nonspecific placebo treatment effect is often considered an interference, something we should control for with double-blind RCTs, instead of a beneficial intervention element [23]. Bensing and Verheul [23] have described three mechanisms through which placebo (or nocebo) work: conditioning, expectancies, and affect manipulation. I will describe how these three mechanisms are relevant with regard to employees’ perception and appraisal of organizational level, primary preventive occupational health interventions.

Firstly, participants can have either positive or negative previous experiences with interventions for work stress, which have resulted in a reduction of work stress (positive experience) or not (negative experience) [23]. Via classical conditioning, these previous experiences influence the results of future work stress interventions. In case of the negative experience, a neutral stimulus (e.g. the intervention) coincided with a negative outcome (e.g. feeling stressed). The formerly neutral stimulus (intervention) then becomes a conditioned stimulus, which is in itself enough to produce the negative outcome (feeling stressed). Occupational health research has demonstrated that if organizations failed to learn from previous, similar intervention projects, this will negatively impact participants’ perceptions of other interventions, as well as their willingness to participate in intervention activities [24, 25]. From this we can learn that in conducting an intervention we should always take previous experience with (an) intervention(s) into account, because participants’ conditioned responses to (the announcement of) an intervention can produce (unintended) outcomes in itself.

Secondly, expectancies of an intervention can result in a (placebo) effect, simply because the patient expects the intervention to lead to an effect [23]. While a conditioned response is unconscious, holding expectancies is a conscious process (e.g. “this time something really is going to change”)[23]. Expectancies can also exist on a group level, occupational health research has demonstrated that employees who work closely together can develop a shared way of perceiving and responding to events in the work context [26], and thus also respond similar to interventions in the work context [27]. Expectancies can become strong enough to overcome (negative) conditioned responses [28]. Timely, sufficient, and tailored communication plays an important role in shaping expectancies before, during, and after an intervention.

Affect manipulation, the third mechanism that can produce a placebo effect, works through patients’ conscious, subjective feeling (affective state), which

is influenced positively or negatively by how a treatment is applied [23]. Applied to work stress interventions, this would mean that how an intervention is conducted can be a positive or negative experience, which can result in consequent positive or negative feelings. This has indeed been demonstrated in organizational level primary preventive work stress interventions: the change process such an intervention implies can be perceived as stressful in itself by participants and thus cause negative effects [29]. Furthermore, medical research has demonstrated that the negative placebo or nocebo response is even stronger in patients who are already stressed or anxious [30]. Therefore, it seems important to make participation in an intervention a positive experience that induces a positive affective state, which then would increase the chance of a placebo effect. This is especially relevant in organizations where many employees are stressed.

Based on the findings of the current study combined with the insights from the placebo/nocebo literature, I argue that employees' appraisal should be acknowledged as an influence in every occupational health intervention. More efforts should be directed to exploring and influencing employees' appraisal upfront, by using the three mechanisms which are known to produce placebo effects, i.e.: recognizing conditioned reactions, shaping expectancies, and inducing positive affective states. Open-ended interviews or observations in the work context are tools for assessing what are current conditioned responses, expectancies, and dominant affective states with regard to the intervention. Tailored communication and making the intervention a fun experience can be tools for changing these elements.

Explanation 3: Leadership and organizational culture did not enable the (implementation of the) intervention

Impression 3: The advisory report stated the wishes, needs, and opinions of all employees in the intervention group, obtained by using a participatory process. Once the intervention facilitator has presented the advisory report to the Executive Board, their first reaction is to ask for an 'objective evaluation' of its content.

Since the content of the advisory report was the result of a participatory process, asking for a sort of 'second opinion', could also be interpreted as the implicit questioning of the employees' perspective. In doing so, the Board actually disregards employees' wishes, needs, and opinions. The example could be seen as indicative of a particular leadership style. A leadership style is a relatively stable pattern of behavior shown by a leader [31]. School principals leadership behaviors directly and indirectly affect teachers achievement, especially via the structuring of teachers working conditions [32]. A leadership style can be effective in one situation, but not in the other. In the above impression, an authoritarian leadership style seems to have been taken, which is characterized by controlling and top down communication, with little room for dialogue

[33]. While style can have positive effects in chaotic situations (e.g. after a natural disaster [33]), this result is outnumbered by the studies wherein authoritarian leadership was associated with negative outcomes [31]. More specifically, in a lab study supervisor dominance was associated with lower perception of trust among teachers [34]. Whereas a dissertation in the educational field demonstrated that top management humility predicted teachers' positive perceptions of leader effectiveness [35]. I argue that a leadership style that relies on dialogue, two way communication, and joint decision making – such as a facilitative, democratic, or transformational leadership style [9, 10] – would best suit a participatory intervention. If an authoritarian leadership style is present, it seems unlikely that organizational and employee readiness for a participatory intervention is high enough to conduct it successfully.

Leadership style is known to influence organizational culture [33] and vice versa. The following example of a middle management meeting demonstrates how top level management style has spilled over to the middle managers:

Impression 4: The (middle) management team asks the intervention facilitator to attend their meeting and again explain what will be done in step three of the interventions' first phase, the group sessions with teams. The facilitator explains that the teams will work on formulating solutions to problems related to happy and healthy working. One manager states: how can we know that their solutions will be feasible? Another middle manager adds that he can already predict what solutions will be suggested by some teams. The conversation continues in this atmosphere until one middle manager exclaims: "It almost seems you have no confidence in the teams". The others do not really respond to this, although they do say that they have no need for "grumpy group meetings".

Impressions three and four characterize the relations between top and middle management, and employees. How management and employees relate to each other, to their work, and to the outside world has been defined as 'organizational culture' [36]. Hofstede and colleagues found that the core of an organizational culture is made up of employees' shared perceptions of daily practices [36]. Daily practices are considered 'the way things generally go in my organization', or in the words of Hofstede and others: "conventions", "customs", "habits", "mores", "traditions," or "usages." (p. 311 [36]). Based on impression three (top management perspective) and four (middle management perspective), it seemed to be managers' shared perception that employees' point of view is of limited value. This practice is in line with one of the two independent organizational culture dimensions that are associated with the management style of (top) leaders, that is the 'Employee-Oriented vs. Job-Oriented' dimension [36]. The dimension ranges from a primary concern for employees (characterized by statements such as "where I work, decisions are not centralized at the top"), to a primary concern for the work (characterized by statements such as "where I work, changes are imposed by management decree") [36]. Participatory interventions (for work stress) probably fit best with organizations that are 'employee-oriented' rather than 'job-oriented'. In the current intervention

study, the limited trust in the employees' point of view was not indicative of an 'employee-orientation'. Therefore, this probably hindered the implementation of the participatory intervention. Another intervention study pointed to a second reason for the importance of an 'employee-orientation' for the success of an intervention [37]. The researchers found that divisions higher in employee-orientation (referred to as 'strong human relations values') had higher levels of readiness for change pre-intervention, which predicted the post-intervention score on the outcome.

In conclusion, it has been described (in other settings) that elements of organizational culture can make (enable) or break (hinder) an intervention [38]. Shared perceptions of daily practices are at the core of organizational culture and can inhibit the intervention process as well as pre-intervention readiness for change. Future participatory, organizational level intervention studies should take the organizational culture into account, by trying to capture important cultural aspects in the preparation phase of the psychosocial risk management approach. This can be done by a mixed methods approach referred to as 'qualitative orientation and quantitative verification' [36], wherein interviews or observations are used to tailor questionnaires that are supposed to measure dimensions of organizational culture.

Determine the degree of participation in decision making that is feasible and suitable

I have described how the context factors (i.e. employees' behavior and perceptions, leadership style and organizational culture) did not fit the participatory nature of the intervention described in this thesis. Nevertheless, stakeholder involvement or participation is considered an essential element of occupational health interventions, throughout the whole research process of intervention development, implementation and effectiveness study [39]. Participation is considered important because it can increase the support base for the intervention as well as the feasibility of intervention activities [40]. Furthermore, participating can be an empowering experience in itself [40]. A last reason for employee participation in decision making can be that it influences employee appraisal: being involved reduces the number of alternative explanations of a certain decision that will circulate in an organization [41]. In the literature on participation in decision making, a continuum of participation is distinguished with choice control at one end and voice control at the other end [42]. In a situation of complete 'choice or decision control' the participant directly influences the decision that is made, whereas in a situation of 'voice or process control' the participant can give his or her input on the problem formulation, gathering information, or defining alternatives. This process ends before the actual decision is made. In occupational health interventions both types of participation are encouraged [40, 43] and applied (for an example of choice control, see [44]; for an example of voice control, see [45]). In the organizational level intervention under study in this thesis, a degree of both choice and voice control was present. Choice control because of the participatory working

group, and voice control in the formulation of problems and measures in the three steps of the needs assessment. Neither form of participation functioned really well in the current intervention. In organizations with adverse context factors for participation, some degree of 'voice or process control' might be the highest achievable. In preparing an intervention it seems relevant to consider what degree of participation in decision making is feasible and suitable for that specific organization.

A need for a more elaborate 'preparation phase'

I have discussed how employees' behavior and perceptions, as well as leadership style and organizational culture affected the implementation process directly, or indirectly via (lacking) readiness for change and failing participation. If we conduct a more elaborate first step in a psychosocial risk management approach (preparation), lacking readiness for change will be detected earlier and the context factors could be taken into account in the design of the intervention already. Elaboration of the first step in education could consist of: 1) assessing (by observations) and addressing (by cognitive behavioral techniques) learned helplessness, 2) recognizing conditioned reactions, shaping expectancies, and inducing positive affective states with regard to the intervention, 3) assessing cultural aspects (by mixed methods) that are indicative of leadership style (which preferably is democratic), and organizational culture (which preferably is employee-oriented).

Considerations

In the following paragraphs several considerations with regard to the methodology and findings of this thesis are presented.

Methodological difficulties due to the definition of work stress

Work stress is defined in different ways as exemplified by the meta-analytic finding that in 55 work stress interventions, 60 different outcome variables were used to measure work stress related determinants or outcomes [46]. Following this finding, it is concluded that work stress is considered multifactorial in its causes and consequences. Due to the multifactorial nature, the choice of suitable outcome measures is difficult. Three such difficulties are described hereafter, that is: stress as the result of an imbalance between job demands and job resources; stress as a stage in a process rather than a definite outcome; and the measurement level of stress.

Most work stress models are based on a 'balance' idea, in short: work stress can arise if job demands outweigh job resources. In both the cohort study (**chapter 2**) and the intervention effectiveness study (**chapter 4**), we measured several job demands and job resources, but we did not combine them in order to determine whether a balance or imbalance was present. Instead, we used a burnout

measure and two proxies of work stress (i.e. need for recovery and vitality) in order to determine whether the intervention and control group had stress complaints or not. Few studies have tried to capture the imbalance in itself. If we still believe work stress is the result of imbalanced demands and resources, we might try to calculate the imbalance in order to develop a more sensitive measure of work stress determinants. More advanced statistical techniques could make this possible. Jenny and colleagues [47] did an interesting attempt and computed a ratio between demands and resources, which was then related to outcomes such as burnout.

The second difficulty is that stress is a stage in a process rather than one definite outcome. Evaluations of work stress interventions could take the stress process into account by means of developing a logic model of expected changes, wherein proximal, intermediate, and distal effectiveness outcomes are distinguished [48, 49], instead of only distal outcomes. Proximal outcomes have been described as those effects that arise immediately, whereas intermediate effects concern changes in factual processes (e.g. changes in resources-demands ratio, the work organization, supervisor behavior) and social processes (e.g. changes in collaboration), and distal effects are psychosocial health outcomes [48].

The third difficulty is related to the measurement level. The measurement level of the intervention (organization) differed from the measurement level of the outcomes (individual) in the effectiveness study in this thesis. The only organizational or 'collective' outcome measure was organizational efficacy (i.e. 'an individual's perception of the general capabilities of an organization' (p.127, [50])). This might be a suboptimal way of trying to capture the intervention effect, which might also contribute to an erroneous conclusion that the intervention did not render an effect (i.e. type II error). In future studies the measurement level of intervention and measures should preferably be kept the same by (at least) including constructs at the group level, such as team level performance, aggregated sickness absence registration data, or an assessment of interpersonal relations. Ideally, an analysis can be based on a combination of self-reports and organizations' (or even societies') routinely collected data.

Overlap between definitions organizational level intervention and primary preventive intervention

Organizational level interventions for work stress are defined as follows: "Planned actions designed to reach relatively large groups of individuals in a relatively uniform way by changing the way work is designed, organized or managed" (p.220,[51]). This definition overlaps with the definition of primary preventive interventions for work stress: "primary intervention aims to prevent the incidence of work-related mental health problems; it is 'work-directed' - aiming to reduce job stressors at their source by modifying the job or the work environment." (p.3, [52]; [53]). The classifications seem to have been used interchangeably in the literature, which can cause confusion about the interventions under study, the most suitable study design, and most appropriate

outcomes. Because, by (these) definition(s) organizational level interventions are always primary preventive, whereas primary preventive interventions are always conducted at the organizational level. I believe this classification is not completely justified, because organizational level interventions can also be secondary (or even tertiary) preventive (e.g. offering all managers a coaching course to teach them how to recognize stress symptoms), whereas primary preventive interventions can also be applied at the individual level (e.g. tailored coaching to prioritize tasks and thus manage job demands or utilize job resources). The 'individual/organizational level' distinction should be used for the interventions' level of application, and 'primary/secondary/tertiary' for the target group and its health status. Following this line of reasoning, it might be useful to further specify the 'individual/organizational level' distinction into four levels, as has been proposed in the organizational change literature [54] and used in the context of restructuring interventions [55]: 1) interventions directed at the individual, 2) interventions aimed at a group (e.g. team or department), 3) interventions targeting the managers or supervisors, and 4) interventions involving changes in organizational procedures and practices. The last parts of the definitions (i.e. "changing the way work is designed, organized or managed" and "reduce job stressors at their source by modifying the job or the work environment.") relate to the intervention strategies that are applied, which typically include changing (i) task characteristics, (ii) working conditions, and (iii) social conditions [56]. Creating more clarity about the intervention's level of application, the target group and its health status, and the intervention strategies applied will make it easier to choose the most suitable study design and appropriate outcomes.

Primary preventive intervention applied in a situation that required secondary prevention

The HM intervention (**chapter 3-5**) was presented to the schools as an organizational level primary preventive intervention, which is an intervention to eliminate the sources of stress by changing the way work is designed, organized, or managed. A study among 81 organizations demonstrated that those organizations with low levels of stress and balanced demands and resources had the best chance of successfully developing and implementing interventions [57]. However, both schools wanted to participate in the intervention project because work stress was already present in the intervention departments (amongst other problems). Because a substantial amount of employees already suffered from work stress at baseline, this might have led to employees' unreasonably high expectations of the possible effectiveness and scope of the intervention, as was found in earlier work stress interventions [58]. In the literature it has been proposed that primary and secondary preventive interventions should be combined in order to promote the positive and prevent harm, respectively [52]. Based on the current thesis I would like to add the nuance that such a combination indeed seems a good idea, however this should be done sequentially rather than simultaneously in highly stressed organizations. I believe it is not likely that a participatory, primary preventive, organizational

level intervention (that aims to alter the sources of stress at work by changing the way work is designed, organized or managed) will lead to a stress reduction in organizations that are stressed already. In these organizations, stressed employees and managers should rather be offered a secondary preventive intervention first (that aims to help individuals or groups cope effectively with existing stressors). Addressing the existing work stress complaints at both employee and middle manager level first, might create an atmosphere of readiness for change, after which the stressors at their source can be changed more efficiently and maybe even effectively. This order of intervening might prepare employees and managers for the change process that a primary preventive, organizational level intervention implies and shape their expectancies. Participating in a secondary, coping intervention requires employees' development of meta-analytical skills with regard to one's own thoughts, affects, and behaviors. Developing these skills might also prepare employees for their participating role in a primary intervention, which can be a new experience in sectors and organizations not used to shared decision-making.

Recommendations

In this section I formulate recommendations for future practice and research, by describing that conducting primary preventive, organizational level interventions often implies a change process, and by exploring how to meaningfully combine process and effect measures.

Recommendation for practice: A primary preventive, organizational level intervention is a change process that should be managed as such

Based on the experience with the intervention in this thesis I believe that conducting an organizational level, primary preventive intervention often implies managing an organizational change process, because the intervention activities are (by definition) changes in how the work is organized, designed, or managed. Change processes in organizations generally tend to be tough because they are slow [59], and fail often or make things worse [60]. But foremost they are hard because organizations are dynamic rather than static [61], and requirements of intervention activities could change almost on a day to day basis especially in 'stressed organizations'. Several recommendations can be formulated for practice, for example for intervention facilitators, HR-professionals, middle managers, and top level managers in schools and other organizations, to manage the change process.

- *Conduct an elaborate preparation phase prior to starting the change process* (the implementation of the intervention), to identify preexisting context conditions, which could be indicative of (lacking) readiness for change. Especially:

- Capture (by observations or recollection of previous experiences) and change (by cognitive behavioral techniques) what particular negative automatic thoughts employees might have with regard to the intervention.
- Shape expectancies by timely, adequate and tailored communication about the intervention. Induce positive states by making the intervention a fun thing to do.
- Assess cultural aspects (by mixed methods) that are indicative of leadership style (which preferably is democratic), and organizational culture (which preferably is employee-oriented).
- *Look before you leap.* Starting the change process but aborting it halfway (or after problem definition or action planning), is the most efficient way to make the change fail, and possibly future change processes as well. It is furthermore detrimental for the trust between managers and employees. Employees might learn that they are not taken seriously and they will become more cynical towards future changes.
- *Design participation in a manner that suits the organizational context.* The change process needs to be a participatory effort of both employees and staff. If employees might have difficulties with their participating role, efforts should be directed at making sure they could participate well. If management or supervisors might have difficulties with employees' participation, efforts should be directed at changing this attitude.

Recommendation for research: combine process and effect outcomes

The importance of measuring the implementation process in intervention studies is increasingly recognized (e.g. [51]). An evaluation of the implementation process is often conducted in a separate process evaluation analysis, alongside a (randomized) trial, in order to answer research questions such as 'what factors hindered and helped the implementation?'. The analysis typically relies on both qualitative (e.g. interviews) and quantitative measures (e.g. questionnaires), which are operationalized according to common implementation models such as Steckler and Linnan [62] or RE-AIM [63]. Reports of the effect and process analyses are often provided in different articles, thereby complicating the possibility to integrate the findings from the separate analyses. An example of such a process evaluation is the implementation of a worksite social and physical environmental intervention in order to reduce need for recovery (amongst other variables)[64].

Researchers have pointed to the relevance of combining both process and outcome variables [51]. Combining these measures supposedly leads to a greater understanding of what works for whom under which circumstances. Whereas many have pointed to the relevance of this type of research, not many have tried to do this: only 39 of the 84 studies (46%) included in a review on process components in stress management interventions [65].

Nevertheless, some examples of combining process and effect measures in intervention evaluations are present in the literature. In the following I will use some of these intervention evaluations as illustrations of ways to further explore the combination of process and effectiveness variables in future studies. These explorative ways are: (i) subgroup analyses wherein allocation to subgroup is based on process data, (ii) studying an implementation component as a predictor of the effectiveness outcome, (iii) integrating process and effect measures. The statistical procedures applied differ, although the type of research is the same. The three ways are thus considered variations on the same theme.

Firstly, researchers have conducted regular *subgroup analyses*, in epidemiology also referred to as per protocol analyses (i.e. comparing that part of the intervention group that followed the intervention protocol more rigorously, to the control group). If a small effect of the intervention exists, it would be found more easily in a group that received a 'higher dose' of the intervention. For example, in a pragmatic RCT of a web-based intervention aiming to empower disability claimants it was found that 33% of the target population did not even log onto the website [66]. Therefore, per protocol analyses were performed, comparing subgroups of participants who used the intervention for more than 1, 2, 3 or 4 hours, on all outcomes [67]. The researchers found significant dose-response relationships for context-specific empowerment, knowledge, coping, and claimant's active participation. The strength of the relationship increased with the hours spent on the intervention. Also in the current thesis, per protocol analyses showed a significant effect on occupational self-efficacy among those participants in the intervention group who received a higher dose of the first phase of the intervention, indicating that an intervention effect might have been found if the dose would have been high enough for all participants.

Secondly, researchers have used implementation components as a *predictor of the primary outcome* in two manners: 1) using the intervention dosage as a predictor, 2) using another specific known process component as a predictor.

As an example of the first manner, the effect of a problem solving intervention delivered by occupational physicians and targeted at employees who returned to work after experiencing common mental disorders was tested in a cluster RCT design [68]. The primary effect outcome was the incidence of recurrent sickness absence (yes/no). The process components of the Steckler and Linnan model [62] were assessed, and participant and physician satisfaction with the intervention was added. The researchers combined process and effect variables by conducting multilevel logistic regression analysis with recurrent sickness absence as dependent variable and participation in each of the intervention components as independent variables. The analyses showed specific effects for specific groups: a lower risk of recurrent sickness absence was found among participants who participated in two specific intervention activities (i.e. filling out the problem solving inventory and having discussed with the physician how to realize work opportunities). A higher risk was found

among participants who participated in only one specific activity (i.e. the problem solving inventory). Participating in more elements of the intervention per se did not result in a larger effect on the primary outcome.

As an example of the second manner, researchers quantitatively assessed specific known implementation components to assess variance in the outcomes under study [69]. The components were all related to perception or appraisal of the implementation process, the importance of which has been described earlier (paragraph 'main findings in context'). More specifically, in a teamwork intervention in a Danish elderly care setting, Randall and colleagues [69] used employee perceptions of participation, line manager attitudes and actions, perceived exposure, and readiness for change to explain variance in the outcomes (i.e. self-efficacy, job satisfaction, well-being). Regression analyses demonstrated firstly that line manager attitudes and actions were associated with higher post-intervention levels of self-efficacy and job satisfaction. Secondly, employee readiness for change was associated with higher levels of self-efficacy and intervention history was associated with job satisfaction. All betas were in positive direction, indicating that more positive appraisals of the process were related to more positive outcomes. This could be a ground for measuring these process components upfront, in order to either adapt the planned intervention or to conduct another intervention first (namely increasing line manager attitudes and actions, or employee readiness for change).

Thirdly, researchers have *integrated process and effect measures*. In a Swiss primary intervention to reduce stress among 1400 workers in eight diverse medium and large sized companies in industry and services, participants were asked (among many other variables) to retrospectively assess the impact the intervention had within their organization [47]. This 'impact assessment' of the intervention was then related to longitudinal changes in the ratio between job resources and job demands, the (proximal) intervention outcome. Then the intervention participants were, again retrospectively, assigned to the high impact subgroup or the low/medium subgroup. Repeated General Linear Model analysis showed that those who appraised the intervention as having had 'high impact' within the organization, also improved or maintained a certain positive resources/demands ratio after two years. The one exception to this general finding were managers who already had a favorable resources demands ratio.

Looking over these options, I believe the first can be quite easily applied to ongoing intervention studies, or even in studies that have finished already. The second option would require more literature research in order to formulate specific hypotheses upfront, and it furthermore requires good, validated measures (which could be based on for example Randall, Nielsen and Tvedt's Intervention Process Measure [20]). The third option is interesting because it also dealt with one of the definition problems with regard to stress (a possible result of an imbalance between job demands and job resources) by computing a ratio. I conclude that several options for combining process and effect data exist, and are feasible to apply. However, a systematic review on complex social interventions

found that the quality of reporting of implementation variables was often poor and anecdotal [70], which makes the combination of process and effect difficult. Adequate reporting of implementation variables deserves more attention in research. Ideally, process components are measured upfront as part of an elaborate preparatory phase, in order to either adapt the planned intervention or to conduct another intervention first (e.g. increasing line manager attitudes and actions, and employee readiness for change). Conducting combined effect-process analyses can help to obtain better interventions, smoother intervention processes, and disclosure of the 'black box' of intervention mechanisms.

General conclusion

The main objective of this thesis was to explore ways to decrease work stress in teachers from an individual, organizational and intervention evaluation perspective.

From an individual perspective work stress in teachers might be decreased by addressing mastery in secondary preventive, individual level interventions by using cognitive behavioral techniques, although this finding requires replication and experimentation.

From an organizational perspective work stress cannot be decreased by the currently applied participatory, primary preventive, organizational level intervention. The intervention was ineffective in reducing (determinants of) work stress. Furthermore, the implementation process of the intervention was hindered by an insufficient implementation strategy for the intervention activities, insufficient communication, and unexpected events. It was argued that employee and organizational readiness for change was low, due to employee learned helplessness, employees' negative appraisal of the intervention, an authoritarian leadership style, and an organizational culture low in employee-orientation. These characteristics also did not fit the nature of a participatory intervention. Future interventions of this type should conduct a more elaborate preparatory step, as to establish these factors beforehand.

The most relevant evidence in intervention studies can be obtained by gathering both effect and process data. The most relevant effect data are obtained when researchers consider beforehand the specific research question, the complexity of the intervention, the availability of existing data, the context, and the costs and then choose the most appropriate research design. Other designs than the RCT are available and feasible, of which the stepped-wedge randomized trial and the multiple baseline design seem especially suitable to primary preventive, organizational level interventions. Although every research design poses its own challenges, such as the increased number of participants and measurement moments that are needed. The most relevant process data are obtained based on a framework that includes variables related to the intervention, participants' mental models, and context factors. Ideally, process and effect data are combined in several subgroup analyses.

References

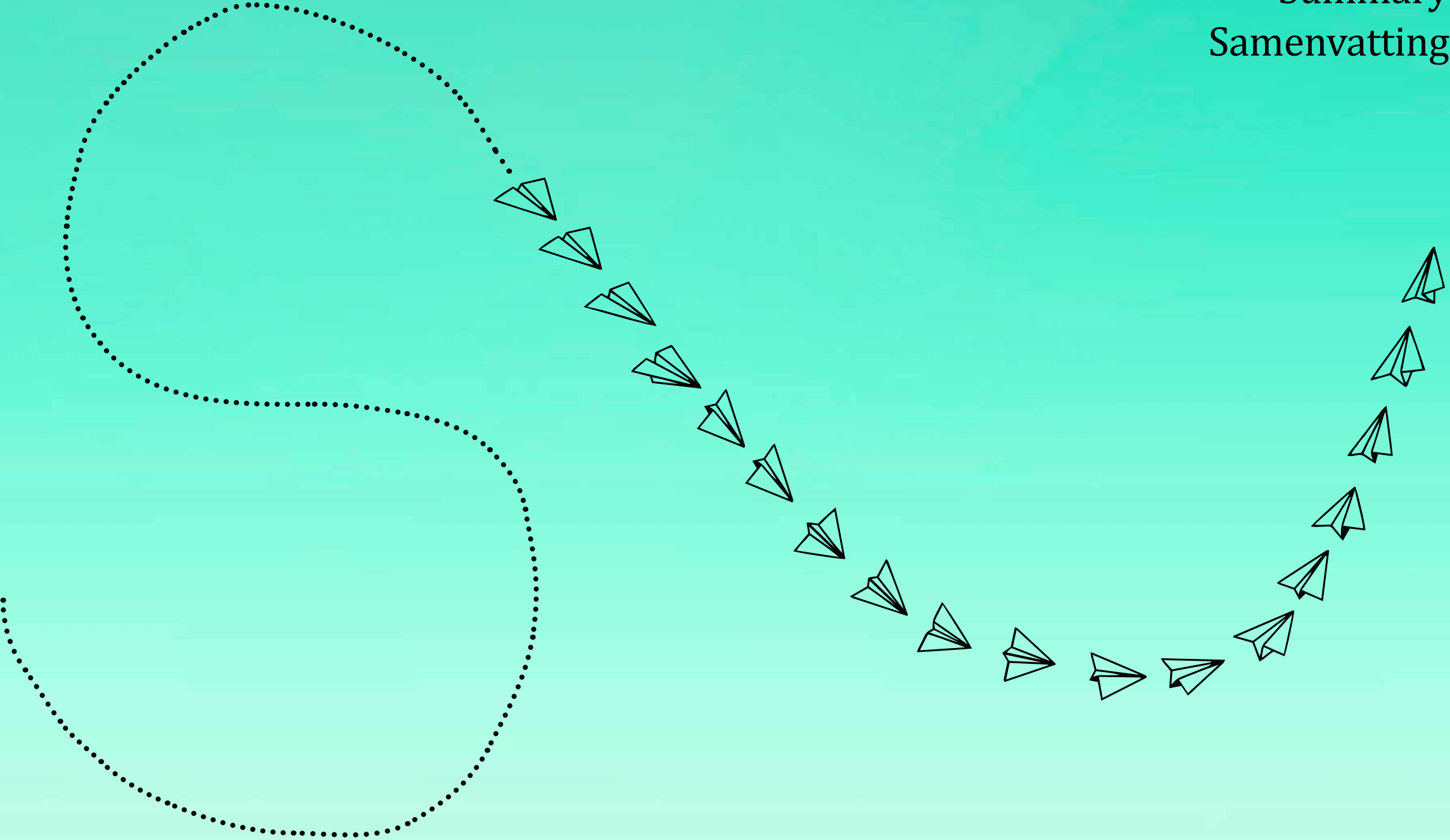
- Cox T, Griffiths A, Randall R: **Chapter 10: A risk management approach to the prevention of work stress.** In *The Handbook of Work and Health Psychology. Volume 191*. 2nd edition. Edited by Schabracq MJ, Winnubst JAM, Cooper CL. Chichester, West Sussex: John Wiley & Sons Ltd; 2003:191-206.
- Leka S, Jain A, Cox T, Kortum E: **The development of the European framework for psychosocial risk management: PRIMA-EF** *Journal of occupational health* 2011, **53**(2):137-143.
- World Health Organization: **PRIMA-EF: guidance on the European framework for psychosocial risk management: a resource for employers and worker representatives.** 2008, **9**.
- Nielsen K, Randall R, Holten A-, González ER: **Conducting organizational-level occupational health interventions: What works?** *Work Stress* 2010, **24**(3):234-259.
- Kompier M, Marcelissen F: **Handboek werkstress [Handbook of work stress]**. Amsterdam, The Netherlands: NIA 1990.
- Weiner BJ: **A theory of organizational readiness for change.** *Implementation science* 2009, **4**(1):67.
- Armenakis AA, Harris SG, Mossholder KW: **Creating readiness for organizational change.** *Human relations* 1993, **46**(6):681-703.
- Weiner BJ, Amick H, Lee SY: **Conceptualization and measurement of organizational readiness for change: a review of the literature in health services research and other fields.** *Med Care Res Rev* 2008, **65**(4):379-436.
- Griffith J: **Relation of principal transformational leadership to school staff job satisfaction, staff turnover, and school performance.** *Journal of educational administration* 2004, **42**(3):333-356.
- Barnett K, McCormick J: **Leadership and individual principal-teacher relationships in schools.** *Educational Administration Quarterly* 2004, **40**(3):406-434.
- Seligman ME: **Learned helplessness.** *Annu Rev Med* 1972, **23**(1):407-412.
- Maier SF, Seligman ME: **Learned helplessness: Theory and evidence.** *J Exp Psychol: Gen* 1976, **105**(1):3.
- Weiner B: **An attributional theory of achievement motivation and emotion.** In *An attributional theory of motivation and emotion*. Springer; 1986:159-190.
- Timms C, Graham D, Cottrell D: **"I just want to teach": Queensland independent school teachers and their workload.** *Journal of Educational Administration* 2007, **45**(5):569-586.
- Fink D: **The law of unintended consequences: the 'real' cost of top-down reform.** *Journal of Educational Change* 2003, **4**(2):105-108.
- Timms C, Graham D, Caltabiano M: **Perceptions of school administration trustworthiness, teacher burnout/job stress and trust: the contribution of morale and participative decision-making.** In *Advances in Organisational Psychology*. Edited by Anonymous Bowen Hills, QLD, Australia: Australian Academic Press; 2007:135-151.
- Parlementaire Commissie Onderwijsvernieuwingen o.l.v. Dijsselbloem, J.R.V.A.: **Parlementair Onderzoek Onderwijsvernieuwingen: Tijd voor Onderwijs.** Vergaderjaar 2007-2008, **31007, nr. 6**.
- Beck AT: *Cognitive therapy of depression*: Guilford press; 1979.
- Butler AC, Chapman JE, Forman EM, Beck AT: **The empirical status of cognitive-behavioral therapy: a review of meta-analyses.** *Clin Psychol Rev* 2006, **26**(1):17-31.
- Randall R, Cox T, Griffiths A: **Participants' accounts of a stress management intervention.** *Human Relations* 2007, **60**(8):1181-1208.
- Nielsen K, Randall R, Albertsen K: **Participants' appraisals of process issues and the effects of stress management interventions.** *J Organ Behav* 2007, **28**(6):793-810.
- Arnold M, Finnis DG, Kerridge I: **Medicine's inconvenient truth: the placebo and nocebo effect.** *Intern Med J* 2014, **44**(4):398-405.
- Bensing JM, Verheul W: **The silent healer: the role of communication in placebo effects.** *Patient Educ Couns* 2010, **80**(3):293-299.
- Nytrø K, Saksvik PØ, Mikkelson A, Bohle P, Quinlan M: **An appraisal of key factors in the implementation of**

- occupational stress interventions.** *Work & Stress* 2000, **14**(3):213-225.
25. Saksvik PØ, Nytrø K, Dahl-Jørgensen C, Mikkelsen A: **A process evaluation of individual and organizational occupational stress and health interventions.** *Work & Stress* 2002, **16**(1):37-57.
26. Mathieu JE, Heffner TS, Goodwin GF, Cannon-Bowers JA, Salas E: **Scaling the quality of teammates' mental models: Equifinality and normative comparisons.** *J Organ Behav* 2005, **26**(1):37-56.
27. Nielsen K, Randall R: **Opening the black box: Presenting a model for evaluating organizational-level interventions.** *European Journal of Work and Organizational Psychology* 2013, **22**(5):601-617.
28. Harrington A: *The placebo effect: An interdisciplinary exploration:* Harvard University Press; 1999.
29. Aust B, Rugulies R, Finken A, Jensen C: **When workplace interventions lead to negative effects: Learning from failures.** *Scand J Public Health* 2010, **38**(SUPPL. 3):106-119.
30. Wasan AD, Kaptchuk TJ, Davar G, Jamison RN: **The association between psychopathology and placebo analgesia in patients with discogenic low back pain.** *Pain Medicine* 2006, **7**(3):217-228.
31. Eagly AH, Johannesen-Schmidt MC, Van Engen ML: **Transformational, transactional, and laissez-faire leadership styles: a meta-analysis comparing women and men.** *Psychological Bulletin* 2003, **129**(4):569-591.
32. Leithwood K, Jantzi D: **The effects of transformational leadership on organizational conditions and student engagement with school.** *Journal of Educational Administration* 2000, **38**(2):112-129.
33. Goleman D: **Leadership that gets results.** *Harv Bus Rev* 2000, **78**(2):4-17.
34. Chamberlin CR: **Nonverbal behaviors and initial impressions of trustworthiness in teacher-supervisor relationships.** *Communication Education* 2000, **49**(4):352-364.
35. Oyer BJ: **The relationship between principals' confidence, humility, and effectiveness: a study of teacher perceptions.** *Doctor of Education.* Bowling Green State University, 2011.
36. Hofstede G, Neuijen B, Ohayon DD, Sanders G: **Measuring organizational cultures: A qualitative and quantitative study across twenty cases.** *Adm Sci Q* 1990, :286-316.
37. Jones RA, Jimmieson NL, Griffiths A: **The impact of organizational culture and reshaping capabilities on change implementation success: The mediating role of readiness for change.** *Journal of Management Studies* 2005, **42**(2):361-386.
38. Borg M, Waisfisz B, Frank U: **Quantitative assessment of organizational culture within hospitals and its relevance to infection prevention and control strategies.** *J Hosp Infect* 2015, **90**(1):75-77.
39. Goldenhar LM, LaMontagne AD, Katz T, Heaney C, Landsbergis P: **The intervention research process in occupational safety and health: an overview from the National Occupational Research Agenda Intervention Effectiveness Research team.** *J Occup Environ Med* 2001, **43**(7):616-622.
40. Nielsen K, Randall R, Holten A-, González ER: **Conducting organizational-level occupational health interventions: What works?** *Work Stress* 2010, **24**(3):234-259.
41. Shalit B: **Structural ambiguity and limits to coping.** *J Human Stress* 1977, **3**(4):32-45.
42. Thibaut J, Walker L: **A theory of procedure.** *Calif Law Rev* 1978, :541-566.
43. Van Eerd D, Cole D, Irvin E, Mahood Q, Keown K, Theberge N, Village J, St. Vincent M, Cullen K: **Process and implementation of participatory ergonomic interventions: a systematic review.** *Ergonomics* 2010, **53**(10):1153-1166.
44. Driessen MT, Proper KI, Anema JR, Knol DL, Bongers PM, van der Beek AJ: **Participatory ergonomics to reduce exposure to psychosocial and physical risk factors for low back pain and neck pain: results of a cluster randomised controlled trial.** *Occup Environ Med* 2011, **68**(9):674-681.
45. Oude Hengel KM, Joling CI, Proper KI, van der Molen H, Bongers PM: **Intervention mapping as a framework for developing an intervention at the worksite for older construction workers.** *American Journal of Health Promotion* 2011, **26**(1):e1-e10.
46. Richardson KM, Rothstein HR: **Effects of Occupational Stress Management Intervention Programs: A Meta-Analysis.** *J Occup Health Psychol* 2008, **13**(1):69-93.
47. Jenny GJ, Brauchli R, Inauen A, Fullemann D, Fridrich A, Bauer GF: **Process and outcome evaluation of an organizational-level stress management intervention in Switzerland.** *Health Promot Int* 2015, **30**(3):573-585.
48. Fridrich A, Jenny GJ, Bauer GF: **The context, process, and outcome evaluation model for organisational health interventions.** *BioMed research international* 2015, :1-12.
49. Thiele Schwarz U, Lundmark R, Hasson H: **The dynamic integrated evaluation model (DIEM): achieving sustainability in organizational intervention through a participatory evaluation approach.** *Stress Health* 2016, **32**(4):285-293.
50. Van Vuuren M, Elving WJ, De Jong MDT, Seysel ER: **What keeps you here? Organizational efficacy and person-organization fit as antecedents organizational commitment.** In *Why work? The contribution of value perceptions and efficacy expectations to organizational commitment.* Edited by Van Vuuren HA. 2006:121-147.
51. Nielsen K, Taris TW, Cox T: **The future of organizational interventions: Addressing the challenges of today's organizations.** *Work Stress* 2010, **24**(3):219-233.
52. LaMontagne AD, Martin A, Page KM, Reavley NJ, Noblet AJ, Milner AJ, Keegel T, Smith PM: **Workplace mental health: developing an integrated intervention approach.** *BMC Psychiatry* 2014, **14**(1):1.
53. Hargrove MB, Quick JC, Nelson DL, Quick JD: **The theory of preventive stress management: a 33-year review and evaluation.** *Stress Health* 2011, **27**(3):182-193.
54. Anderson DL: *Organization development: The process of leading organizational change:* Sage Publications; 2013.
55. Wiezer N, Nielsen K, Pahkin K, Widerszal-Bazyl M, de Jong T, Mattila-Holappa P, Mockallo Z: *Exploring the link between restructuring and employee well-being:* Warsaw: Central Institute for Labour Protection - National Research Institute; 2011.
56. Semmer NK: **Job stress interventions and the organization of work.** *Scandinavian Journal of Work, Environment and Health* 2006, **32**(6):515-527.
57. Taris TW, Kompier MA, Geurts SA, Schreurs PJ, Schaufeli WB, de Boer E, Sepmeijer KJ, Watterz C: **Stress management interventions in the Dutch domiciliary care sector: Findings from 81 organizations.** *International Journal of Stress Management* 2003, **10**(4):297.
58. Bourbonnais R, Brisson C, Vinet A, Vezina M, Abdous B, Gaudet M: **Effectiveness of a participative intervention on psychosocial work factors to prevent mental health problems in a hospital setting.** *Occup Environ Med* 2006, **63**(5):335-342.
59. Amburgey TL, Kelly D, Barnett WP: **Resetting the clock: The dynamics of organizational change and failure.** *Academy of Management Proceedings;* Academy of Management; 1990.
60. Beer M, Eisenstat RA, Spector B: **Why change programs don't produce change.** *Harvard business review* 1990, **November-December.**
61. Griffiths A: **Organizational interventions: facing the limits of the natural science paradigm.** *Scand J Work Environ Health* 1999, **25**(6):589-596.
62. Steckler A, Linnan L: *Process evaluation. For public health intervention and research:* San Francisco: Jossey-Bass; 2002.
63. Glasgow RE, Vogt TM, Boles SM: **Evaluating the public health impact of health promotion interventions: the RE-AIM framework.** *Am J Public Health* 1999, **89**(9):1322-1327.
64. Coffeng JK, Hendriksen IJM, Van Mechelen W, Boot CRL: **Process evaluation of a worksite social and physical environmental intervention.** *Journal of Occupational and Environmental Medicine* 2013, **55**(12):1409-1420.
65. Murta SG, Sanderson K, Oldenburg B: **Process evaluation in occupational stress management programs: a systematic review.** *Am J Health Promot* 2007, **21**(4):248-254.
66. Samoocha D, Snels IA, Bruinvels DJ, Anema JR, van der Beek A: **Effectiveness of an interactive website aimed at empowerment of disability benefit claimants: results of a pragmatic randomized controlled trial.** *J Occup Rehabil* 2011, **21**(3):410-420.
67. Samoocha D, Snels IA, Bruinvels DJ, Anema JR, van der Beek AJ: **Chapter 7: Efficacy, dose-response relationships and predictors of compliance in a web-based intervention aimed at**

empowerment of disability benefit claimants, in: Empowerment of Disability Benefit Claimants prior to their Disability Assessment. VUmc, Amsterdam; 2011

68. Arends I, Bruinvels DJ, Rebergen DS, Nieuwenhuijsen K, Madan I, Neumeyster-Gromen A, Bultmann U, Verbeek JH: **Interventions to facilitate return to work in adults with adjustment disorders.** *Cochrane Database Syst Rev* 2012, **12**:CD006389.
69. Randall R, Nielsen K, Tvedt SD: **The development of five scales to measure employees' appraisals of organizational-level stress management interventions.** *Work Stress* 2009, **23**(1):1-23.
70. Egan M, Bambra C, Petticrew M, Whitehead M: **Reviewing evidence on complex social interventions: appraising implementation in systematic reviews of the health effects of organisational-level workplace interventions.** *J Epidemiol Community Health* 2009, **63**(1):4-11.

Summary
Samenvatting



Summary

Good education advances our knowledge-based society. Good education will become excellent education if the general level of teaching is improved. And the level of teaching will more easily improve if the school organization is functioning well. However, this improvement in the level of teaching is challenged by the high level of work stress among teachers. Throughout the Western and Eastern developed world, the educational sector is (one of the) front runner(s) in the experience of work stress. European estimates of the stress levels of teachers more than doubled those found in other occupations. Work stress is likely to result in several mental and physical health problems, negative organizational outcomes, and societal costs. These substantial consequences are especially alarming in the light of an already shrinking workforce. Researchers have posed that some of these consequences might be prevented if adequate measures are taken. However, the evidence on the most effective ways to decrease work stress in teachers is inconclusive. To advance our understanding of decreasing work stress, more knowledge on the individual and organizational determinants of work stress for specific occupational groups is needed. Furthermore, methodological issues need to be addressed when implementing interventions, such as determining causality. The objective of this thesis was to explore ways to decrease work stress in teachers, and thereby eventually contribute to excellent education (**chapter 1**). The exploration was conducted along three key questions: (1) How can we decrease work stress in teachers from an individual perspective? (2) How can we decrease work stress in teachers from an organizational perspective? (3) How can we gather the most relevant evidence in intervention studies in the occupational setting, for example to decrease work stress?

How can we decrease work stress in teachers from an individual perspective?

The individual perspective was explored by conducting a longitudinal study with three waves, among 549 older teachers (aged 45-64 years) (**chapter 2**). In this study the role of mastery was studied, as a personal resource for decreasing work stress. Path analysis was applied to explore the interplay between mastery and job demands (i.e. psychological demands and emotional demands) and job resources (i.e. autonomy and social support), in influencing work stress-related outcomes (i.e. depression and work engagement).

Psychological job demands contributed to an increase in mastery; emotional job demands contributed to a decrease in mastery. Psychological demands at baseline (T1) were related to both an increase in depression at T3 (via T2 job demands) and a decrease in depression (via T2 mastery), but not to changes in work engagement. Emotional demands (T1) were related to an increase in depression (T3) and a decrease in work engagement (T3) (partially via T2 mastery). These findings suggest that the type of job demand matters. A job

demand can be either a challenge (a chance for personal growth or development), a hindrance (an obstacle to growth or accomplishment), or a threat (a risk of personal harm or loss). Labeling job demands as challenge, hindrance or threat can help to determine the type of intervention that is best to apply in order to achieve optimal individual and organizational outcomes. Contrary to expectations, job resources were neither longitudinally related to mastery nor to depression or work engagement, and therefore no mediation by mastery was found.

Finally, mastery (T2) was directly related to lower depression and higher work engagement one year later (T3). Increasing mastery in older teachers can be explored further as an intervention strategy to decrease depression and increase work engagement in this occupational group. The findings require replication and actual experimentation.

How can we decrease work stress in teachers from an organizational perspective?

In order to decrease work stress in teachers, a participatory, primary preventive, organizational level intervention was conducted among the employees of two secondary vocational education schools in the Netherlands (school A and school B). The intervention was expert-based, since it was developed by a consultancy firm and applied over a hundred times in both public and private organizations. However, it was never evaluated scientifically. As described in detail in **chapter 3**, the intervention comprised a needs assessment phase (consisting of interviews, a questionnaire, and group sessions) and a phase of implementing intervention activities, which were derived from the needs assessment phase (e.g. structured performance reviews, or creating a staff room). A facilitator supervised the needs assessment phase and he was assisted by a participatory group of employees and staff members. In the phase of implementation of intervention activities the facilitator's help was optional, and only one of the two schools (school A) purchased this aid.

The primary hypothesis was that the intervention would decrease work stress by decreasing need for recovery and increasing vitality (**chapter 4**). The hypothesis was tested in a controlled trial with 356 employees in two schools. The intervention group consisted of 204 employees, in 2 departments, within 24 teams, and the control group consisted of 152 employees, in 2 departments, within 24 teams. Mixed model analysis was applied in order to assess differences between the intervention and control group on average over time. All analyses were corrected for baseline values and several covariates (e.g. gender, educational level). The hypothesis could not be confirmed as no statistical differences were found between the intervention and control group on need for recovery and vitality. Two small but statistically significant effects in unfavorable direction were found on two of the secondary outcomes (i.e. absorption and organizational efficacy). However, post-hoc per protocol analyses demonstrated that employees who participated in at least two of the three steps of the

needs assessment (e.g. questionnaire and group session) had a significantly higher post-intervention score on occupational self-efficacy (i.e. the belief in one's ability at work) than those in the control group, which is favorable. This result indicates that employees with a higher participation grade regarding the intervention might have had a mastery experience, which then positively affected occupational self-efficacy.

The process evaluation in **chapter 5** describes that the lack of positive effects on the primary outcomes could be partly due to implementation failure. Most process evaluations mainly studied the implementation process and the quality of the implementation (fidelity). However, in adopting this approach for participatory organizational level occupational health interventions, important aspects, such as context and participants' perceptions, are missing. Therefore, for this intervention study a process evaluation framework was applied covering aspects of the intervention and its implementation as well as the context and participants perceptions. The process evaluation revealed that the needs assessment phase was implemented successfully in school A, but not in school B where participation and readiness for change were insufficient. The results for school A and B diverged further in the phase of implementation of intervention activities: in school A several intervention activities were implemented, whereas this was hardly the case in school B. Regardless of actual participation and actual implementation of intervention activities, participants in both schools felt not involved in the choice of intervention activities. Furthermore, in both schools participants generally perceived the intervention's second phase negatively, for example because the intervention activities were not seen as the right solutions for the problems (school A). Three types of context factors further inhibited the implementation. That is, the schools had no capacity (school B) or only partial capacity (school A) to implement the intervention activities, foremost because middle managers were not fully equipped to do so. Secondly, the organizational culture in both schools was characterized by a lack of mutual trust between managers and teachers, which hindered implementation of this participatory intervention. Thirdly, several external events interfered with the planned intervention activities (e.g. a national change was announced by the Ministry of Education, implying an intensification of classroom-bound lessons from 850 to 1000 hours per year).

Based on chapters 4 and 5 it is concluded that further implementation of the intervention in its current form is not recommended. The general discussion (chapter 8) describes how the intervention could be changed in order to better fit the organizational context in schools.

How can we gather the most relevant evidence in intervention studies?

Designing evaluations of occupational health interventions can be challenging with regard to methodology (e.g. difficulties with randomization and control group), the intervention itself (e.g. the organization wants to adjust the inter-

vention protocol), and the context (e.g. new policy interferes with the intervention). The current intervention study also suffered from design challenges. For example, both schools wanted to participate in the study with a specific department, for which they believed the intervention would be most relevant. Although conceivable from the school perspective, this posed a methodological issue because randomization procedures were thereby obstructed. A controlled trial was conducted, making the effect evaluation prone to confounding and selection bias. As a consequence, baseline differences on several outcomes were present between the intervention and control group. The narrative review in **chapter 6** describes several of these challenges occupational health intervention researchers encountered and outlines research designs that solve (some of) them. In the review we furthermore explored alternative research designs for the situation wherein a regular Randomized Controlled Trial (RCT) is not feasible. One experimental alternative design (i.e. stepped wedge design) and several observational designs were found and discussed (i.e. propensity scores, instrumental variables, multiple baseline design, interrupted time series, difference-in-difference, and regression discontinuity).

Future primary preventive, participatory, organizational level interventions could benefit especially from the stepped wedge randomized design whereby the intervention is randomly applied to all groups but sequentially over time, or the multiple baseline design, which does the same although not randomly. Of course these designs pose their own challenges, such as the increased number of participants and measurement moments that are needed. In order to gather the most relevant evidence in intervention studies, researchers should consider beforehand the specific research question, the complexity of the intervention, the availability of existing data, the context, and the costs before choosing the most appropriate design.

Another issue in gathering the most relevant evidence based on intervention studies is related to the design of process evaluations. The field of process evaluation is rapidly developing and we noticed that many different theoretical frameworks exist, which include many overlapping process variables. In order to understand the most relevant evidence with regard to the implementation process, a systematic literature review of 44 process evaluations of organizational level stress management interventions was conducted (**chapter 7**). The review demonstrated that there is indeed great heterogeneity in methods and process variables studied: among the 44 evaluations, 47 different process variables were found. Also a theoretical framework or program theory, which could guide the choice of process variables, was missing in half of the articles. Furthermore, the process evaluation data were mostly collected after the intervention took place, instead of also before and during the intervention. We concluded that a standardized framework including variables with regard to the intervention, participants' mental models, and context factors would improve the theory development in the field of process evaluation for stress management interventions.

General discussion

In the general discussion (**chapter 8**) the main findings of this thesis are discussed from an individual, organizational, and intervention evaluation perspective. The organizational perspective is discussed in most detail. The expert-based approach evaluated in this part of the thesis is compared to the evidence-based approach. Several explanations are provided for the lack of positive effects. Firstly, I argue that employees' participation in the intervention was hindered by a phenomenon known as 'learned helplessness' (i.e. refraining to respond to (and thus influence) an event, while it could have been controlled by our own actions). Secondly, I describe that the intervention did not include a strategy to address employees' perceptions and appraisal. Thirdly, leadership and organizational culture did not enable the (implementation of the) intervention. Based on this discussion it is proposed that the degree of participation in interventions should be tailored to the context, whereby feasibility and suitability are leading principles. Furthermore, a more elaborate 'preparation phase' is proposed.

Next, several considerations with regard to methodology are discussed. Firstly, methodological difficulties due to the definition of work stress are provided, that is: (i) stress as the result of an imbalance between job demands and job resources; (ii) stress as a stage in a process rather than a definite outcome; and (iii) the measurement level of stress. It is proposed, respectively: (i) to actually calculate the imbalance in order to develop a more sensitive measure of work stress determinants instead of using proxies for stress or endpoints in the stress process, (ii) to develop a logic model of expected changes for an intervention study, distinguishing proximal, intermediate, and distal effectiveness outcomes; and (iii) to keep the measurement level of a group intervention and the measures the same by (at least) including constructs at the group level (e.g. aggregated sickness absence registration data, team level performance). Secondly, in the considerations paragraph of the general discussion it is discussed that an overlap exists between the definitions for organizational level intervention and for primary preventive intervention, which contributes to lack of clarity. Finally, it is described that in the current study a primary preventive intervention was applied, where secondary prevention would have been more appropriate.

The general discussion concludes with recommendations for practice and research. The main recommendation for practice is that a primary preventive, organizational level intervention is a change process that should be managed as such. This can be done by conducting a thorough preparation phase, to ensure commitment prior to the start, and to design participation in a manner that suits the organizational context. The main recommendation for research is to combine process and effect measures in future intervention studies, because this supposedly leads to a greater understanding of what works for whom under which circumstances.

Samenvatting

Goed onderwijs is de drijfveer van onze kennismaatschappij. Goed onderwijs kan excellent onderwijs worden als het niveau van doceren verbetert. Het niveau van doceren zal gemakkelijker verbeteren als een schoolorganisatie goed functioneert. Het verbeteren van het docerniveau wordt echter bemoeilijkt door de mate van stress waar docenten mee kampen. In geïndustrialiseerde landen is de onderwijssector (één van de) koploper(s) in de mate waarin werkstress ervaren wordt. Europese schattingen van het stressniveau van docenten zijn twee keer zo hoog als de stressniveaus in andere beroepen.

Werkstress leidt veelal tot verscheidene mentale en fysieke gezondheidsproblemen, maar ook tot negatieve organisatieuitkomsten (bijv. uitval uit werk) en verhoogde maatschappelijke kosten. Deze gevolgen zijn vooral alarmerend in het licht van een krimpende beroepsbevolking. Onderzoekers hebben gesteld dat sommige van deze gevolgen mogelijk voorkomen kunnen worden als adequate maatregelen worden getroffen. Echter, het wetenschappelijke bewijs voor de meest effectieve manieren om werkstress bij docenten te verminderen is niet sluitend. Om beter te begrijpen hoe we werkstress kunnen verminderen, is meer kennis over individuele en organisatorische determinanten van werkstress nodig voor specifieke beroepsgroepen waaronder medewerkers in het onderwijs. Ook zijn er nog methodologische problemen op te lossen, zoals het vaststellen van causaliteit bij de implementatie van interventies voor werknemersgezondheid. De doelstelling van dit proefschrift was het verkennen van manieren om werkstress bij docenten te verminderen en zo uiteindelijk bij te dragen aan excellent onderwijs (**hoofdstuk 1**). Drie kernvragen waren leidend in deze verkenning: (1) Hoe kunnen we werkstress bij docenten verminderen vanuit individueel perspectief? (2) Hoe kunnen we werkstress bij docenten verminderen vanuit organisatieperspectief? (3) Hoe kunnen we het meest relevante bewijs verzamelen in interventieonderzoek in de werkcontext?

Hoe kunnen we werkstress bij docenten verminderen vanuit individueel perspectief?

Het individuele perspectief werd verkend in een longitudinaal onderzoek met drie meetmomenten, onder 549 oudere docenten (45-64 jaar) (**hoofdstuk 2**). In dit onderzoek is de rol van *mastery* (nl.: gegeneraliseerd zelfvertrouwen, resulterend in de overtuiging dat men omstandigheden die van invloed zijn op het leven gewoonlijk wel aan kan) bestudeerd als een persoonlijke hulpbron om werkstress te verminderen. Pad-analyse is toegepast om de wisselwerking te verkennen tussen *mastery* enerzijds en taakeisen (nl.: psychologische taakeisen en emotionele taakeisen) en hulpbronnen anderzijds (nl.: autonomie en sociale steun), evenals hun invloed op werkgerelateerde stressuitkomsten (nl.: depressieve klachten en bevlogenheid).

Psychologische taakeisen droegen bij aan een toename van mastery, terwijl emotionele taakeisen bijdroegen aan het verminderen van mastery. Psychologische taakeisen waren op baseline (T1) gerelateerd aan zowel een toename van depressieve klachten op T3 (via T2 - taakeisen) als een afname in depressieve klachten (via T2 - mastery), maar niet aan veranderingen in bevologenheid. Emotionele taakeisen (T1) waren gerelateerd aan een toename van depressieve klachten (T3) en een vermindering van bevologenheid (T3) (gedeeltelijk via T2 - mastery). Deze bevindingen suggereren dat het type taakeis ertoe doet. Een taakeis kan een uitdaging zijn (een kans op persoonlijke groei of ontwikkeling), een hindernis (een belemmering voor groei of prestatie), of een bedreiging (een risico op persoonlijk letsel of verlies). Vaststellen of taakeisen een uitdaging, hindernis of bedreiging vormen kan helpen bepalen welk type interventie het beste kan worden toegepast om optimale individuele en organisatorische uitkomsten te bereiken. In tegenstelling tot de verwachtingen waren de hulpbronnen niet longitudinaal gerelateerd aan mastery, depressieve klachten of bevologenheid, wat een mogelijk mediatie-effect van mastery uitsluit.

Tenslotte was mastery (T2) direct gerelateerd aan minder depressieve klachten en meer bevologenheid een jaar later (T3). Het vergroten van mastery bij oudere docenten kan verder worden verkend als een interventiestrategie om depressieve klachten te verminderen en de bevologenheid in deze doelgroep te vergroten. De bevindingen vereisen replicatie en toetsing in experimenten.

Hoe kunnen we werkstress bij docenten verminderen vanuit organisatieperspectief?

Om werkstress bij docenten te verminderen werd een participatieve, primair preventieve, organisatorische interventie onder de werknemers van twee middelbare beroepsopleidingsscholen (MBO's) in Nederland uitgevoerd (school A en school B). De interventie was gebaseerd op expertkennis: een adviesbureau ontwikkelde de interventie en paste deze meer dan honderd keer toe in zowel publieke als private organisaties. De interventie was echter nooit wetenschappelijk geëvalueerd. Zoals in detail beschreven in **hoofdstuk 3**, omvatte de interventie een behoeftepeilingsfase (bestaande uit interviews, een vragenlijst en groepsessies) en een fase van implementatie van de interventieactiviteiten die afkomstig waren uit de behoeftepeilingsfase (bijvoorbeeld gestructureerde prestatiebeoordelingen, of het inrichten van een docentenkamer). Een facilitator begeleidde de behoeftepeilingsfase en werd bijgestaan door een participatieve groep docenten en stafmedewerkers. In de fase van implementatie van interventieactiviteiten was de hulp van de facilitator optioneel. Slechts een van beide scholen kocht deze hulp in (school A).

De primaire hypothese was dat de interventie werkstress van medewerkers zou verminderen door de herstelbehoefte te verlagen en vitaliteit te vergroten (**hoofdstuk 4**). De hypothese werd getest in een gecontroleerde experiment met 356 medewerkers van twee scholen. De interventiegroep bestond uit 204

medewerkers van twee afdelingen en 24 teams. De controlegroep bestond uit 152 medewerkers van twee afdelingen en 24 teams. *Mixed model* analyse werd toegepast om de (gemiddelde) verschillen over de tijd (nulmeting, 12 maanden, 24 maanden) tussen de interventie- en controlegroep te meten. In alle analyses werd gecorrigeerd voor scores op de nulmeting en verschillende covariaten (bijvoorbeeld geslacht, onderwijsniveau). De hypothese kon niet bevestigd worden, aangezien er geen statistische verschillen werden gevonden tussen de interventie- en controlegroep op herstelbehoefte en vitaliteit. Twee kleine maar statistisch significante effecten in ongunstige richting werden gevonden op twee van de secundaire uitkomsten (nl. absorptie en organisatorische effectiviteit). Uit post-hoc per protocolanalyses op sensitiviteit bleek echter een effect in gunstige richting: medewerkers die deelnamen aan minstens twee van de drie stappen van de behoeftepeilingsfase (bijv. vragenlijst en groepsessie) hadden een significant hogere post-interventiescore op beroepsmatige eigen effectiviteit (het geloof in het eigen vermogen om het werk goed te kunnen uitvoeren) dan de medewerkers in de controlegroep. Dit resultaat geeft aan dat medewerkers met een hogere participatiegraad in de interventie mogelijk een mastery-ervaring hebben gehad, die vervolgens een positieve invloed heeft op de beroepsmatige eigen effectiviteit.

De procesevaluatie in **hoofdstuk 5** beschrijft dat het uitblijven van positieve effecten op de primaire uitkomstmaten gedeeltelijk kan worden verklaard door het falen van de implementatie. Eerdere procesevaluaties bestudeerden vooral het implementatieproces en de kwaliteit van de implementatie (nl. het naleven van het protocol). Als deze aanpak wordt toegepast op participatieve, organisatorische interventies voor het bevorderen van de gezondheid van werknemers, ontbreken echter belangrijke aspecten, zoals de rol van de context en de perceptie van de deelnemers. Daarom werd in dit interventieonderzoek een procesevaluatieraamwerk toegepast dat niet alleen betrekking heeft op de interventie en de uitvoering daarvan, maar ook op de context en de percepties van deelnemers. Uit de procesevaluatie bleek dat de behoeftepeilingsfase succesvol werd geïmplementeerd in school A, maar niet in school B, waar deelname en bereidheid tot verandering ontoereikend waren. De resultaten voor school A en B divergeren verder in de fase van implementatie van interventieactiviteiten: op school A zijn verschillende interventieactiviteiten geïmplementeerd, terwijl dit nauwelijks het geval was op school B. Ongeacht de feitelijke deelname en de daadwerkelijke implementatie van interventieactiviteiten, voelden medewerkers van beide scholen zich niet betrokken bij de keuze voor interventieactiviteiten. Daarnaast beschouwden de deelnemers van beide scholen de tweede fase van de interventie over het algemeen negatief, bijvoorbeeld omdat de interventieactiviteiten niet gezien werden als de juiste oplossingen voor de problemen (school A). Ook drie soorten contextfactoren belemmerden de implementatie. Ten eerste hadden de scholen geen (school B) of slechts gedeeltelijke capaciteit (school A) om de interventieactiviteiten uit te voeren, vooral omdat het middenmanagement niet volledig toegerust was om deze taak op zich te nemen. Ten tweede werd de organisatiecultuur in beide scholen gekenmerkt door een gebrek aan wederzijds vertrouwen tussen managers en docenten, wat de implementatie van deze participatieve

interventie belemmerde. Ten derde interfereerden diverse externe gebeurtenissen met de ingevoerde of geplande interventieactiviteiten (bijvoorbeeld de aankondiging van het Ministerie van Onderwijs dat het aantal klasgebonden lessen moest toenemen van 850 uur naar 1000 uur per jaar).

Op basis van hoofdstukken 4 en 5 wordt geconcludeerd dat verdere implementatie van de interventie in de huidige vorm niet aan te bevelen is. De algemene discussie (**hoofdstuk 8**) beschrijft hoe de interventie kan worden gewijzigd om beter aan te sluiten op de (organisatorische) context van scholen.

Hoe kunnen we het meest relevante bewijs verzamelen in interventieonderzoek in de werkcontext?

Het ontwerpen van evaluaties van gezondheidsinterventies op het werk kan uitdagend zijn met betrekking tot methodologie (bijv. moeilijkheden met randomisatie en een vergelijkbare controlegroep), de interventie zelf (bijv. de organisatie wil het interventieprotocol aanpassen) en de context (bijv. nieuw overheidsbeleid interfereert met de interventie). Het interventieonderzoek uit hoofdstukken 3-5 had ook te lijden onder uitdagingen in de onderzoeksopzet. Zo wilden beide scholen deelnemen aan het onderzoek met medewerkers van een specifieke afdeling, voor wie zij meenden dat de interventie het meest relevant zou zijn. Hoewel dit vanuit schoolperspectief goed voorstelbaar was, leverde het een methodologisch probleem op omdat randomisatie daardoor niet mogelijk was. Een gecontroleerde trial werd uitgevoerd, wat de effectevaluatie kwetsbaar maakte voor verstoringen en selectiebias. Op de nulmeting bleken er inderdaad significante verschillen tussen de interventie- en controlegroep te zijn. De narratieve review in **hoofdstuk 6** beschrijft een aantal van dergelijke uitdagingen die onderzoekers van gezondheidsinterventies op het werk ondervonden, en schetst onderzoeksopzetten die (sommige van) deze uitdagingen oplossen. In de review zijn daarnaast alternatieve onderzoeksopzetten onderzocht en beschreven voor situaties waarin een reguliere *Randomized Controlled Trial* (RCT) niet haalbaar is. Een alternatieve, experimentele onderzoeksopzet (nl. *stepped wedge design*) en diverse observationele onderzoeksopzetten werden gevonden en besproken (nl. *propensity scores*, *instrumental variables*, *multiple baseline design*, *interrupted time series*, *difference-in-difference*, en *regression discontinuity*).

Toekomstige primair preventieve, participatieve, organisatorische interventies zouden vooral kunnen profiteren van een gerandomiseerde *stepped wedge* of een *multiple baseline* onderzoeksopzet. Bij een *stepped wedge* onderzoeksopzet wordt de interventie willekeurig op alle groepen toegepast, maar na elkaar in plaats van parallel. Een *multiple baseline* onderzoeksopzet behelst hetzelfde, maar dan niet willekeurig. Natuurlijk kennen deze onderzoeksopzetten zo weer hun eigen uitdagingen, zoals dat er meer deelnemers en meetmomenten nodig zijn. Om het meest relevante bewijs in interventieonderzoek te verzamelen, zouden de onderzoekers de specifieke onderzoeksvraag, de complexiteit van de interventie, de beschikbaarheid van bestaande data, de context en de kosten moeten afwegen alvorens de meest geschikte onderzoeksopzet te kiezen.

Een ander probleem bij het verzamelen van het meest relevante bewijs in interventieonderzoek gaat over het ontwerpen van procesevaluaties. Het onderzoeksveld dat zich richt op procesevaluaties ontwikkelt zich snel en het valt op dat er veel verschillende theoretische kaders bestaan, waarin veel overlappende procesvariabelen beschreven worden. Om te begrijpen wat het meest relevante bewijs is als het gaat over het implementatieproces, is een systematische literatuurreview uitgevoerd, waarin 44 procesevaluaties van organisatorische stressmanagement interventies werden geïnccludeerd (**hoofdstuk 7**). Uit de review blijkt inderdaad een grote verscheidenheid in gehanteerde methoden en procesvariabelen: in de 44 evaluaties werden 47 verschillende procesvariabelen gevonden. In de helft van de artikelen ontbrak een theoretisch kader of programmatheorie, waarop de keuze van procesvariabelen kon worden gebaseerd. Bovendien werden de procesevaluatiegegevens meestal verzameld nadat de interventie plaatsvond, in plaats van ook voor en tijdens de interventie. De conclusie luidt dat een gestandaardiseerd kader met variabelen over de interventie, de mentale modellen van de deelnemers en contextfactoren de theorieontwikkeling op het gebied van procesevaluatie voor stressmanagement interventies zou kunnen verbeteren.

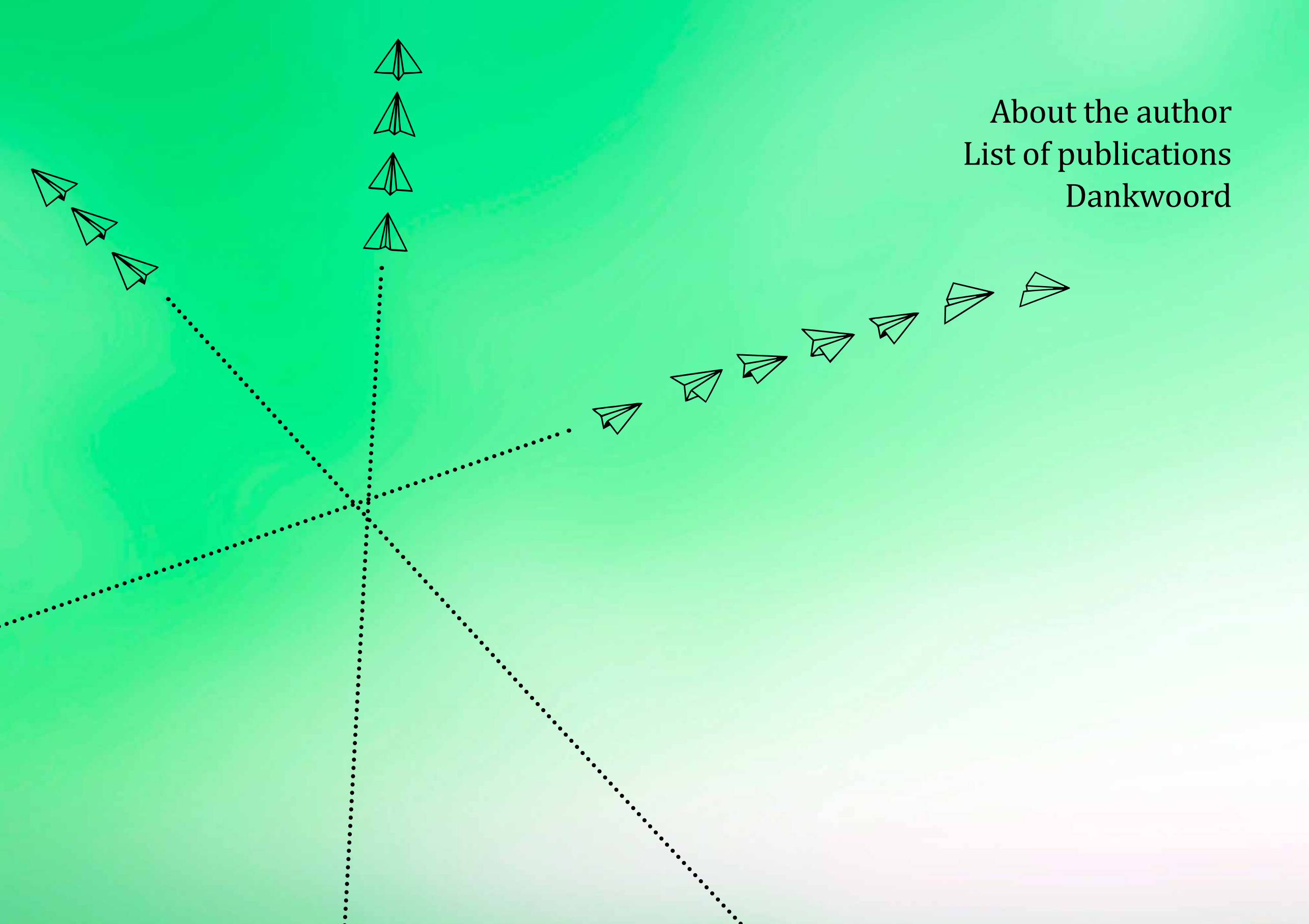
Algemene discussie

In de algemene discussie (**hoofdstuk 8**) worden de belangrijkste bevindingen van dit proefschrift besproken vanuit een individueel, organisatorisch en interventie-evaluatieperspectief. Het organisatorische perspectief wordt in het bijzonder besproken. De *expert-based* aanpak die is geëvalueerd in dit proefschrift wordt vergeleken met de *evidence-based* aanpak zoals beschreven in de literatuur. Verschillende verklaringen voor het uitblijven van positieve interventie-effecten worden gegeven. In de eerste plaats betoog ik dat de deelname van de werknemers aan de interventie werd belemmerd door een fenomeen dat bekend is als 'aangeleerde hulpeloosheid' (nl. afzien van een reactie op en dus ook het beïnvloeden van een gebeurtenis, terwijl dat wel binnen de eigen invloedssfeer had gelegen). Ten tweede beschrijf ik dat de interventie geen strategie bevatte om de percepties van en de beoordeling door de werknemers te beïnvloeden. Ten derde heeft de stijl van leidinggeven en de organisatiecultuur de (implementatie van) interventie niet voldoende gefaciliteerd. Op grond van deze discussie stel ik voor om de mate van participatie in interventies af te stemmen op de context, waarbij haalbaarheid en maatwerk leidende principes moeten zijn. Bovendien stel ik voor om voorafgaand aan de interventie een meer uitgebreide 'voorbereidingsfase' uit te voeren.

Vervolgens bespreek ik verschillende overwegingen ten aanzien van de methodologie. In de eerste plaats beschrijf ik methodologische moeilijkheden als gevolg van de definitie van werkstress, namelijk: (i) stress als gevolg van een disbalans tussen taakeisen en hulpbronnen; (ii) stress als een fase in een proces in plaats van een uitkomst; en (iii) het meetniveau van stress. Ik stel voor om in interventieonderzoek respectievelijk (i) de disbalans daadwerkelijk te berekenen om zo een sensitievere maat voor determinanten van werkstress te ontwikkelen, in plaats van indirecte maten voor stress of eindpunten in het

stressproces te gebruiken, (ii) een logisch model van de te verwachten veranderingen te ontwikkelen, waarin directe, intermediaire en distale uitkomsten worden beschreven; en (iii) het meetniveau van een (groeps)interventie en het meetniveau van de instrumenten gelijk te houden door (tenminste) constructen op groepsniveau te gebruiken (bijv. geaggregeerde ziektegegevens, beoordeling teamprestaties). In de tweede plaats beschrijf ik in de 'overwegingen'-paragraaf dat er een overlap bestaat tussen de definities van organisatorische interventies en primair preventieve interventies, wat niet verhelderend is. Ten slotte beschrijf ik dat in het huidige onderzoek een primaire preventieve interventie werd toegepast, waar secundaire preventie beter gepast zou zijn geweest.

Ik beëindig de algemene discussie met aanbevelingen voor praktijk en onderzoek. Mijn belangrijkste aanbeveling voor de praktijk is om een primaire preventieve, organisatorische interventie te beschouwen als een organisatorische verandering en die ook als zodanig te managen. Dit kan gedaan worden door een grondige voorbereidingsfase uit te voeren, om betrokkenheid voorafgaand aan de interventie te waarborgen en de deelnemerparticipatie vorm te geven op een manier die aansluit bij de organisatorische context. Mijn belangrijkste aanbeveling voor onderzoek is het combineren van proces- en effectmaatregelen in toekomstige interventiestudies, omdat dit vermoedelijk leidt tot een beter begrip van wat voor wie werkt onder welke omstandigheden.



About the author
List of publications
Dankwoord

About the author



Roosmarijn Schelvis was born on March 3rd, 1986 in Heemskerk (The Netherlands) as the second child of two primary school teachers. Growing up amidst three siblings, she completed her secondary education (Gymnasium) at Bonhoeffer College in Castricum in 2004. Because she had always been curious about what drives human behavior, she enrolled in the Bachelor of Science in Psychology at Leiden University. Roosmarijn moved to Leiden and held several voluntary positions such as the president of the psychological students' association Labyrint. Various extracurricular courses in psychology, anthropology and educational studies were part of her academic training, including a semester at the Università degli Studi di Firenze in Florence (Italy). After having successfully completed an Honours Research Bachelor Project on non-verbal behavior in socially anxious young children, she obtained her Bachelor of Science degree in Psychology in 2008.

During her subsequent Master of Science in Health Psychology at Leiden University, Roosmarijn wrote a master thesis consisting of a meta-analysis and literature review on the effectiveness of school-based life skills programs for adolescents' health behaviors. In parallel with her thesis she worked as a trainee-consultant during a six month internship at the Management Development department of management consulting firm Berenschot. After this internship she aspired to learn more about organizations and how they work, and thus added a minor in Business Administration at Leiden University to her curriculum. In 2010 she graduated with honours (cum laude) as a Master of Science in Health Psychology and she completed the minor in Business Administration.

In 2010 she started her career as a junior research scientist at the Netherlands Organization for Applied Scientific Research (TNO). She worked on several applied projects on work and psychosocial health, while living in Amsterdam with Willem ten Teije. In 2012 Roosmarijn and Willem got married. From 2013 onwards Roosmarijn was also affiliated to VU University Medical Center as a PhD-student, which resulted in the present thesis. Currently, Roosmarijn still works as research scientist at the department of Work, Health and Technology at TNO on various research projects related to work stress, health behaviors and the prevention of work-related disease. She recently moved back to Leiden with Willem and their daughter Isolde (born in 2016). The family is expecting a new member in January 2018.

Contact: roos.schelvis@tno.nl / roos.schelvis@gmail.com.

List of publications

Scientific articles in this thesis

Schelvis RMC, Wiezer NM, van der Beek AJ, Twisk JW, Bohlmeijer ET, Oude Hengel KM: **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**(1):141.

Schelvis RMC, Wiezer NM, Blatter BM, van Genabeek JAGM, Oude Hengel KM, Bohlmeijer ET, van der Beek AJ: **Evaluating the implementation process of a participatory organizational level occupational health intervention in schools.** *BMC Public Health* 2016, **16**(1):1212.

Havermans BM, Schelvis RMC, Boot CR, Brouwers EP, Anema JR, van der Beek AJ: **Process variables in organizational stress management intervention evaluation research: a systematic review.** *Scand J Work Environ Health* 2016, **42**(5):371-381.

Schelvis RMC, Oude Hengel KM, Burdorf A, Blatter BM, Strijk JE, van der Beek AJ: **Evaluation of occupational health interventions using a randomized controlled trial: challenges and alternative research designs.** *Scand J Work Environ Health* 2015, **41**(5):491-503.

Schelvis RMC, Oude Hengel KM, Wiezer NM, Blatter BM, Van Genabeek JAGM, Bohlmeijer ET, van der Beek AJ: **Design of the Bottom-up Innovation project - A participatory, Primary preventive, Organizational level intervention on work-related stress and well-being for workers in dutch vocational education.** *BMC Public Health* 2013, **13**(1).

Other scientific articles

Schelvis RMC, Zwetsloot GI, Bos EH, Wiezer NM: **Exploring teacher and school resilience as a new perspective to solve persistent problems in the educational sector.** *Teachers and Teaching* 2014, **20**(5):622-637.

Blok MM, Groenesteijn L, Schelvis RMC, Vink P: **New ways of working: does flexibility in time and location of work change work behavior and affect business outcomes?** *Work* 2012, **41**(Supplement 1):2605-2610.



International presentations

Schelvis RMC. **Implementation of a comprehensive intervention for workplace mental health: lessons learned**. Invited as a junior keynote at Wellbeing at Work conference 2016, Amsterdam, the Netherlands.

Van der Doef, M. and Schelvis RMC. **Predictors of work ability in employees with severe headaches**. Oral presentation at Wellbeing at Work conference 2016, Amsterdam, the Netherlands.

Schelvis RMC, Wiezer NM, Oude Hengel KM, Bohlmeijer ET, Van der Beek AJ. **To what extent does implementation influence intervention effects in an organizational-level intervention for work-related stress?** Oral presentation at APA Work, Stress and Health congress 2015 in Atlanta, United States of America.

Schelvis RMC, Leijten FRM, Wiezer NM, Oude Hengel KM, Bohlmeijer ET, Van der Beek AJ, Ybema, JF. **Interplay between self-efficacy and job demands and resources in influencing work engagement and depression in older teachers: a three-wave longitudinal study**. Poster presentation at APA Work, Stress and Health 2015 in Atlanta, United States of America.

Schelvis RMC, Wiezer NM, Blatter BM, Oude Hengel KM, Bohlmeijer ET, Van der Beek AJ. **Bottom-up innovation in vocational education; Lessons from implementing a participatory intervention in the Netherlands**. 2014, invited lecture at La Trobe University in Melbourne, Australia.

Schelvis RMC, Wiezer NM, Blatter BM, Oude Hengel KM, Bohlmeijer ET, Van der Beek AJ. **Bottom-up innovation in vocational education; Lessons from implementing a participatory intervention in the Netherlands**. Oral presentation at ICOH-WOPS Congress 2014 in Adelaide, Australia.

Selection of related national publications

De Koning, N. & Schelvis, R. **Werkdruk ontrafeld – u kunt pas wat aan werkdruk doen als u weet wat u eronder verstaat**. *Arborendement* 2014, december, p. 18-19.

Wiezer, N., Schelvis, R., Van Zwieten, M., Kraan, K., Van der Klauw, M., Houtman, I., Kwantes, J.H., Bakhuis Roozeboom, M. **Werkdruk**. *TNO-rapportage: R12-10877*. 2012.

Schelvis, R., Bakhuis Roozeboom, M., De Kraker, H., Wiezer, N., & Joling, C. **Meer bevlogenheid, minder stress: beter bedrijfsresultaat**. *Management Executive*, 2012 (juli-augustus), p. 34-37.

Schelvis, R. & Oude Hengel, K. (2011). **Gezondheidsmanagement in de praktijk volgens IM; Intervention Mapping bij gezondheidsinterventies in organisaties**. *Gids voor Personeelsmanagement* 2011, 11, p. 26-29.

Schelvis, R., Kraan, K. & Wiezer, N. (2011). **Schoolleider onmisbare coach op weg naar bevlogenheid; Werkstressoren en Energiebronnen**. *Kader Primair; vakblad voor leidinggevend in het primair onderwijs*. Algemene Vereniging Schoolleiders (AVS). 2011, 16(5), 14-17.



Dankwoord

Als ik van smartlappen hield zou ik nu Hazes opzetten. Want dit proefschrift kon er alleen maar komen omdat een heleboel mensen in mij geloofden. Het is mijn bedoeling al deze mensen hierna stuk voor stuk te bedanken.

Promotieteam

Allard van der Beek, mijn promotor. Tijdens onze kennismaking gebaarde je naar een volle boekenplank met de proefschriften van promovendi die je eerder begeleidde. Toen wist ik dat ook dit boekje op een dag af zou komen. Ik heb groot plezier beleefd aan de vinkjes die je maakte in de kantlijn van mijn stukken - één voor elke verbetering - zodat ik die na aanpassing weer kon wegstrepen. Voortgang was namelijk nog nooit zo concreet! Ik ben blij dat ik van je heb mogen leren. En welke andere promotor stuurt je nou je favoriete nummer van Cake als het even tegenzit?!

Noortje Wiezer, copromotor van het eerste uur. Ik vind het geweldig dat jij mij, als naaste collega, wilde begeleiden in de totstandkoming van dit proefschrift. Je associatieve vermogen en creativiteit inspireren mij nog steeds. Van jou leerde ik om me te blijven verwonderen tijdens langdurige observaties, zoals ik deed op de scholen. Ik ben je dankbaar voor de aanmoediging om te vertrouwen op mijn observatievermogen en dus mijn eigen verhaal te vertellen. Zonder jouw eindeloos puzzelen met de budgetten zouden al mijn vakantie-dagen een laptopscherm als uitzicht hebben gehad. Dank dat je me daarvoor hebt behoed.

Karen Oude Hengel, copromotor van het tweede uur. Jouw pragmatisme en besliskracht hebben menig proces versneld. Ik hoop nog steeds dat dat aanstekelijk werkt. Ik ben blij dat ik ook voor antwoorden op alle voor de hand liggende vragen bij jou terecht kon. Je wist me altijd goed gemutst weer op weg te helpen. Ook ben ik je dankbaar voor alle steuntjes in de rug, je hebt een uitstekend gevoel voor timing, zelfs als het over prosecco gaat.

Andere begeleiders

Birgitte Blatter, als teamleider 'Preventie' nam jij mijn wens om promotieonderzoek te doen direct serieus. In je latere managementfunctie bij TNO dacht je aan mij voor het uitvoeren van het interventieproject 'Innovatie van Onderop', waar uiteindelijk drie artikelen uit voortvloeiden. Deze stevige basis gaf mij en vele anderen genoeg vertrouwen om met dit proefschrift door te gaan, ook al kon jij zelf niet meer betrokken zijn als copromotor. Veel dank daarvoor!

Ernst Bohlmeijer, lange tijd was jij mijn medepsycholoog-medestander in het promotieteam; die invloed is in de interventie-artikelen zeker terug te zien. De groene campus in Twente gaf me bovendien altijd een beetje een vakantiegevoel onder werktijd! Ik vind het erg leuk dat ik jou en je vakgroep een beetje

heb leren kennen, ook al kon je niet tot het einde toe betrokken blijven als tweede promotor. Wellicht werken we in de toekomst nog eens samen om de mentale gezondheid van Nederlanders te bevorderen.

Joost van Genabeek, als projectleider van het interventieproject was jij voor mij heel belangrijk. Welke hobbel we ook tegenkwamen, jij was er nooit erg van onder de indruk en bleef kalm. Ik volgde je graag. In het bijzonder bedankt voor je aanmoediging om de procesevaluatie gedetailleerd en 'anders dan anders' aan te pakken, want veel van wat ik in de algemene discussie uitwerkte is geïnspireerd op wat ik leerde tijdens het procesevaluatie onderzoek.

Co-auteurs

Jan Fekke Ybema, jij was nooit te beroerd om mij de analyses van het STREAM-artikel (hoofdstuk 2) nog een keer uit te leggen, ook al werkte ik met grote tussenpozen aan het stuk, waardoor het geleerde niet erg goed bekliffde. Veel dank dus voor je geduld en de nauwgezetheid waarmee je de analyses uitvoerde, de methode en resultaten schreef, en het manuscript controleerde. Voor haar rol bij hetzelfde artikel bedank ik **Fenna Leijten** hartelijk. Ik ben onder de indruk van hoe snel jij denkt en werkt. De automatische check op het Engels die jij als *native speaker* uitvoert maakte dit ingewikkelde artikel eenvoudiger. Zonder jou was het schrijven van dit stuk lang niet zo leuk geweest!

Jos Twisk, dank voor je geduldige toelichting op al mijn vragen over de effectevaluatie (hoofdstuk 4)! Ik heb er veel van geleerd.

Lex Burdorf, dank voor je richtinggevende hulp bij het artikel over alternatieven voor de gerandomiseerde gecontroleerde trial (hoofdstuk 6). Die kwam op een cruciaal moment, want jij zag een structuur die ik al een tijdje kwijt was in dit stuk. Ik heb veel geleerd van je systematische manier van redeneren en over het opvoeren van welke bron voor welk argument. Bovendien herinner ik me dat we ook een hoop hebben gelachen!

Bo Havermans, mijn VU-kamergenoot. Op onze systematische review ben ik trots (hoofdstuk 7). In wetenschappelijk opzicht zijn we aan elkaar gewaagd, dus wie weet komt er nog eens een Schelvis & Havermans naast deze Havermans, Schelvis en anderen. Ik heb met mijn 'alle details erbij'-schijfstijl geleerd van jouw 'meer is minder'-schijfstijl. Met plezier denk ik terug aan onze congresgangen naar Australië en Amerika, je altijd opbouwende non-verbale communicatie heeft menige saaie sessie opgevrolijkt. Overigens zou dit artikel lang niet zo puntig zijn geweest zonder **Cécile Boot**. Jouw doelgerichte manier van werken beschouw ik als een heel goed voorbeeld. Bovendien zou ik zo weer een internationaal wetenschappelijk congres of retraite met je willen organiseren!



Lees- en promotiecommissie

Prof. Paulien Bongers, voorzitter van de lees- en promotiecommissie. Ik maak je nu al zeven jaar mee als directeur bij TNO, in functies en bedrijfsonderdelen die steeds van naam veranderen. Soms lijkt de enige constante jouw sturing, die altijd stevig is op de inhoud. Ik ben nog steeds onder de indruk van hoe je dat doet, voor mij het bewijs dat leidinggeven vanuit die inhoud heus wel kan werken. Daarom ben ik trots dat jij de lees- en promotiecommissie wilde voorzitten.

Mijn grote dank gaat eveneens uit naar de andere leden van de lees- en promotiecommissie: **Prof. Ijmert Kant, Prof. Tinka van Vuuren, Prof. Toon Taris, Dr. Aukje Nauta en Dr. Maaïke Huysmans**. Op het randje van de komkommertijd keurde u het manuscript goed, waardoor de verdediging van mijn proefschrift nog voor mijn zwangerschapsverlof gepland kon worden. De baby kan zich nu al verheugen in het vooruitzicht op de onverdeelde aandacht van zijn moeder.

Betrokkenen bij de interventie

Mijn grote dank gaat uit naar de twee deelnemende MBO-instellingen (**Horizon College en Friesland College**) en de betrokkenen bij het adviesbureau dat de interventie bedacht en uitvoerde (Heijting Weerts Groep, **Edda en Jan Anton**). In het bijzonder bedank ik de deelnemers aan de Ontwikkelmotor, de participatieve groep, van beide scholen: **Hans, Hendrik, Kees en Marian** enerzijds, en **Elly, Fokelien, Henk, Hennie, Henny, Jan, Joke en Ronald** anderzijds. Ook de collegeleden, directeuren en opleidingsmanagers wil ik in deze dankzegging betrekken. Er lag een behoorlijke klus op jullie bord en het onderwijsveld is werkelijk ingewikkeld. Tot slot dank aan alle medewerkers die tot drie keer toe die eindeloze vragenlijst hebben ingevuld, soms zelfs midden in de zomervakantie teruggestuurd, beplakt met een heel velletje Franse postzegels. Zonder jullie inspanningen waren hoofdstuk 3 t/m 5 er niet geweest!

Collega's

Sandra Eikhout, research manager Work, Health en Technology en mijn leidinggevende bij TNO gedurende het grootste deel van mijn proefschriftwerkzaamheden. Bedankt dat je me de kans gunde om me wetenschappelijk te ontwikkelen in een promotietraject, ook al was je misschien niet zo fan van de buitenpromovendus-constructie. Zonder jouw steun was het niet gelukt. Met plezier blijf ik de ontwikkelde kennis inzetten in lopende en toekomstige projecten.

Lieve **collega's van Work, Health & Technology** en haar voorlopers (team Preventie, expertisegroep Work & Health), we zijn met ongeveer 45 en ik kan jullie hier niet één voor één bedanken. Jullie initiële interesse in mijn promotietraject maakte dat het echt bestond, dat ik echt 'van start' was. Jullie tus-

sentijdse interesse maakte dat ik zag dat er heus wel enige voortgang was, al voelde dat soms niet zo. Jullie interesse tegen het eind maakte dat ik me plots realiseerde dat het boek bijna af was! Bijzondere dank aan een aantal naaste collega's die nog niet genoemd zijn en met wie ik al zeven jaar vier tot vijf dagen per week werk aan beter, gezonder en veiliger werk: **Maartje, Marianne, Marloes, Irene Houtman, Maaïke, Seth, Goedele, Jan Harmen, Swenneke, Ernest, Karolus, Wendela, Romy, Jorien, Luuk en Dick. Irene Niks** dank ik in het bijzonder voor het aanscherpen van mijn Nederlandse samenvatting, leuk dat ik je na onze ontmoeting op ICOH-WOPS in Australië sinds kort ook 'collega' kan noemen. En natuurlijk alle collega's uit het mooie **Team 2**: ik ben trots dat ik jullie teamcoördinator mag zijn! Ook de trouwe krachten op het secretariaat, **Caroline, Carolina en Henny**, bedankt voor jullie interesse en voor de manier waarop jullie mijn werkende leven vereenvoudigen.

Fijne **collegapromovendi op de VU**. Veel dank voor het samen trotseren van muizen, asbest en moeilijke reviewers!

Vrienden en familie

Lieve **Annelies en Rosalie**, twee kanjers van vriendinnen als twee kanjers van paranimfen. Fijn om te merken dat 'het cluster' waarin we elkaar blijven ontmoeten zich nu al uitstrekt tot aan Berlijn. Met trots voer ik jullie straks mee het podium op. Ik weet dat het mijn zenuwen zal indammen. Dank voor jullie aanmoedigingen en de steevaste overtuiging dat deze promotieplechtigheid ooit zou komen.

Lieve **Vera**, we zijn talloze kopjes thee, glazen wijn en tegenwoordig ook boterhammen appelstroop verder dan toen we bevriend raakten tijdens onze opleiding Psychologie. Ik kijk uit naar meer wijn en appelstroop als ik straks weer in Leiden woon! Dat geldt ook voor jou, **Dagmar**! Dat er nog maar vele Buitenkunsten mogen volgen, waarop we het totaal niet over proefschriften hebben - omdat dat iets uit het verleden is.

Lieve **Yvonne, Claudia, Elise en Jantien**, lieve vriendinnen. Dank voor ontelbare fijne momenten sinds 1998! Fijn dat het altijd wel even over mijn proefschrift kon gaan, maar juist ook niet de hele tijd. Lieve **Nikki**, jouw nagedachtenis houdt me bij de les op moeilijke momenten. Juist op die momenten probeer ik met de jou kenmerkende lichtheid te leven of beslissen.

Lieve **Carel en Erna; Ilonka en Jan; Heleen, Vincent, Bruno en Felix; Irene, Thijs Jan en Nora; Lidewij, Stephan en Bella**. Lieve schoonfamilie, dank voor de interesse in m'n proefschrift. Maar vooral bedankt voor de afleiding ervan op gezellige familiebijeenkomsten. Ik geef het buitenpromovendusstokje graag over aan Thijs Jan.

Lieve **opa**, dank voor uw niet aflatende interesse in mijn proefschrift. Met ome Gerard, Mieke en ik hebben we in elke generatie een gepromoveerde 'Reijners', ook al verschillen de achternamen misschien enigszins. Nu mijn boekje af is



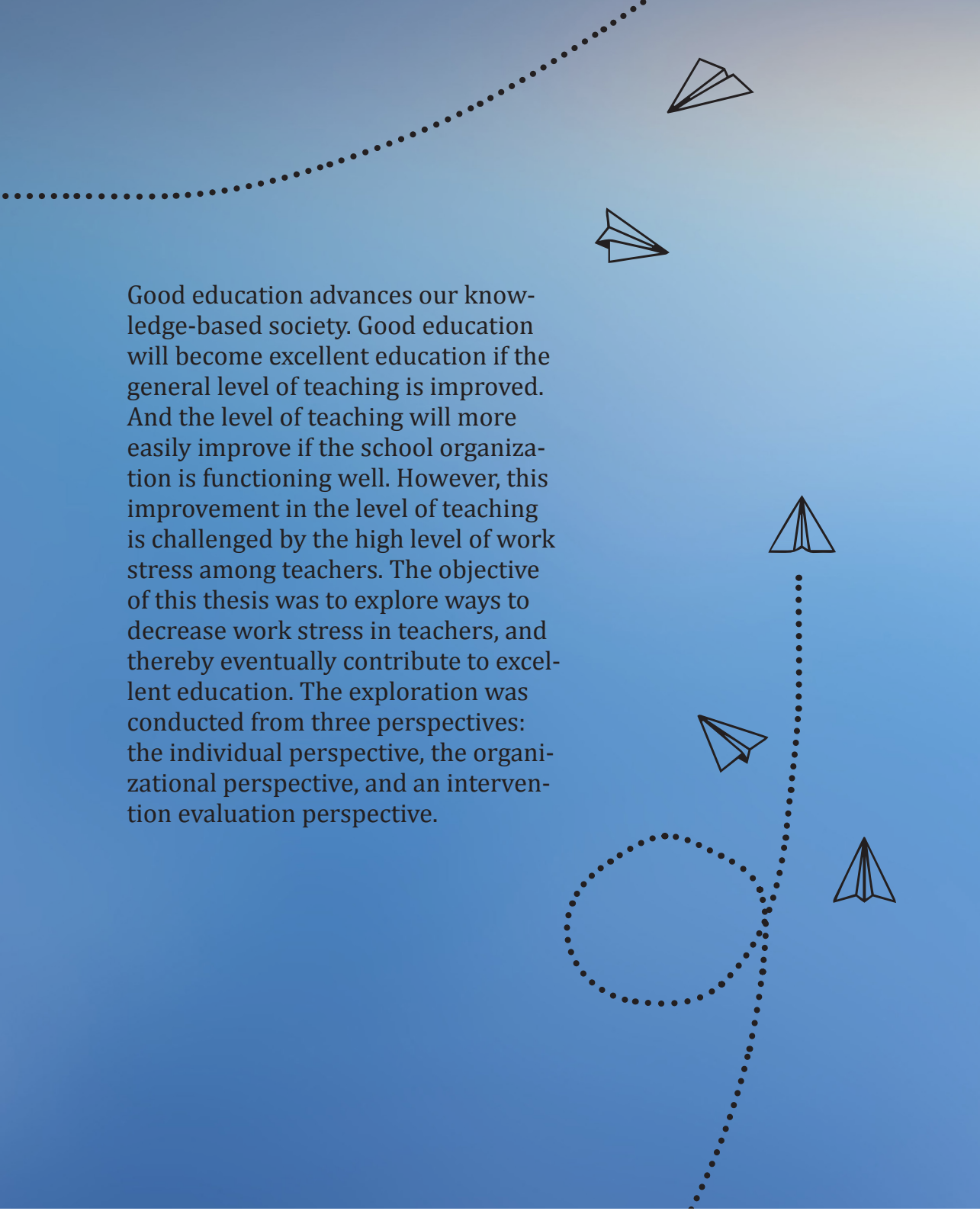
kunnen we wat mij betreft aan dat van u beginnen.

Lieve **Femke, Sjeng** en **Liselot** – Schelvis siblings. Dank dat ik gewoon jullie zus(je) kan zijn zonder dat het uitmaakt of ik gepromoveerd ben!

Lieve pap en mam, **Cees en Inge**. Ik herinner me jullie vroegere grapje nog goed: “je mag alles worden wat je wilt, maar ga niet het onderwijs in”. En dat terwijl jullie beiden al bijna jullie hele leven in en om datzelfde onderwijs werkzaam zijn. En meestal met plezier. Ik ben ervan overtuigd dat jullie geintje-met-een-seintje alles te maken heeft met het thema van dit proefschrift. Op megalomane dagen had het schrijven van dit proefschrift daarom wel wat weg van een persoonlijke missie om werkstress het onderwijs uit te krijgen. Met plezier heb ik jullie advies om “niet het onderwijs in te gaan” dus genegeerd. En onder grote dank voor jullie onbedoelde aansporing draag ik dit proefschrift aan jullie op.

Liefste **Willem**. Speciaal voor jou zou ik Hazes willen instuderen op de piano, terwijl je m’n houding tegenover het smartlappengenre kent en je weet dat ik geen piano speel. Ook al “geloof” je niet per se in wetenschap, je geloof in mij was ontzettend belangrijk om het promoveren naast mijn baan bij TNO vol te houden. Je goed getimed vraag op één van mijn jammerklachten (“Waarom ga je er eigenlijk mee door?”) had dezelfde uitwerking als voornoemd ouderlijk advies. Ik beloof hierbij plechtig dat ik nooit meer ga promoveren.

Lieve, kleine **Isolde**. Jij bent mijn persoonlijke, bewezen effectieve werkstressremedie. Ik geef je een exemplaar van dit proefschrift, maar neem er vooral geen voorbeeld aan. Aan de andere kant, er is een kans dat je op mij lijkt en dit advies volstrekt in de wind slaat. Dan blijft alleen het advies van Annie M.G. Schmidt over, in die dichtregel over een regenworm uit Sneek: “[...] *doe nooit wat je moeder zegt, dan komt het allemaal terecht.*”

A decorative graphic on the right side of the page features a dotted line that starts at the top left, curves upwards, then downwards, then upwards again, and finally curves to the right. Several paper airplane icons are scattered around the path, appearing to fly along it. The background is a solid blue color.

Good education advances our knowledge-based society. Good education will become excellent education if the general level of teaching is improved. And the level of teaching will more easily improve if the school organization is functioning well. However, this improvement in the level of teaching is challenged by the high level of work stress among teachers. The objective of this thesis was to explore ways to decrease work stress in teachers, and thereby eventually contribute to excellent education. The exploration was conducted from three perspectives: the individual perspective, the organizational perspective, and an intervention evaluation perspective.



TNO

eMGO+
Institute for Health and Care Research

BODY@WORK