Freie Universität Berlin
Fachbereich Erziehungswissenschaft und Psychologie

Work Engagement – An Active State of Occupational Well-Being

Dissertation
zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
M.Sc. Tino Lesener

Berlin, 2020
Erstgutachter: Prof. Dr. Dieter Kleiber

Zweitgutachterin: Prof. Dr. Bettina Hannover

Drittgutachterin: Prof. Dr. Renate Soellner

Betreuer: PD Dr. Dr. Burkhard Gusy

Tag der Disputation: 12/06/2020
Table of Contents

Abstract ......................................................................................................................................... II
Abstract (deutsch).......................................................................................................................... V
Theoretical Background ................................................................................................................ 8
  Introduction ........................................................................................................................... 8
  Defining Work Engagement............................................................................................... 9
  Work Engagement and its Integration in the JD-R Framework................................. 11
  Empirical Evidence on the JD-R Framework and its Motivational Process............ 13
  Research Questions............................................................................................................... 14
  References ............................................................................................................................. 15

The Job Demands-Resources Model: A Meta-Analytic Review of Longitudinal Studies ....... 20
  Abstract.................................................................................................................................. 21
  Introduction .......................................................................................................................... 22
  Method .................................................................................................................................. 27
  Results ................................................................................................................................... 31
  Discussion ............................................................................................................................. 42
  References ............................................................................................................................. 48

The Drivers of Work Engagement: A Meta-Analytic Review of Longitudinal Evidence.......64

Relationships between Effort-Reward Imbalance and Work Engagement in Police Officers: Taking a Salutogenic Perspective.................................................................................................94

Measuring Well-Being with the Utrecht Work Engagement Scale – Student Form: Validation of a 9- and a 3-Item Measure of Student Engagement ................................................................. 121
  Abstract................................................................................................................................ 122
  Introduction ........................................................................................................................ 123
  Method ................................................................................................................................ 125
  Results ................................................................................................................................... 129
  Discussion ........................................................................................................................... 132
  References ........................................................................................................................... 134

General Discussion .................................................................................................................... 138
  Limitations .......................................................................................................................... 139
  Recommendations for Future Research ......................................................................... 141
  Implications ....................................................................................................................... 144
  Conclusions ....................................................................................................................... 147
  References .................................................................................................................................. 147

Statement of Independence (Erklärung zur Dissertation)......................................................... 151
Abstract

Work engagement as a positive, fulfilling, and affective-motivational state of occupational well-being is currently one of the most popular outcomes in occupational health psychology. In the last decade, there has been a considerable amount of research on the concept. For instance, a quick search in Google Scholar (December 2019) reveals almost 80.000 hits for «work engagement». Researchers’ and practitioners’ interest in work engagement is understandable given its association with better health and performance. Despite the growing number of articles on the topic, there are still blind spots that need to be addressed. Hence, I was interested in bringing some more clarity to the concept. Specifically, I aimed to get detailed insights into the job-related antecedents and outcomes of work engagement, and its applicability in specific occupational and non-occupational contexts.

To tackle these aims, I realized two meta-analytic reviews, where I used the innovative meta-analytic structural equation modeling (metaSEM) approach. This approach is still very uncommon in psychological research, especially on occupational health psychology topics. However, metaSEM makes it possible to integrate data from all available studies into one meta-analytic structural model, and to test this model even if it had not been tested in the primary studies. Furthermore, I used data from a health monitoring program at a large urban police department, and from a health monitoring program at a large German university. Since I conducted the studies in collaboration with my colleagues, I will use the terms we and our when presenting results and implications.

We were able to validate the essential assumptions of the job demands-resources (JD-R) framework, which was the first theoretical model that introduced work engagement conceptually: job resources lead to work engagement even in the long run, whereas job demands do not have a substantial impact on work engagement. The JD-R framework has the crucial advantage over other occupational frameworks that it allows for both salutogenic and pathogenic effects of job characteristics on occupational well-being. Based on our research, we concluded, that the JD-R framework is an excellent theoretical basis to examine work engagement in a broad range of organizations and occupational fields.
We were also interested to find out which job resources are most important for work engagement and why some job resources might be more significant than others for facilitating work engagement. Hence, we investigated which job resources would be the key drivers of work engagement in various occupational settings. We differentiated job resources in organizational-level, group-level, and leader-level resources and concluded that organizational-level resources (e.g., autonomy) are the most important predictors for work engagement. This provides clarity on where interventions to enhance occupational well-being should start. Although interventions at any of the three levels are promising, organizations are well-advised to predominantly strengthen resources at the organizational level to increase employees’ work engagement.

With this in mind, we aimed to broaden the theoretical and contextual application of work engagement. Hence, we integrated work engagement as a positive, salutogenic outcome into the effort-reward imbalance (ERI) model. Since the ERI model was originally designed to identify health complaints and their job-related antecedents, the integration of work engagement broadens the applicability of this model. We found that an imbalance between effort and reward is associated with reduced work engagement. However, especially the reward dimensions esteem and security foster work engagement within the police context. Hence, we gained insights into which specific factors within the ERI model are most important for work engagement.

Furthermore – since work engagement is associated with various positive health- and performance-related outcomes – we applied the concept and its operationalization to the context of university students. We examined the psychometric structure of a short (9 items) and an ultra-short (3 items) measure of student engagement with the Utrecht Work Engagement Scale – Student Form (UWES-SF), and its relations to other pathogenic and salutogenic health indicators. We confirmed the three-factor structure of the German UWES-9-SF and the one-factor structure of the UWES-3-SF. Both measures appear to be reliable and valid indicators of student engagement that can be used as an alternative to Schaufeli’s original version, the UWES-SF. Due to its brevity, the UWES-3-SF is most appropriate to capture student engagement in national and epidemiological surveys.
Taken together, my research sharpens the understanding of work engagement as a theoretical concept and its applicability in various occupational and non-occupational contexts. I firmly believe that work engagement – defined as a positive, fulfilling state – is one of the most beneficial states of well-being for both employees and students.
Abstract (deutsch)


Darüber hinaus wollten wir klären, welche spezifischen arbeitsbezogenen Ressourcen die wesentlichen Determinanten für Work Engagement sind, und warum manche arbeitsbezogenen Ressourcen wichtiger für Work Engagement sind als andere. Wir haben arbeitsbezogene Ressourcen differenziert in (1) aufgabenbezogene, (2) soziale sowie (3) führungsbezogene Ressourcen und konnten zeigen, dass insbesondere aufgabenbezogene Ressourcen (z. B. Handlungsspielraum bei der Arbeit) die wichtigsten arbeitsbezogenen Determinanten für Work Engagement sind. Daraus schlussfolgern wir, dass Interventionen am Arbeitsplatz vor allem an aufgabenbezogenen Ressourcen ansetzen sollten, um das Work Engagement der Beschäftigten zu fördern.


Aufgrund der positiven Zusammenhänge zu verschiedenen Gesundheits- und Leistungsindikatoren haben wir darüber hinaus das Konzept auf den Kontext des Studiums übertragen. Wir haben eine Kurzversion (9 Items) sowie eine Ultrakurzversion (3 Items) zur Messung von Student Engagement mit der Utrecht Work Engagement Scale – Student Form (UWES-SF) eingeführt, deren psychometrische Qualitäten überprüft sowie deren Assoziationen zu diversen pathogenen und salutogenen Gesundheitsindikatoren berichtet. Wir konnten die dreifaktorielle Struktur der Kurzversion (UWES-9-SF) sowie die einfaktorielle Struktur der Ultrakurzversion (UWES-3-SF) bestätigen. Beide Instrumente sind reliable und valide Instrumente zur Messung von Student Engagement und können als Alternative zur Originalversion eingesetzt werden. Durch die Kürze ist die UWES-3-SF besonders geeignet für nationale, epidemiologische Gesundheitsmonitorings.
Theoretical Background

Introduction

Occupational Health Psychology (OHP) emerged as a novel psychological field just 20 years ago (Schaufeli, 2004). The main purpose of OHP is to develop, maintain, and promote employees health (Tetrick & Quick, 2011). Others have remarked that OHP concerns the application of psychology to improve the quality of work life, and to protect and promote the safety, health, and well-being of employees (Schaufeli, 2004). In OHP, »health« is considered as a positive concept in line with the definition of the World Health Organization (1946, June 19-22): “Health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity.” However, most research has mainly focused on negative aspects of work such as damage, disease, disorder, and dysfunction (Houdmont & Leka, 2010). In a review ten years ago, Macik-Frey, Quick, and Nelson (2007) found only two studies that were published between 1996 and 2006 (in the Journal of Occupational Health Psychology) which focused on positive aspects of work and its relations to positive health and well-being. Tetrick and Quick (p. 17) concluded in 2011, that “more theoretical development and supporting research are needed to define health not just as the absence of illness but as something more.”

Even so, the last decade has seen a growing number of researchers exploring positive indicators of work-related health and well-being, such as job satisfaction, work engagement, enthusiasm, pleasure, or flow (Bakker & Oerlemans, 2011; Gusy, 2017). Conceptually, Bakker and Oerlemans (2011) have introduced the circumplex model of occupational well-being, which proposes that affective states can be differentiated on two dimensions: (1) a pleasure-displeasure continuum, and (2) an activation-deactivation continuum. In the model, there are three major positive forms of work-related well-being: job satisfaction, happiness at work, and work engagement. All of them are characterized by high levels of pleasure, but they differ in terms of activation. Job satisfaction reflects a high level of pleasure, but a low level of activation (Bakker & Oerlemans, 2011). Employees who are satisfied may feel good at work, but may also have limited energy or aspirations (Grebner, Semmer, & Elfering, 2005). Happiness at work reflects a high level of pleasure, and a moderate level of activation. Employees who are happy may have a somewhat higher level of activation than those who are merely satisfied or content, but they are not enthusiastic or excited about their job (Bakker & Oerlemans, 2011). Work
engagement, however, reflects a high level of pleasure – and a high level of activation. Employees experience their work as stimulating, energetic, meaningful, and engrossing (Bakker & Oerlemans, 2011). Since modern organizations expect their employees to be proactive, responsible, committed, creative, and energetic, the concept of work engagement can truly make a difference for organizations as a competitive advantage (Leiter & Bakker, 2010). Work engagement, not just job satisfaction, may be needed to increase employees’ job performance (Bakker & Oerlemans, 2011).

Hence, I was interested in bringing some clarity to the concept of work engagement. Despite the growing number of articles on the topic, there are still blind spots that need to be addressed. Specifically, I aimed to get detailed insights into the job-related antecedents and outcomes of work engagement, and its applicability in specific occupational and non-occupational contexts.

Defining Work Engagement

Work engagement as a positive, fulfilling, and affective-motivational state of occupational well-being is currently one of the most popular outcomes in OHP. Contrary to job satisfaction, work engagement is not only a positive, pleasurable state, but also a state with high levels of activation (Bakker & Oerlemans, 2011). In the last decade, there has been a considerable amount of research on the concept of work engagement. Almost ten years ago, Schaufeli and Bakker (2010) reported 785 hits for »work engagement« in Google Scholar. Nowadays (December 2019), the same search reveals almost 80,000 hits. Researchers’ and practitioners’ interest in work engagement is understandable given its association with better health and performance (Halbesleben, 2010).

Due to its popularity, there have been many different constructs labeled as »engagement« in both business and academia. However, these definitions are not always consistent (Christian, Garza, & Slaughter, 2011). Macey and Schneider (2008) showed that the term is used differently to refer to either psychological states, traits, or behaviors.

Engagement as a state. Some authors have defined engagement in terms of various psychological states such as satisfaction, commitment, or involvement (for an overview see Macey & Schneider, 2008). For example, Harter, Schmidt, and Hayes (2002, p. 269) define engagement as “the individual’s involvement and satisfaction with as well as enthusiasm for
work.” Similar to satisfaction, engagement is considered as a positive, work-related state (Macey & Schneider, 2008), but – opposed to satisfaction – engagement connotes high levels of activation and energy (Bakker & Oerlemans, 2011). Similar to commitment and involvement, engagement reflects attachment and identification, and the willingness to exert energy in support of the organization (Macey & Schneider, 2008). However, in contrast to commitment, engagement’s attachment and identification refers rather to the work itself than to the organization as a whole (Christian et al., 2011). In contrast to involvement – which is a rather cognitive state (Kanungo, 1982) – engagement has a strong affective connotation (Macey & Schneider, 2008). Even though (state) engagement may share substantial aspects with other constructs, in its combination these aspects create a unique concept, which adds value over and above these constructs (Leiter & Bakker, 2010). More specifically, (state) engagement seems to be a more affective and energetic state than satisfaction, commitment, empowerment, or involvement.

Engagement as behavior. Some authors consider engagement rather as an organizational behavior that contains something special, extra, or atypical (for an overview see Macey & Schneider, 2008). For example, Simbula, Guglielmi, Schaufeli, and Depolo (2013, p. 43) define engagement as “giving it their all.” Behavioral engagement includes several aspects of other well-established constructs like extra-role behavior, personal initiative, or proactive behavior. However, it is not just about doing something more, but it is rather about doing something different (Leiter & Bakker, 2010). Similar to personal initiative (Frese & Fay, 2001), engagement contains self-starting, proactive, persistent, and also highly adaptive behaviors to serve organizational purposes (Macey & Schneider, 2008). Taken together, engagement behaviors include especially innovative, proactive behaviors that go beyond what is ordinarily expected.

Engagement as a trait. Some authors regard engagement as a disposition, or as an enduring tendency to experience state affect over time (for an overview see Macey & Schneider, 2008). In line with the concept of autotelic personality, engaged employees proactively pursue activities for their own sake rather than for extrinsic reasons (Macey & Schneider, 2008). Furthermore, similar to highly conscientious people, engaged employees can be characterized as hard working, ambitious, confident, and resourceful individuals (Macey & Schneider, 2008). However – in contrast to workaholics – engaged employees work hard because for
them work is challenging and fun, and not because they cannot resist their inner urge to do so (Leiter & Bakker, 2010). All these aspects emphasize the disposition or tendency to experience work as something predominantly (and permanently) positive, active, and energetic (Macey & Schneider, 2008). Leiter and Bakker (2010) conclude that some employees are dispositionally more prone to being engaged at work than others.

All these definitions emphasize that engagement is a desirable condition, has an organizational purpose, and connotes involvement, commitment, passion, enthusiasm, and energy (Macey & Schneider, 2008). However, taking into consideration all these aspects of engagement and using it as an umbrella term for trait, state, and behavioral engagement leads to obscuring rather than solving the conceptual problem (Leiter & Bakker, 2010). To make the most of »work engagement«, a clear conceptualization and operationalization is indispensable. In line with Schaufeli, Salanova, González-Romá, and Bakker (2002), I consider work engagement as a motivational concept and – referring to the definitions mentioned above – as a psychological state. Schaufeli, Salanova et al. (2002, p. 74) define work engagement as “a positive, fulfilling, work-related state of mind that is characterized by vigor, dedication, and absorption.” This is a specific, well-defined, and – with the Utrecht Work Engagement Scale (UWES) – properly operationalized psychological state, which is open to both empirical research and practical application (Leiter & Bakker, 2010).

Work Engagement and its Integration in the JD-R Framework

The first theoretical model which introduced work engagement conceptually was the job demands-resources (JD-R) framework (Bakker & Demerouti, 2007; Demerouti, Nachreiner, Bakker, & Schaufeli, 2001). The JD-R framework was originally designed to identify possible causes of burnout (Demerouti et al., 2001). The framework implies that every job includes job demands as well as job resources (Schaufeli, 2017). Job demands can be defined as “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” (Demerouti et al., 2001, p. 501), or – as Schaufeli (2017, p. 121) stated – the “bad things” at work. Contrary to job demands, job resources can be defined as “positively valued physical, psychological, social, or organizational aspects of the job that are functional in achieving work goals, reduce job demands, or stimulate personal growth and development” (Schaufeli & Taris, 2014, p. 56), or – as Schaufeli (2017, p. 121) stated – the “good things” at work.
According to Demerouti et al. (2001), excessive job demands require additional effort to achieve work-related goals, resulting in exhaustion. Similarly, a lack of job resources also hinders the achievement of work-related goals and leads to disengagement – the motivational component of burnout. **Burnout** as the central pathogenic outcome within the JD-R framework is defined as a consequence of extended exposure to specific job demands like intense physical, affective, and cognitive strain (Demerouti, Bakker, Vardakou, & Kantas, 2003). However, given the recent observable trend in OHP to also focus on positive aspects of health, Schaufeli and Bakker (2004) established a revision of the JD-R framework, which includes work engagement as a second dimension of well-being. **Work engagement** is defined as “a positive, fulfilling, work-related state of mind that is characterized by vigor, dedication, and absorption” (Schaufeli, Salanova et al., 2002, p. 74). This positive psychology twist enables the description of salutogenic mechanisms, expressed through the motivational process within the revised JD-R. This revised JD-R framework implies two causal – essentially independent – processes: the **health impairment process** and the **motivational process**: High job demands increase the risk for burnout and lead to **negative outcomes** such as health complaints, whereas job resources play a motivational role, stimulate work engagement and foster **positive outcomes**, such as job performance or organizational commitment (see Figure 1).

**Figure 1.** The job demands-resources framework (Schaufeli, 2017).

The JD-R framework has stimulated a considerable amount of research in OHP. In the following paragraphs, I will shed light on the existing empirical evidence on the JD-R framework and, more specifically, on the motivational process – and on what still needs to be done.
Empirical Evidence on the JD-R Framework and its Motivational Process

Several reviews and meta-analyses have examined several aspects of the JD-R framework and – more specifically – of the motivational process within the JD-R framework (e.g. Christian et al., 2011; Crawford, LePine, & Rich, 2010; Halbesleben, 2010; Mauno, Kinnunen, Mäkikangas, & Feldt, 2010; Nahrgang, Morgeson, & Hofmann, 2011). Halbesleben (2010) and Christian et al. (2011) found significant positive correlations between job resources and work engagement ($\rho$ varies between .24 and .53). Nahrgang et al. (2011) included associations from work engagement on safety outcomes in their meta-analysis and found that work engagement strengthened employees’ behavior towards working safely ($\rho$ varies between .08 and .32). Furthermore, Crawford et al. (2010) calculated a meta-analytic structural equation model to test the assumptions of the JD-R framework. In line with the JD-R framework, they found strong relationships between job resources and work engagement ($\rho = .36$), and between job demands and burnout ($\rho = .25$). However, the results of these meta-analyses and reviews rely mostly on cross-sectional studies. Only Mauno et al. (2010) differentiated between cross-sectional and longitudinal studies. In their qualitative review, they conclude that especially autonomy and social support lead to work engagement over time.

These systematic reviews and meta-analyses reflect the multifaceted approach of the JD-R framework and the enormous boom of research on job-related antecedents of work engagement within the last two decades. They support the JD-R’s associations between job characteristics and occupational well-being, and specifically the motivational process within the JD-R framework: job resources are associated with work engagement. However, these meta-analyses and reviews suffer from several shortcomings. Since they do not predominantly rely on longitudinal data, they are inappropriate to test the causal relationships implied by the JD-R framework. To validate the JD-R framework and its essential assumptions (i.e. the motivational process), longitudinal studies are necessary. Another aspect which has been neglected is the question of whether specific job resources have differential effects on work engagement over time. Saks and Gruman (2014, p. 163) conclude that we still need to explain “what [job] resources will be most important for [work] engagement or why some [job] resources might be more important than others for facilitating [work] engagement.”
Research Questions

Due to these research gaps, my aims were to examine and potentially validate the essential assumptions made by the JD-R framework longitudinally in order to empirically test its claim as the best theoretical model for research on work engagement. Also, as called for by Saks and Gruman (2014), I intended to identify the key drivers of work engagement over time. Hence, I formulated the following research questions:

Research Question 1: Does the longitudinal evidence confirm the essential assumptions made by the JD-R framework?

Research Question 2: Are job resources at different levels related to work engagement over time?

Research Question 3: Which job resources are the most important drivers of work engagement over time?

However, since work engagement as a concept is not restricted to research on the JD-R framework, I was also interested in its applicability in specific occupational and non-occupational contexts. Specifically, I aimed to broaden the theoretical and contextual application of work engagement. Hence, I intended to integrate work engagement into another well-known theoretical model that assesses occupational stress and well-being: the effort-reward imbalance (ERI; Siegrist, 1996) model. Above all, I was interested to see which factors within the ERI model are most important for work engagement.

Research Question 4: Within the ERI model, which factor is the key driver for work engagement – high effort, or any of the reward systems?

Furthermore – since work engagement is associated with various positive health- and performance-related outcomes – I wanted to apply the concept and its operationalization to the context of university students. From a psychological perspective, studying may be considered as work (Ouweneel, Le Blanc, & Schaufeli, 2011). Hence, the concept of student engagement has also raised more and more interest. With the Utrecht Work Engagement Scale – Student Form (UWES-SF; Schaufeli, Martinez, Marques Pinto, Salanova, & Bakker, 2002) there already existed a measurement to capture student engagement with 17 items. However, to date, an evaluated German version of the shortened 9-item UWES-SF does not exist. Hence,
I intended to validate a German version of the UWES-9-SF. Another purpose was to develop and validate a German ultra-short version of the UWES-SF with only 3 items for the application in national and epidemiological surveys.

**Research Question 5**: Do the shortened versions of the UWES-SF – the UWES-9-SF and the UWES-3-SF – appear to be reliable and valid indicators of student engagement?

Two answer the first three research questions, I conducted two meta-analytic reviews. In the first meta-analytic review, I included all available longitudinal studies on the JD-R framework to validate its essential assumptions. In the second meta-analytic review I used data only on the longitudinal associations between job resources and work engagement – the motivational process within the JD-R framework. In both reviews, I used the innovative meta-analytic structural equation modeling (metaSEM; Cheung, 2014) approach to examine these research questions. This approach is still very uncommon in psychological research, especially on OHP topics. However, metaSEM makes it possible to integrate data from all available studies into one (theoretical) model, and to test this model even if it had not been tested in the primary studies. To answer the fourth research question, I used data from a health monitoring program at a large urban police department in Germany. To tackle the fifth research question, I used data from a health monitoring program at a large German university. Taken together, my research should emphasize the broad applicability of the concept »work engagement«.

**References**

[https://doi.org/10.1108/02683940710733115](https://doi.org/10.1108/02683940710733115)


The Job Demands-Resources Model: A Meta-Analytic Review of Longitudinal Studies


https://doi.org/10.1080/02678373.2018.1529065
Abstract

The Job Demands-Resources (JD-R) Model is an influential framework to understand how job characteristics foster employee well-being. Differing from the cross-sectional focus of most JD-R model reviews, this meta-analytic review uses longitudinal evidence to validate the essential assumptions within the JD-R model. We highlight two aspects: (1) The assessment of the methodological quality of the available longitudinal studies, and (2) the examination of the essential assumptions – the impact of job characteristics on well-being, and vice-versa – with meta-analytic structural equation modeling (metaSEM). We included 74 studies and evaluated their quality: Eighteen studies suffered from serious methodological shortcomings, whereas 29 studies (39 %) were considered high-quality studies. Our meta-analytic structural equation models confirmed the essential assumptions simultaneously. The quality of the studies also had an influence on the goodness-of-fit indices. The models with reciprocal assumptions between job characteristics and well-being fitted the data best. The findings of this meta-analytic review suggest that the JD-R model is an excellent theoretical basis to assess employee well-being for a broad range of organizations. However, more research is needed to clarify the reciprocal relationships between job characteristics and employee well-being, as well as the different influences of challenging and hindering job demands on work engagement.

Keywords

Engagement, Burnout, Demands, Resources, Meta-Analytic Review, Employee Well-Being
Introduction

The Job Demands-Resources Model (JD-R; Bakker & Demerouti, 2007; Demerouti, Nachreiner, Bakker, & Schaufeli, 2001) is currently the most popular framework in occupational health psychology to investigate the relationships between job characteristics and employee well-being. For instance, a quick search in Google Scholar (July 2018) reveals about 14,000 hits for the »Job Demands-Resources Model«, against only 3,400 for Karasek’s »Job Demand-Control Model«. The JD-R model claims that specific job characteristics lead to well-being, which in turn influences job performance. To examine these causal assumptions – as defined by Cook and Campbell (1979) – properly, longitudinal studies are necessary (Dormann, Zapf, & Perels, 2010).

However, most empirical studies using the JD-R model support only the conclusion that variables are correlated but do not imply causal relations concerning the impact of job characteristics on well-being. Systematic reviews and meta-analyses available to date have also neglected this aspect of causality. Cross-sectional designs are ill-suited to test causal relationships, since they do not provide evidence regarding the temporal order of the variables. Also, they are not suited to examine reverse and reciprocal causal relationships (de Lange, Taris, Kompier, Houtman, & Bongers, 2003). Our study aims to close this knowledge gap by providing a necessary and timely longitudinal meta-analysis. To this end, we aggregated only data of longitudinal studies and used meta-analytic structural equation modeling to test for causal relationships. One major advantage of meta-analytic structural equation modeling is that it allows an integration of the given data from all studies into one model and to specify models that have not been tested in the primary studies. To make the most of this approach, we therefore not only examined the essential assumptions of the JD-R model but also the reciprocal effects from well-being on job characteristics.

Additionally, we set out to rectify another oversight of earlier reviews and meta-analyses by taking into account the methodological quality of the studies. Quality assessment of primary studies is an essential and mandatory part of meta-analyses (Dreier, 2013). Insufficient rigor and stringency in study design can lead to biased results, which in turn also biases the conclusions reached by reviews and meta-analyses (Kristensen, 1995). Quite possibly, using only high-quality studies enables a more precise examination of the pathways proposed by
the JD-R model. Therefore, we evaluated not only the results but also the methodological quality of longitudinal JD-R model research.

An Overview of the JD-R Model

The JD-R model was introduced in 2001 to identify possible antecedents of burnout (Demerouti et al., 2001). According to Demerouti et al. (2001), achieving work-related goals requires additional effort in the presence of excessive job demands, thus resulting in exhaustion. Similarly, a lack of job resources also encumbers the achievement of work-related goals and leads to disengagement. Given the recent observable trend in occupational health psychology to focus on positive aspects of health, Schaufeli and Bakker (2004) later supplemented these ideas by adding work engagement as a positive dimension of well-being into the JD-R model. This positive psychology twist allows for a description of salutogenic mechanisms, expressed through the motivational process within the revised JD-R model. The central idea of the JD-R model is that working conditions, which are specific to every occupation, can generally be classified as either job demands or job resources (Bakker & Demerouti, 2007). Thus, the JD-R model should be applicable to various occupational settings in order to detect the consequences of specific work environments on employee well-being and employee performance.

*Job demands* can be described as “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” (Demerouti et al., 2001, p. 501). Prevalent examples are time and work pressure, role conflicts, or quantitative workload. Crawford, LePine and Rich (2010) enlarge this definition through the differentiation between hindering and challenging job demands. Challenging job demands may promote employee's personal growth and future gains and tend to be perceived as opportunities to learn, whereas hindering job demands may thwart employee's personal growth and tend to be perceived as constraints or barriers (Crawford et al., 2010). Contrary to job demands, *job resources* can be defined as “physical, psychological, social, or organizational aspects of the job that may [...] be functional in achieving work goals, reduce job demands and its related costs, or stimulate personal growth and development” (Demerouti et al., 2001, p. 501), such as autonomy, and organizational or supervisor support.
According to Demerouti et al. (2001) the central pathogenic health indicator within the JD-R model is burnout, most often operationalized as exhaustion – the energetic component of burnout. Demerouti, Bakker, Vardakou and Kantas (2003) define exhaustion as a consequence of extended exposure to specific job demands like intense physical, affective, and cognitive strain. As a salutogenic health indicator, Schaufeli, Salanova, González-Romá and Bakker (2002) introduce the concept of work engagement. This concept emphasizes the affective-cognitive state and can be described as “a positive, fulfilling, work-related state of mind that is characterized by vigor, dedication, and absorption” (Schaufeli et al., 2002). The JD-R model implies two causal – essentially independent – processes: the health impairment process and the motivational process. High job demands increase the risk for burnout and lead to negative outcomes, such as health complaints or turnover intentions, whereas job resources play a motivational role, stimulate work engagement and foster positive organizational outcomes, such as performance or organizational commitment (Bakker & Demerouti, 2017). Furthermore, the absence of job resources leads to burnout, whereas job resources can also buffer the impact of job demands on burnout (Bakker & Demerouti, 2017). This model (Schaufeli, 2017) is depicted in Figure 1.

Even though there were several modifications – e.g. through the inclusion of job crafting, personal resources, or job performance (see Bakker & Demerouti, 2017) – the essential assumptions within the JD-R model remain unaffected: (1) job demands predict burnout, (2) job resources predict work engagement, and (3) job resources also have an impact on burnout (see Figure 1). As can be seen in the next paragraph, these assumptions are confirmed by cross-sectional analyses. However, correlation is not causation – the assumptions should be confirmed by longitudinal analyses, too. Therefore, our purpose is to validate these essential assumptions longitudinally. Note that we limit our meta-analytic review to the longitudinal associations between job characteristics and employee well-being because the outcomes within the JD-R are still too heterogeneous to compound. In addition, we were not able to consider the differentiation in challenging and hindering job demands in our models, since this differentiation is still not applied in most of the studies concerning the JD-R model.
Meta-Analytic Evidence

A large body of evidence has examined several aspects of the JD-R model (meta-analyses and reviews: Alarcon, 2011; Christian, Garza, & Slaughter, 2011; Crawford et al., 2010; Halbesleben, 2010; Mauno, Kinnunen, Mäkikangas, & Feldt, 2010; Nahrgang, Morgeson, & Hofmann, 2011). Halbesleben (2010) and Christian et al. (2011) focused on work engagement and the motivational process in their reviews. Their analyses revealed significant correlations between work engagement and job resources ($\rho$ varies between .24 and .53), as well as health indicators ($\rho$ varies between .08 and .59). Alarcon (2011) focused on burnout and the health impairment process, and revealed significant positive correlations between burnout and job demands ($\rho$ varies between .32 and .53), and negative correlations between burnout and job resources ($\rho$ varies between .32 and .53), and negative correlations between burnout and health ($\rho$ varies between -.36 and -.51). Nahrgang et al. (2011) included associations of well-being and safety outcomes in their meta-analysis. Burnout was detrimental to working safely ($\rho$ varies between -.13 and -.32), whereas work engagement strengthened employees’ behavior towards working safely ($\rho$ varies between .08 and .32). Crawford et al. (2010) calculated a meta-analytic structural equation model to test the assumptions of the JD-R model. In line with the JD-R model, they found strong relationships between job resources and work engagement ($\rho = .36$), and between (hindering) job demands and burnout ($\rho = .25$). The results of these meta-analyses and reviews rely mostly on cross-sectional studies. Only Mauno et al. (2010) differentiate between cross-sectional and longitudinal studies. In their qualitative review, they conclude that especially autonomy and social support lead to work engagement over time. Results concerning reciprocal effects between job characteristics and well-being were inconsistent. However, to date, there is no meta-analysis of only longitudinal studies.
The JD-R model has stimulated a considerable amount of research in occupational health psychology. Systematic reviews and meta-analyses reflect the multifaceted approach of the JD-R model and support its proposed associations between job characteristics, well-being, and health or job performance. However, these meta-analyses and reviews suffer from several shortcomings. First, they are not based on longitudinal studies. All reviews rely on cross-sectional studies that are inappropriate to test the causal relationships implied by the JD-R model. To validate the JD-R model and its assumptions, longitudinal studies are necessary. Even more so because longitudinal designs also allow for an examination of reversed or reciprocal relationships, which are – due to the addition of job crafting – one of the major recent extensions of the JD-R model. Second, earlier reviews did not assess the methodological quality of the included studies, which may bias their results.

**Research Questions**

Therefore, the purpose of our study is to conduct a meta-analytic review of longitudinal studies of the JD-R model to test its essential assumptions. Due to the additional research quality assessment, it is possible to examine if methodologically weak studies affect stability and longitudinal effects sizes (Kristensen, 1995).

Specifically, we investigate the following research questions:

1. How many longitudinal studies based on the JD-R model can be considered high-quality studies?
2. Do the longitudinal causal effect sizes confirm the essential assumptions of the JD-R model as stated by Schaufeli (2017)? In other words, does the data confirm the following hypotheses?
   - Hypothesis a: Excessive job demands lead to burnout.
   - Hypothesis b: Abundant job resources lead to work engagement.
   - Hypothesis c: Poor job resources contribute to burnout.

To address the several modifications of the JD-R model we also tested two more aspects: First, through the inclusion of job crafting, the JD-R model also predicts a reciprocal impact of employee well-being on job characteristics. Second, because all job demands are
combined into one category – i.e. not differentiated as hindering and challenging job demands as proposed by Crawford et al. (2010) – the path from job demands to work engagement (and vice versa) should be non-significant.

3. Do the longitudinal causal effect sizes confirm the additional assumptions of the JD-R model? In other words, does the data confirm the following hypotheses?
   - Hypothesis d: Low job demands do not lead to work engagement.
   - Hypothesis e: Abundant work engagement does not impact job demands.
   - Hypothesis f: The reciprocal model will be the most appropriate model.

**Method**

**Literature Search**

The purpose of our meta-analytic review was to identify preferably all longitudinal studies using the JD-R model as a theoretical framework. In a first step, we identified ten key studies that met the inclusion criteria (see next paragraph) to create a replicable search strategy with relevant keywords. In a second step, we searched several psychology-specific and general databases: PsycARTICLES, PsycINFO, PSYNDX and Academic Search Ultimate. In a third step, we inspected the references of the identified studies to find further longitudinal studies of the JD-R model. The respective steps are documented in the results section.

**Inclusion Criteria and Study Coding**

To be included, studies had to meet the following four criteria: The study (1) investigated employees based on the JD-R model; (2) used a longitudinal design with two or more waves; (3) measured at least one job characteristic (demand or resource), and (4) a dimension of well-being (work engagement or burnout).

To select the final body of evidence we chose a two-step approach. First, we checked all abstracts for the inclusion criteria. If the abstract was relevant, we acquired the full text and examined it in a second step. Studies utilizing (partly) the same data set were synthesized into one study to consider independent effect sizes. We documented the following information for identification and quality assessment of the study, and the statistical analyses afterwards: bibliographic information, sector of the sample, sample size, effect sizes between relevant indicators, study design, number of and interval between measurements, psychometric quality of the measures, method of analysis, and response rate and nonresponse analysis.
Quality Assessment

For quality assessment, we used a scheme from de Lange, Taris, Kompier, Houtman, and Bongers (2003), which was developed to evaluate the quality of studies in occupational health psychology. It consists of five criteria: (1) study design, (2) number and interval of measurements, (3) psychometric quality of the measures, (4) statistical analyses, and (5) nonresponse analysis. Studies could obtain between one (insufficient) and four (very good) stars for each criterion.

The criterion study design refers to the possibility of examining causal assumptions. Only if there is a statistical significant relationship between predictor and outcome, the predictor precedes the effect in the outcome over time, the relationship is not due to a third variable, and there is a plausible argument for the relationship, cross-lagged effects can be interpreted as causal (Shadish, Cook, & Campbell, 2002). Studies that used a full panel design with more than three measurements were rated as very good.

The criterion number and interval of measurements is necessary to detect and interpret cross-lagged effects. There are different recommendations for time lags in longitudinal studies. Dormann and Griffin (2015) try to deduce optimal time lags statistically by focusing on variable stabilities. They recommend pilot studies (with a quite short interval) to identify the stabilities, and to adjust the time lags to these stabilities. Usually the optimal time lags are much shorter than those frequently used in occupational health psychology. Studies that realized more than one interval with a plausible methodological or theoretical argument were rated as very good.

The psychometric quality of the measures of the variables has decisive influence on the validity of the statistical conclusions. We choose to use the reliability of the instruments measuring job characteristics and well-being as the core criterion for our assessment. If additional details concerning the validity of the instruments were given, the rating was further upgraded. Studies that used instruments with reliabilities of at least .70, and additionally validated one or more external criteria, were rated as very good.

The criterion method of analysis addresses the calculated statistical models used to prove causal effects. Therefore, the stabilities of the variables of interest are taken into account.
Studies that used structural equation modeling and tested different models (standard, reversed, reciprocal) against each other were rated as very good.

The last evaluation criterion, *nonresponse analysis*, assesses possible response selectivity. Relevant for the rating is the response rate as well as selectivity being statistically controlled for. Studies that controlled for demographic variables (age, gender) and focal variables (job characteristics and well-being), and had a response rate over 80% were rated as very good.

Table 1 summarizes the criteria and ratings. The authors used this scheme to independently rate the studies. Information concerning the execution, problems, and results of the quality assessment is given in the result section.

### Table 1.

*Criteria for evaluating quality of longitudinal research – modified scheme from de Lange et al. (2003)*

<table>
<thead>
<tr>
<th>Criteria</th>
<th>*</th>
<th>**</th>
<th>***</th>
<th>****</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study Design</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predictor and outcome measured on different occasions</td>
<td>At least one variable measured on more than one occasion</td>
<td>All variables measured on two occasions (full panel design)</td>
<td>All variables measured on more than two occasions (full panel design with &gt;2 occasions)</td>
<td></td>
</tr>
<tr>
<td><strong>Number and Interval of Measurements</strong></td>
<td>One time lag without argument</td>
<td>One time lag with plausible theoretical/methodological argument</td>
<td>More than one time lag without argument</td>
<td>More than one time lag with plausible theoretical/methodological argument</td>
</tr>
<tr>
<td><strong>Psychometric Quality of the Measures</strong></td>
<td>Insufficient/controversial information</td>
<td>Acceptable reliabilities of the measures (Cronbach’s α = .6-.7)</td>
<td>Good reliabilities of the measures (Cronbach’s α &gt; .7)</td>
<td>Good reliabilities of the measures (Cronbach’s α &gt; .7) and validation with at least one external criteria</td>
</tr>
<tr>
<td><strong>Method of Analysis</strong></td>
<td>Correlational analysis</td>
<td>Regression analysis considering stability of predictor or outcome</td>
<td>Regression analysis considering stability of predictor and outcome</td>
<td>Structural equation modeling and testing of different models (standard, reversed, reciprocal)</td>
</tr>
<tr>
<td><strong>Nonresponse Analysis</strong></td>
<td>No check of selectivity of the sample</td>
<td>Check on selectivity referring demographic or focal variables</td>
<td>Check on selectivity referring demographic and focal variables</td>
<td>Check on selectivity referring demographic and focal variables, and response rate over 80%</td>
</tr>
</tbody>
</table>

*Note. * = insufficient; ** = sufficient; *** = good; **** = very good*

#### Analysis Strategy

We used a two-stage structural equation modeling (TSSEM) approach (Cheung & Chan, 2005) to test the essential assumptions of the JD-R model. Unlike the meta-analytic structural equation modeling (MASEM) approach (Viswesvaran & Ones, 1995), TSSEM takes...
the homogeneity of the covariances into account. Thus, more homogenous effect sizes are weighted more heavily in the calculation of the pooled correlation matrix. Landis (2013) recommends TSSEM as the method of choice if at least one study provides a full correlation matrix—a great potential to evaluate research questions which have not been examined within the individual studies before. Therefore, we chose TSSEM.

In a first step, we combined the relevant effect sizes into matrices to calculate a pooled correlation matrix. Since we concentrated on the essential assumptions of the JD-R model, we decided to compute a matrix that contains information about the latent variables assumed by the JD-R model (job demands T1, T2; job resources T1, T2; burnout T1, T2; work engagement T1, T2). Luchman and Gonzalez-Morales (2013) used the Spearman-Brown formula from Hunter and Schmidt (2015, p. 444) to compute composite correlations in their review about the Demand-Control-(Support) Model. We, too, used the Spearman-Brown formula to obtain a composite effect size estimate if a study provided several correlations between the variables (e.g. more than one job resource with work engagement). After calculating an appropriate matrix for each study, we used the metaSEM package (Cheung, 2014) version 0.9.14 in R 3.3.3 to estimate a pooled correlation matrix. Since the studies used different scales and instruments—especially concerning job demands and job resources—we chose random effects models.

In a second step, we input the pooled correlation matrix to conduct the structural equation models in metaSEM. Following Dormann et al. (2010), we allowed correlations between latent variables at T1, but not at T2. To assess the causality of the essential assumptions, we specified a model (M1), which only included the stabilities of the variables. Afterwards, we added the paths of interest as follows: The standard model (M2) included the cross-lagged paths from job characteristics to well-being. The reversed model (M3) contained the inversed paths from well-being to job characteristics. Last, the reciprocal model (M4) resulted from a combination of M2 and M3.

We assessed the models’ goodness-of-fit by chi square test statistic, Root Mean Square Error of Approximation (RMSEA), Standardized Root Mean Square Residual (SRMR), Tucker-Lewis Index (TLI), and Comparative Fit Index (CFI) as recommended by Hu and Bentler (1998). A non-significant chi square indicates good model fit. A SRMR value of .08 or lower and a
RMSEA value of less than .06 indicate good model fit (Hu, L. & Bentler, 1999). For CFI and TLI values of .95 may be interpreted as an acceptable fit, while .97 is indicative of a good model fit (Schermelleh-Engel, Moosbrugger, & Müller, 2003). We compared the models using the Satorra-Bentler scaled chi-squared difference test (Satorra, 2000) to determine the best-fitting model.

Results

Systematic Literature Search

The electronic search in the selected databases yielded 625 potentially relevant documents. We manually added 11 documents by looking through the references of the identified documents. PsycINFO yielded most search results ($k = 326$). PSYNDEX ($k = 39$) and PsycARTICLES ($k = 43$) contributed far less search results. Lastly, the search request of the multidisciplinary database Academic Search Ultimate resulted in $k = 217$ documents. After the deletion of duplicates, 438 documents remained in the pool. The abstracts and full texts of approximately half ($k = 189$) of the studies in the pool were assessed independently by two raters. The initial inter-rater reliability (Cohens $\kappa = .94$) was found to be very good according to the guidelines of Wirtz and Caspar (2002). In case of disagreements, we returned to the primary study and resolved ambiguous coding through a consensus discussion.

The final sample consisted of $k = 83$ documents that fulfilled the inclusion criteria after reviewing the abstracts and full texts. Documents, which included the same data, were synthesized into one study. Following the document synthesis, $k = 74$ studies and 77 samples remained as the body of evidence of our meta-analytic review. Figure 2 shows a flowchart of our literature search.
The Job Demands-Resources Model: A Meta-Analytic Review of Longitudinal Studies

625 documents in following databases
Academic Search Premier \( (k = 217) \)
PSYNDEx \( (k = 39) \)
PsycINFO \( (k = 326) \)
PsycARTICLES \( (k = 43) \)

11 additional documents from references

438 documents after deletion of duplicates

331 documents were excluded after reviewing the abstract

107 potential relevant documents were gathered in full text

24 documents were excluded after reviewing the full text

83 included documents in the meta-analytic review

74 included studies in the meta-analytic review (77 samples)

Since the JD-R model is supposed to be valid for diverse work environments, it is required to cover different occupational groups that present variation on the job characteristics dimensions. Kristensen (1995) and de Lange et al. (2003) argue (concerning the Demand-Control-(Support) Model) that heterogeneous populations and therefore a broader range of job characteristics are even more important than the representativeness of the study samples. We extracted detailed information concerning the population, the number and interval of measurements, the measurements of job characteristics and well-being, and the analysis strategy.

Population. Altogether, the virtual population consists of 37,324 employees. The average sample size is 450. The studies were conducted in various cultural contexts and occupational groups, with different gender ratios. The majority of the studies was realized in Europe \( (k = 46) \), followed by Asia \( (k = 11) \), Australia \( (k = 8) \), North America \( (k = 6) \), and Africa \( (k = 2) \). The gender distribution also varies between the studies: Some samples consist of females only (Hu, Q., Schaufeli, & Taris, 2017; Lu, Siu, Chen, & Wang, 2011), while one sample is completely...
The sample size varies between $N = 75$ (Wirtz, N., Rigotti, Otto, & Loeb, 2017) and $N = 2897$ (Frins, van Ruysseveldt, van Dam, & van den Bossche, 2016).

**Number and interval of measurements.** Most of the studies used a two-wave ($k = 50$) or three-wave ($k = 25$) study design, whereas two studies realized more than three measurements (Fong et al., 2016; Kühnel & Sonnentag, 2011). The time lags between the measures varied between four days (Park & Lee, 2015) and 10 years (Airila et al., 2014) with a mean interval of eleven months.

**Measurement of job characteristics.** The operationalization of job characteristics was as expected very diverse, too. Only the Job Content Questionnaire (JCQ; Karasek, 1979), the Questionnaire on the Experience and Evaluation of Work (VBBA; van Veldhoven & Meijman, 1994), the Copenhagen Psychosocial Questionnaire (COPSOQ; Kristensen, Hannerz, Høgh, & Borg, 2005), and the Instrument for Stress-Oriented Job Analysis (ISTA; Zapf, 1993) were used more than five times. This goes to show that a broad variety of job characteristics and instruments are utilized in the context of JD-R research.

**Measurement of well-being.** The operationalization of well-being was more homogeneous: Primarily, the Utrecht Work Engagement Scale (UWES; Schaufeli & Bakker, 2003) and the Maslach Burnout Inventory (MBI; Maslach, Jackson, & Leiter, 1996) were used to measure well-being. Furthermore, the Oldenburg Burnout Inventory (OLBI; Halbesleben & Demerouti, 2005) was used in six studies to measure burnout, and the Copenhagen Burnout Inventory (CBI; Kristensen, Borritz, Villadsen, & Christensen, 2005) in one study. All studies except two assessed work engagement with the UWES (de Beer, Pienaar, & Rothmann Jr., 2016; Oliver, 2013).

**Analysis strategy.** In most studies structural equation modeling ($k = 42$) or multiple/hierarchical regressions ($k = 24$) were realized. Beyond that, two studies used growth curve modeling, four studies conducted analysis of variance/covariance (ANOVA, ANCOVA, MANCOVA), one study used hierarchical linear modeling, and one study growth mixture modeling. In doing so, 27 studies focused explicitly on the motivational process, 20 studies focused only on the health impairment process, while 25 studies considered both processes simultaneously. One study examined only reversed causation effects (de Beer, Pienaar, & Rothmann Jr., 2013),
another study modelled merely the association between job demands and work engagement (Lu et al., 2011).

**Additional information.** Our body of evidence contains two dissertations (Kim, 2008; Oliver, 2013) as well as two articles from book chapters (Millear & Liossis, 2010; Nakamura & Otsuka, 2013). All remaining studies were published in peer-reviewed journals. An overview of the studies and further information are available from the first author.

**Quality Assessment of the Identified Studies**

Table 2 presents the results of the quality assessment based on the described scheme shown in Table 1.

**Study design.** In four studies (5 %), job characteristics and well-being were measured only once on different occasions, that is why these studies were classified as **insufficient.** Sixteen studies measured at least one construct at only one occasion. These studies were classified as **sufficient** because the stability for at least one construct could be considered in the analysis. Altogether 40 studies realized a full panel design with two measurements and were evaluated as **good.** The remaining studies (k = 18; 24 %) measured all constructs on more than two occasions and were classified as **very good.**

**Number and interval of measurements.** As mentioned above, most studies realized two or three measurements – with only two exceptions (Fong et al., 2016; Kühnel & Sonntag, 2011). Thirty-two studies (43 %) used a two-wave study design without any theoretical or methodological argument to justify the time lag and were rated as **insufficient.** In 19 studies, the authors gave an adequate argument to justify the time lag for their two-wave design. These studies were evaluated as **sufficient.** Since the possibility is given to test different time lags against each other, studies with three or more occasions were rated as **good** or **very good.** 19 studies lacked an appropriate theoretical or methodological argument, and were thus classified as **good,** while seven studies (9 %) realized more than two occasions and justified their time lags with plausible arguments. These studies were rated as **very good.**
Table 2.

**Evaluation of the identified studies of the JD-R model**

<table>
<thead>
<tr>
<th>Primary Study</th>
<th>Number/Interval of Measurements</th>
<th>Nonresponse Analysis</th>
<th>Quality of Measures</th>
<th>Study Design</th>
<th>Method of Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adriaenssens et al., 2015</td>
<td>**</td>
<td>*</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Airla et al., 2014</td>
<td>**</td>
<td>***</td>
<td>**</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Akkermans, Breninkmeijer, van den Bossche, Blonk, &amp; Schaufeli, 2013</td>
<td>*</td>
<td>**</td>
<td>***</td>
<td>***</td>
<td>****</td>
</tr>
<tr>
<td>Ángelo &amp; Chambel, 2015</td>
<td>**</td>
<td>***</td>
<td>**</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Barbier, Dardenne, &amp; Hansez, 2013</td>
<td>****</td>
<td>**</td>
<td>***</td>
<td>***</td>
<td>****</td>
</tr>
<tr>
<td>Barbier, Hansez, Chmiel, &amp; Demerouti, 2013</td>
<td>***</td>
<td>**</td>
<td>**</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Bickerton, Miner, Dowson, &amp; Griffin, 2014</td>
<td>***</td>
<td>**</td>
<td>**</td>
<td>***</td>
<td>****</td>
</tr>
<tr>
<td>Bickerton, Miner, Dowson, &amp; Griffin, 2015</td>
<td>***</td>
<td>**</td>
<td>**</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Biggs, Brough, &amp; Barbour, 2014a</td>
<td>****</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>****</td>
</tr>
<tr>
<td>Biggs, Brough, &amp; Barbour, 2014b</td>
<td>****</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>****</td>
</tr>
<tr>
<td>Brauchli, Schaufeli, Jenny, Füllemann, &amp; Bauer, 2013</td>
<td>****</td>
<td>***</td>
<td>**</td>
<td>***</td>
<td>****</td>
</tr>
<tr>
<td>Brough &amp; Biggs, 2015</td>
<td>*</td>
<td>**</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Brough et al., 2013</td>
<td>**</td>
<td>***</td>
<td>***</td>
<td>**</td>
<td>*</td>
</tr>
<tr>
<td>Chambel, Lopes, &amp; Batista, 2016</td>
<td>*</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Cheng, Mauno, &amp; Lee, 2014</td>
<td>****</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>****</td>
</tr>
<tr>
<td>Chrisopoulos, Dollard, Winefield, &amp; Dormann, 2010</td>
<td>****</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>****</td>
</tr>
<tr>
<td>Consiglio, Borgogni, Di Tecco, &amp; Schaufeli, 2016</td>
<td>**</td>
<td>*</td>
<td>***</td>
<td>**</td>
<td>*</td>
</tr>
<tr>
<td>de Beer et al., 2013</td>
<td>*</td>
<td>*</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>de Beer et al., 2016</td>
<td>***</td>
<td>***</td>
<td>**</td>
<td>***</td>
<td>****</td>
</tr>
<tr>
<td>de Lange, Witte, &amp; Notelaers, 2008</td>
<td>**</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>****</td>
</tr>
<tr>
<td>Demerouti, Le Blanc, Bakker, Schaufeli, &amp; Hox, 2009</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>****</td>
</tr>
<tr>
<td>Dicke, Stebner, Linninger, Kunter, &amp; Leutner, 2017</td>
<td>***</td>
<td>*</td>
<td>***</td>
<td>***</td>
<td>****</td>
</tr>
<tr>
<td>Diestel &amp; Schmidt, 2012</td>
<td>**</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>****</td>
</tr>
<tr>
<td>Dikkers, Jansen, de Lange, Vinkenburg, &amp; Kooij, 2010</td>
<td>*</td>
<td>*</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Dollard &amp; Bakker, 2010</td>
<td>***</td>
<td>**</td>
<td>**</td>
<td>***</td>
<td>**</td>
</tr>
<tr>
<td>Fong et al., 2016</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>****</td>
</tr>
<tr>
<td>Frins et al., 2016</td>
<td>*</td>
<td>**</td>
<td>***</td>
<td>***</td>
<td>****</td>
</tr>
<tr>
<td>Gan, T. &amp; Gan, 2014</td>
<td>****</td>
<td>***</td>
<td>***</td>
<td>**</td>
<td>*</td>
</tr>
<tr>
<td>González-Morales, Peiró, Rodríguez, &amp; Bliese, 2012</td>
<td>**</td>
<td>***</td>
<td>***</td>
<td>**</td>
<td>*</td>
</tr>
<tr>
<td>Gregersen, Vincent-Höper, &amp; Nienhaus, 2016</td>
<td>**</td>
<td>*</td>
<td>***</td>
<td>***</td>
<td>****</td>
</tr>
<tr>
<td>Guglielmi et al., 2016</td>
<td>*</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>**</td>
</tr>
<tr>
<td>Hakanen, Peeters, &amp; Perhoniemi, 2011</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>****</td>
</tr>
<tr>
<td>Hakanen, Perhoniemi, &amp; Toppinen-Tanner, 2008</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>****</td>
</tr>
<tr>
<td>Hakanen, Schaufeli, &amp; Ahola, 2008</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>****</td>
</tr>
<tr>
<td>Seppälä et al., 2015</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>****</td>
</tr>
<tr>
<td>Hall, Dollard, Tuckey, Winefield, &amp; Thompson, 2010</td>
<td>*</td>
<td>*</td>
<td>***</td>
<td>***</td>
<td>****</td>
</tr>
<tr>
<td>Houkes, Winants, &amp; Twellaar, 2008</td>
<td>**</td>
<td>***</td>
<td>**</td>
<td>***</td>
<td>****</td>
</tr>
<tr>
<td>Hu, Q. et al., 2017</td>
<td>*</td>
<td>*</td>
<td>***</td>
<td>***</td>
<td>****</td>
</tr>
<tr>
<td>Authors</td>
<td>Quality</td>
<td>Notes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Huang, J., Wang, Wu, &amp; You, 2016</td>
<td>*</td>
<td>**** *** *** *** ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Huang, Y.-H., Du, Chen, Yang, &amp; Huang, 2011</td>
<td>***</td>
<td>**** *** *** * *</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Huynh, Xanthopoulou, &amp; Winefield, 2013</td>
<td>*</td>
<td>*** *** ** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inoue et al., 2013</td>
<td>*</td>
<td>* *** *** ** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jimenez &amp; Dunkl, 2017</td>
<td>*</td>
<td>** *** *** ** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Johnson &amp; Jiang, 2016</td>
<td>**</td>
<td>** *** *** ** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kim, 2008</td>
<td>*</td>
<td>*** *** ** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kinnunen &amp; Feldt, 2013</td>
<td>**</td>
<td>** * *** *** **** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kubicek, Korunka, &amp; Tement, 2014</td>
<td>*</td>
<td>*** *** ** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kühnel &amp; Sonnentag, 2011</td>
<td>***</td>
<td>* *** **** ** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lizano &amp; Mor Barak, 2012</td>
<td>***</td>
<td>*** *** ** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lizano &amp; Mor Barak, 2015</td>
<td>*</td>
<td>*** *** *** ** *</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lorente Prieto, Soria, Martínez, &amp; Schaufeli, 2008</td>
<td>*</td>
<td>*** *** *** *** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vera, Salanova, &amp; Lorente Prieto, 2012</td>
<td>*</td>
<td>* *** *** *** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lu et al., 2011</td>
<td>*</td>
<td>* *** *** *** ****</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mauno, Kinnunen, &amp; Ruokolainen, 2007</td>
<td>*</td>
<td>** *** *** ** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mauno et al., 2016</td>
<td>***</td>
<td>*** *** ** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Millear &amp; Liossis, 2010</td>
<td>***</td>
<td>** *** *** *** ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nakamura &amp; Otsuka, 2013</td>
<td>***</td>
<td>* *** *** * *</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oliver, 2013</td>
<td>*</td>
<td>* * *** *** *** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ouweneel, Le Blanc, &amp; Schaufeli, 2012</td>
<td>*</td>
<td>*** ** *** *** ****</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Park &amp; Lee, 2015</td>
<td>*</td>
<td>** *** *** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perko, Kinnunen, Tolvanen, &amp; Feldt, 2016</td>
<td>*</td>
<td>*** *** *** *** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peters et al., 2016</td>
<td>*</td>
<td>*** *** ** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Qing &amp; Zhou, 2017</td>
<td>*</td>
<td>* *** *** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ragsdale &amp; Hoover, 2016</td>
<td>*</td>
<td>** *** *** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reis et al., 2015</td>
<td>**</td>
<td>* *** *** *** ****</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schaufeli, Bakker, &amp; van Rhenen, 2009</td>
<td>**</td>
<td>*** *** *** *** ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simbula, Guglielmi, &amp; Schaufeli, 2011</td>
<td>****</td>
<td>*** *** *** *** ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sonnentag, Binnewies, &amp; Mojza, 2010</td>
<td>*</td>
<td>*** *** *** *** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stiglbauer, 2016</td>
<td>**</td>
<td>*** *** *** *** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tims et al., 2013</td>
<td>***</td>
<td>*** *** ** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tims, Bakker, &amp; Derks, 2015</td>
<td>**</td>
<td>*** *** *** ** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upadyaya, Vartiani, &amp; Salmela-Aro, 2016</td>
<td>*</td>
<td>* *** *** ** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vahle-Hinz, 2015</td>
<td>**</td>
<td>*** *** *** *** ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>van den Tooren, de Jonge, Vlerick, Daniels, &amp; van de Ven, 2011</td>
<td>*</td>
<td>** * *** ** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>van Wingerden, Bakker, &amp; Derks, 2017</td>
<td>*</td>
<td>*** *** *** *** ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>van Veghel et al., 2010</td>
<td>**</td>
<td>* ** *** *** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vogt, Jenny, &amp; Bauer, 2013</td>
<td>***</td>
<td>* * *** *** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vogt et al., 2015</td>
<td>****</td>
<td>*** *** *** *** ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wang, Y., Huang, &amp; You, 2016</td>
<td>**</td>
<td>** ** *** *** ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weigl et al., 2010</td>
<td>****</td>
<td>** *** *** *** ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wirtz, N. et al., 2017</td>
<td>*</td>
<td>*** *** *** ** **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xanthopoulou et al., 2009</td>
<td>*</td>
<td>*** *** *** *** ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xanthopoulou, Bakker, &amp; Fischbach, 2013</td>
<td>*</td>
<td>*** *** *** *** ***</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: * = insufficient; ** = sufficient; *** = good; **** = very good; *a = studies with more than one sample. Characteristics with empty cells could not be rated with the scheme for quality assessment.
Psychometric quality of the measures. Four studies (5%) were considered as insufficient on this criterion, because the reliability score of at least one construct was below .60 or the authors did not report any psychometric qualities at all. Seventeen studies used scales and instruments with reliabilities between .61 and .70 and, therefore, were classified as sufficient. Most of the studies \((k = 52)\) reported reliabilities higher than .70 and were evaluated as good. Four studies (5%) additionally validated the scales and instruments used and were therefore rated as very good.

Method of analysis. Six studies (8%) realized solely correlative analyses (job characteristics at \(t_1\) and well-being at \(t_2\)) between the focal constructs. These studies were considered as insufficient. Twenty-eight studies were evaluated as sufficient as they used regression analyses and controlled for the stability of at least one construct (predictor or outcome). In 14 studies, the authors considered the stabilities of predictor and outcome, which earned them a study evaluation of good. Twenty-seven (36%) studies controlled not only for stabilities but also tested different models (e.g. standard, reversed or reciprocal) against each other and hence were classified as very good.

Nonresponse analysis. Nineteen studies (26%) did not examine or report any possible response bias on the first and the follow-up measurement and, therefore, were rated as insufficient. Sixteen studies were classified as sufficient, since the authors explored the selectivity of their sample with regard to demographic or focal aspects. Another 38 studies examined the selectivity with regard to demographic and focal aspects. These studies were rated as good. Four studies (5%) were classified as very good, since they examined the selectivity of their sample and had a response-rate over 80%.

Two independent raters carried out the quality assessment on 30 studies simultaneously, whereas the remaining studies were rated by one rater only. The interrater agreement was Cohens \(\kappa = .86\). In case of disagreements, the raters discussed their evaluations and came to a consensual decision.

As de Lange et al. (2003) recommended in their review of the Demand-Control-(Support) Model, those studies which obtained at least sufficient scores on every criteria were considered as high-quality studies. In this meta-analytic review, 29 studies (39%) are labeled as high-quality studies.
Meta-Analytic Structural Equation Modeling

For calculating the pooled correlation matrix, we used 74 correlation matrices. Three matrices could not be handled in metaSEM since they were not positive definite (Adriaenssens, Gucht, & Maes, 2015; Mauno, Mäkikangas, & Kinnunen, 2016; van Vegchel, de Jonge, Söderfeldt, Dörmann, & Schaufeli, 2004). Table 3 gives an overview of the pooled correlation matrix. All correlations were significant on $p < .001$ level. The correlations correspond in extent and direction to the assumptions of the JD-R model as described by Schaufeli and Taris (2014). In a second step, we estimated four nested models using the pooled correlation matrix. The first model (M1) only included the stabilities of the variables. In the second model (M2), the postulated paths from job characteristics to well-being were added. This standard model (M2) fitted the data adequately ($\chi^2 (14) = 319.64; \text{SRMR} = .077; \text{TLI} = .918; \text{CFI} = .959; \text{RMSEA} = .025$). As common in studies based on the JD-R model, we specified a reversed causation model (M3; paths from well-being on job characteristics) and a reciprocal model (M4; including all paths from M1, M2, and M3). The reciprocal model ($\chi^2 (10) = 203.82; \text{SRMR} = .061; \text{TLI} = .927; \text{CFI} = .974; \text{RMSEA} = .024$) fitted the data better than M2 ($\Delta \chi^2 = 115.82, 4 \text{ df}, p < .001$) and M3 ($\Delta \chi^2 = 97.46, 3 \text{ df}, p < .001$).

Table 3.

Meta-analytic correlation matrix; $k = 74$

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. T1 Job Resources</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. T1 Job Demands</td>
<td>−.20 (43)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. T1 Engagement</td>
<td>.44 (44)</td>
<td>−.11 (35)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. T1 Burnout</td>
<td>−.32 (36)</td>
<td>.41 (37)</td>
<td>−.46 (24)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. T2 Job Resources</td>
<td>.61 (44)</td>
<td>−.15 (27)</td>
<td>.38 (34)</td>
<td>−.25 (24)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. T2 Job Demands</td>
<td>−.15 (28)</td>
<td>.62 (35)</td>
<td>−.08 (25)</td>
<td>.34 (26)</td>
<td>−.20 (27)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. T2 Engagement</td>
<td>.37 (50)</td>
<td>−.08 (37)</td>
<td>.66 (49)</td>
<td>−.36 (24)</td>
<td>.49 (35)</td>
<td>−.12 (25)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. T2 Burnout</td>
<td>−.28 (38)</td>
<td>.36 (39)</td>
<td>−.36 (24)</td>
<td>.65 (42)</td>
<td>−.37 (24)</td>
<td>.48 (26)</td>
<td>−.48 (25)</td>
<td></td>
</tr>
</tbody>
</table>

Note. The figure in parentheses reflects the number of studies from which the meta-analytic correlation was derived ($k$). Correlations $\geq |.08|$ are significant on $p < .001$ level.

To answer the second research question, we depicted the standard model (M2) in Figure 3. All stability coefficients of job characteristics and well-being indicators ranged between $\beta = .57$ and .71, which indicates moderate to large stability (Adachi & Willoughby, 2015). As
postulated within the motivational process of the JD-R model, job resources predict work engagement over time ($\beta = .19; p < .001$). As expected job resources also predict burnout negatively ($\beta = -.17; p < .001$). The assumption within the health impairment process was confirmed, too: job demands predict burnout over time ($\beta = .10; p < .001$). However, job demands do not predict work engagement over time ($\beta = -.01; p > .05$). Adachi and Willoughby (2015) point out that effect sizes in longitudinal studies are often much smaller compared to cross-sectional studies. They argue that even small longitudinal effect sizes are substantial – especially if stability effects are moderate to large, which explains most of the outcome variance. The longitudinal effects we found in our study are comparable to those reported in several primary studies (e.g. Kim, 2008; Vogt, Hakanen, Jenny, & Bauer, 2015; Weigl et al., 2010). Moreover, our model explained 50% of the variance in job resources as well as job demands, 52% of the variance in work engagement, and 48% of the variance in burnout at follow-up. Hence, we confirmed hypotheses a, b, c, and d.

Figure 3. Standard model (M2).

Note. $N = 33,761$; *$p < .05$; **$p < .01$; ***$p < .001$. All exogenous variables were allowed to correlate with each other.
Table 4.
Goodness-of-fit indices and comparison of the competing models, N = 33761

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2$</th>
<th>df</th>
<th>SRMR</th>
<th>TLI</th>
<th>CFI</th>
<th>AIC</th>
<th>RMSEA</th>
<th>$\Delta \chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1 Stability Model</td>
<td>420.19</td>
<td>18</td>
<td>.092</td>
<td>.916</td>
<td>.946</td>
<td>384.19</td>
<td>.026</td>
<td>–</td>
</tr>
<tr>
<td>M2 Standard Model</td>
<td>319.64</td>
<td>14</td>
<td>.077</td>
<td>.918</td>
<td>.959</td>
<td>291.64</td>
<td>.025</td>
<td>–</td>
</tr>
<tr>
<td>M3 Reverse Model</td>
<td>301.28</td>
<td>14</td>
<td>.074</td>
<td>.923</td>
<td>.961</td>
<td>273.28</td>
<td>.025</td>
<td>–</td>
</tr>
<tr>
<td>M4 Reciprocal Model</td>
<td>203.82</td>
<td>10</td>
<td>.061</td>
<td>.928</td>
<td>.974</td>
<td>183.82</td>
<td>.024</td>
<td>M1-M2 100.55**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>M1-M3 118.91**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>M1-M4 216.37**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>M2-M4 115.82**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>M3-M4 97.46**</td>
</tr>
</tbody>
</table>

Note. df = degrees of freedom; SRMR = Standardized Root Mean Square Residual; TLI = Tucker-Lewis-Index; CFI = Comparative-Fit-Index; AIC = Akaike Information Criterion; RMSEA = Root Mean Square Error of Approximation. Model comparisons computed with the Satorra-Bentler scaled chi-squared difference test. * $p < .01$; ** $p < .001$.

As can be seen in Table 4, we found the reciprocal model (M4) to be the most appropriate. It is depicted in Figure 4. The stabilities of job characteristics and well-being remain moderate to large ($\beta = .55-.65; p < .001$) compared to those of M2. In addition, the paths from job characteristics to well-being remain significant: job resources predict work engagement ($\beta = .18; p < .001$) and burnout ($\beta = -.16; p < .001$) over time, and job demands predict burnout over time ($\beta = .13; p < .001$). However, job demands do not predict work engagement ($\beta = -.02; p < .001$). Furthermore, the reversed paths from well-being to job characteristics are significant, too: work engagement predicts job resources over time ($\beta = .18; p < .001$), and burnout predicts job demands ($\beta = .20; p < .001$) as well as job resources ($\beta = -.06; p < .05$) over time. Only work engagement does not predict job demands over time ($\beta = .04; p > .05$). This model explained 44% of the variance in job resources as well as job demands, 55% of the variance in work engagement, and 52% of the variance in burnout at follow-up.
To integrate research questions 1 and 2, we re-estimated the meta-analytic structural equation modeling using only the high-quality studies. Similar to the models including all studies, the reciprocal model ($\chi^2 (10) = 53.13$; SRMR = .052; TLI = .969; CFI = .989; RMSEA = .017) fitted the data better compared to the stability, standard, and reverse models. The reciprocal model of high-quality studies is depicted in Figure 5. In comparison to the reciprocal model including all studies (Figure 4), the stability of burnout within the high-quality model is higher ($\beta = .64$ vs. $\beta = .60$). That is why burnout at T1 explains most of the variance of burnout at T2, and the path coefficient from job demands on burnout is smaller, although still significant ($\beta = .10; p < .05$). The reverse path from burnout to job resources is larger ($\beta = -.16; p < .05$) than in the reciprocal model including all studies (Figure 4). Also, the paths between job demands and work engagement are non-significant. All other coefficients are comparable to those we found in the reciprocal model including all studies (see Figure 5). This model explained 46% of the variance in job resources, 45% of the variance in job demands,
57% of the variance in work engagement, and 55% of the variance in burnout at follow-up. Thus, we confirmed hypotheses e, and f.

Figure 5. Reciprocal model of the high-quality studies.

Note. N = 14,486; *p < .05; **p < .01; ***p < .001. All exogenous variables were allowed to correlate with each other; dotted lines represent reversed effects.

Discussion

The aim of the present meta-analytic review was to summarize the existing evidence for the causal pathways proposed by the JD-R model, to assess the quality of the realized studies, and to validate the essential as well as the reciprocal assumptions within the JD-R model. In contrast to previous meta-analyses and reviews, we focused on two aspects: We exclusively considered longitudinal studies to prove the causal assumptions made by the JD-R model, and we used a validated scheme from de Lange et al. (2003) to obtain the methodological quality of all studies in order to detect and combat potential biases caused by poor methodology.
Altogether, we identified 74 studies, which used the JD-R model as their theoretical framework and met the inclusion criteria. These studies were very diverse in terms of the sample, the investigated hypotheses, the statistical analyses – and their quality.

Concerning the first research question, we found that the quality of the included studies was very heterogeneous. Eighteen studies (24 %) had a rating of ten stars or less and were hence of insufficient methodological quality, whereas 29 studies (39 %) obtained at least sufficient scores on all five criteria and were considered high-quality studies.

By using longitudinal data, we were able to bridge the knowledge gap present in previous reviews and meta-analyses of purely cross-sectional studies. In contrast to those earlier analyses, we were able to validate the causal assumptions (hypotheses a, b, and c) made by the JD-R model: job characteristics lead to employee well-being. However, we also tested the path from job demands to work engagement, which was non-significant. Since we were not able to differentiate between challenging and hindering job demands, potential relations of challenging job demands with work engagement may be obscured. Beyond that, we aimed to identify the specific relationships between job characteristics and well-being by testing different models against each other. Our model with reciprocal effects fitted the data best – both when using all studies and when using only the high-quality studies. This finding is coherent with the Conservation of Resources theory by Hobfoll and Shirom (2001) and is in line with previous research concerning the JD-R model (Reis, Hoppe, & Schröder, 2015; Xanthopoulou, Bakker, Demerouti, & Schaufeli, 2009). Job resources foster well-being, which in turn facilitates the acquirement and maintenance of job resources. In contrast, Bakker and Demerouti (2014) explain the reciprocity with self-initiated actions to actively design jobs and thus integrated job crafting into their version of the JD-R model. Hence, we consider the inclusion of job crafting as a substantial improvement to explain these reciprocal effects. Future studies may explore if job crafting can indeed explain the reversed effects of well-being on job characteristics. In addition, the reciprocal model (M4) including solely the high-quality studies had much better goodness-of-fit indices compared to M4 including all studies. The stabilities and path coefficients are comparable between both models. This suggests that research quality can alter the results and subsequent conclusions of reviews and meta-analyses and should therefore be taken into account more consistently.
By now, there is an adequate number of studies examining the causal relationships implied by the JD-R model. From our point of view, research on the JD-R model improved within the last 15 years. The essential assumptions were confirmed throughout all studies. Considering the improved results when taking into account solely the high-quality studies, we strongly advocate realizing high-quality studies with regard to planning, statistical analysis, and reporting of the results in order to reduce potential biases and to increase the validity of the empirical studies. However, since the path from job demands to work engagement is non-significant, we also advocate to consider the differentiation into challenging and hindering job demands as proposed by Crawford et al. (2010).

**Limitations**

Despite our best efforts, the present meta-analytic review, too, is not free of shortcomings.

First, the validity of reviews and meta-analyses always depends on the empirical studies available. The included studies are based on research that is self-reported regarding job characteristics and well-being. No study used objective measures of job demands or job resources. Therefore, common method bias (Podsakoff, P. M., MacKenzie, Lee, & Podsakoff, 2003) due to the utilization of self-reported instruments cannot be ruled out.

Second, in case of multiple effect sizes with regard to a single correlation, we computed a-priori composite correlations between the four constructs – job demands, job resources, burnout, and work engagement – using the Spearman-Brown formula. This is in line with the assumptions made by the JD-R model. Because of methodological limitations of metaSEM, we could not examine differential analyses concerning the impact of specific job demands and job resources. A differentiation, e.g. in challenging and hindering job demands as proposed by Crawford et al. (2010) or in social and structural job resources as proposed by Tims et al. (2013), should be addressed in future meta-analytic reviews. However, these differentiations have not been applied in most of the studies concerning the JD-R model to date.

Third, the results are limited to the essential assumptions of the JD-R model. We did not test any moderation or mediation effects because these associations cannot be tested through meta-analytic structural equation modeling. Moreover, there have been substantial improvements of the JD-R model over the course of its development such as the addition of
personal resources, job crafting, and of job performance as the final outcome. We did not take into consideration any further assumptions because of the limited availability of primary studies or well-conducted longitudinal studies that examined the role of the mentioned constructs.

Fourth, the utilized scheme from de Lange et al. (2003) to assess the quality of the studies was limited to five evaluation criteria. Since we obtained a broad range in all criteria from low-quality to high-quality studies, other criteria may be useful, too. Kmet, Lee, and Cook (2004), for example, proposed to evaluate the sample size as well as the statistical control of confounds. Another important evaluation criterion may be measurement invariance.

Fifth, our systematic literature search yielded a large number of JD-R studies. Most of them were published in peer-reviewed journals, but we also found articles in book chapters as well as dissertations. However, we focused on four databases in our literature search. Due to the fact, that the JD-R model is mostly applied in psychological research, we chose psychology-specific databases namely PsycARTICLES, PsycINFO, and PSYNDEX. Beyond that, we added the multidisciplinary database Academic Search Ultimate to cover a broader range of studies. Nevertheless, it cannot be ruled out, that we did not find every single study, which used the JD-R model as a theoretical framework. Other databases we did not select may reveal additional studies that we did not consider. Furthermore, we did not include grey publications; hence, a file-drawer effect might have biased our results.

Recommendations for Future Research

Based on this meta-analytic review, we would like to point out several issues that should be considered in future research.

Research on the JD-R model with more than three occasions. Our meta-analytic review focuses the essential assumptions made by the JD-R model. Since there are several extensions of the JD-R model, e.g. the inclusion of job crafting, personal resources, or job performance, more studies with three or more occasions are needed (Dormann et al., 2010) to examine the entire JD-R model. These studies should consider well-being (burnout and work engagement) as a mediator in the relation of job characteristics, and positive (e.g. job performance) and negative (e.g. sickness absence) as the final outcomes. To date, there is a lack of studies that focus the associations of well-being and organizational outcomes (Reis et al., 2015). This gap
needs to be closed. To explore the proposed gain and loss spirals through job crafting, four wave designs are needed, which include (longitudinal) effects of job characteristics via well-being on job performance, and of job crafting back on job characteristics.

**Research on the impact of different or optimal time intervals.** The studies included used very diverse time lags between the measures. Mostly, there are no comprehensive reasons why the authors chose a specific time interval (see results section). Therefore, it remains unclear which interval would be ideal to prove causal effects from job characteristics to well-being. To date, clear guidelines are missing. Dormann and Griffin (2015, p. 8) suggested a formula to compute optimal time lags, which consists of the stabilities of the variables as well as the assumed (reciprocal) effects. However, the authors recommend pilot studies to examine the stabilities of the variables and to adapt the optimal time lag for every study, which are probably far shorter than those frequently used in occupational health research.

**Research to clarify reversed and reciprocal relationships.** Schaufeli and Taris (2014) point out that unidirectional causal relationships between job characteristics and well-being may not solely explain the associations. Our specified models include reciprocal effects and thereby emphasize the dynamic nature of the JD-R model. Future research should thus address these dynamic relations in terms of gain and loss spirals more systematically. Referring to the motivational process, Schaufeli and Taris (2014) propose Bandura’s (1997) Social Cognitive Theory as an explanatory approach. Accordingly, job performance would boost work engagement as well as self-efficacy (as a personal resource), since job performance is associated with positive and motivation-enhancing mastery experiences. Thus, job characteristics not only predict job related well-being, but rather the complex interaction between job characteristics, job performance, and well-being. Another explanatory approach suggested by Bakker and Demerouti (2014) refers to employees’ well-being affecting the perceptions of job characteristics. Thus, burnout may lead to negative perceptions of the working environment, and work engagement would lead to a positive appraisal. Objective measures of job characteristics could be helpful to overcome potential biases.

**Research on differentiated job demands and job resources.** Several authors suggest to differentiate job characteristics into more than two general categories (Crawford et al., 2010; Luchman & González-Morales, 2013; van den Broeck et al., 2010). In our meta-analytic review,
we distinguished job demands and job resources. There is evidence that challenging job demands – such as workload, time pressure, or job responsibility – may be appraised as stressful. Nevertheless, challenging demands have potential to foster work engagement, personal growth, and future gains, too. In contrast, hindering demands – such as role conflicts, role ambiguity, or hassles – foster burnout and diminish work engagement (see Crawford et al., 2010). On the other hand, job resources can be categorized as structural and social job resources (Tims et al., 2013). Taken together, more studies need to focus on these differentiations to examine potential benefits when applying the JD-R model.

**Practical Implications**

Our results emphasize that job characteristics influence well-being via two different processes – the motivational and the health impairment process. Accordingly, fostering job resources is beneficial for employee health – beyond the mere absence of illness. Occupational interventions should thus not only focus on prevention (i.e. the health impairment process), but also address health-promoting activities (i.e. the motivational process). Bakker and Demerouti (2014) suggest four possible JD-R interventions: (1) job redesign, and (2) job crafting to optimize the working environment, as well as (3) training, and (4) strengths-based intervention to promote personal resources. Due to the fact that we did not examine personal resources, we recommend organizational interventions such as job redesign and job crafting interventions. Creating a resource-oriented working environment can take various forms: regular feedback from supervisors, more cooperative working conditions, proactive promotion of learning opportunities, various job tasks, and an increase of individual and team autonomy. As implied by the reciprocal associations of job characteristics and well-being, it should be beneficial to promote employees’ proactive behaviors and encourage them to actively change the design of their jobs.

**Conclusion**

Theoretical frameworks in occupational health psychology like the JD-R model try to investigate the relationships between job characteristics and employee well-being. The JD-R model has one crucial advantage over other occupational frameworks: it allows for both, salutogenic and pathogenic effects of job characteristics on well-being and health. We validated the essential assumptions within the JD-R model simultaneously. Despite the remaining research gaps, the JD-R model is an excellent theoretical basis to describe employee well-
being in a broad range of organizations and occupational fields. However, we advocate for high-quality research to further reduce potential biases and increase the validity of empirical studies. By means of our meta-analytic review, we hope to contribute a validated framework to the research community to determine (their) research.

References

References marked with an asterisk indicate papers that were included in the review.


https://doi.org/10.1002/smi.2483


https://doi.org/10.1080/02678370500297720

https://doi.org/10.5271/sjweh.948

https://doi.org/10.1002/smi.2460110104

https://doi.org/10.1016/j.ijnurstu.2014.05.005

https://doi.org/10.1002/job.699

https://doi.org/10.1007/s10869-013-9285-x

https://doi.org/10.1016/j.childyouth.2012.02.006


For copyright reasons the original publication is not included in this PDF. Please access the publication via the DOI provided below.

https://doi.org/10.1080/02678373.2019.1686440
Relationships between Effort-Reward Imbalance and Work Engagement in Police Officers: Taking a Salutogenic Perspective

For copyright reasons the original publication is not included in this PDF. Please access the publication via the DOI provided below.

Measuring Well-Being with the Utrecht Work Engagement Scale – Student Form: Validation of a 9- and a 3-Item Measure of Student Engagement

This version of the article may not completely replicate the final authoritative version published in European Journal of Health Psychology at https://doi.org/10.1027/2512-8442/a000027. It is not the version of record and is therefore not suitable for citation. Please do not copy or cite without the permission of the authors.

Abstract

Our study introduces a 9-item and an ultra-short 3-item version of the Utrecht Work Engagement Scale – Student Form (UWES-SF). Using data from German university students (N = 2,620), our aim was to validate the factorial structure of the UWES-9-SF and to investigate the reliability and validity of an ultra-short measure – the UWES-3-SF. Confirmatory factor analyses (CFA) confirmed the three-factor structure of the 9-item version (CFI = .95; SRMR = .033, RMSEA = .088), the internal consistencies of the subscales were satisfactory. The ultra-short version with three items (one of each dimension) also proved to be a reliable indicator of student engagement with a Cronbach’s alpha of .86. Construct validity with similar and opposite constructs (burnout, satisfaction with life, subjective health, depressive symptoms, challenging study demands) was successfully proven for both measures as well as for the subscales of the UWES-9-SF. The German short and ultra-short versions of the UWES-SF with nine and three items, respectively, are valid and reliable measures to assess student engagement.

Keywords

Work Engagement, Student Burnout, Mental Health, University Students, Student Engagement
Introduction

Work engagement – as an active state of occupational well-being (Bakker & Oerlemans, 2013) – is one of the hottest topics in occupational health psychology. Since – from a psychological perspective – studying may be considered as work (Ouweneel, Le Blanc, & Schaufeli, 2011), the concept of student engagement has also raised more and more interest. A quick search in Google Scholar (December 2018) reveals more than 350,000 hits for »student engagement«. Engaged students tend to be intrinsically motivated in learning activities, they are curious, and perform better than non-engaged students (Bakker, Sanz Vergel, & Kuntze, 2015; Lee & Shute, 2010; Salanova, Schaufeli, Martínez, & Breso Esteve, 2010). Studying should not only prepare students to become well-performing employees, but also stimulate personal growth and well-being. Hence, the concept of student engagement may be a key factor for promoting students’ health and well-being as well as for improving their academic performance.

Schaufeli, Martínez, Pinto, Salanova, and Bakker (2002, p. 465) define student engagement as “a positive, fulfilling state of mind, that is characterized by vigor, dedication, and absorption”. In contrast to student burnout, measured with the Maslach Burnout Inventory – Student Form (MBI-9-SF; Wörfel, Gusy, Lohmann, & Kleiber, 2015), student engagement is conceptualized as a positive, active state of subjective well-being (Bakker et al., 2015). Vigor, the first subscale of the Utrecht Work Engagement Scale – Student Form (UWES-SF) is characterized by “high levels of energy and mental resilience while studying, the willingness to invest effort in one’s study, and persistence even in the face of difficulties” (Schaufeli, Shimazu, Hakanen, Salanova, & de Witte, 2017) and it stands in opposition to the MBI-9-SF sub-factor of exhaustion (decrease of mental energy). Dedication, the second subscale, comprises “feelings of a sense of significance, enthusiasm, inspiration, pride, and challenge”. It is the antagonist of cynicism (second subscale of the MBI-9-SF, negative attitude towards studies). Absorption, the third dimension of the UWES-SF, describes a state of “being fully concentrated and deeply engrossed in one’s work, whereby time passes quickly and one has difficulties with detaching oneself” (Schaufeli et al., 2017). Absorption results from vigor and dedication, and is thus not the opposite of reduced personal accomplishment, the third dimension of the MBI-9-SF (lack of efficacy).
The UWES-SF was originally developed as a self-report questionnaire with 17 items (Schaufeli et al., 2002) and was shortened a few years later to a 9-item version by selecting the most characteristic item of each dimension on face validity (Schaufeli & Bakker, 2004). This item was regressed on the remaining items of the particular scale; the item with the highest β-value was added. To select a third item the sum of these two items was regressed on the remaining items, and again the item with the highest β-value was added (Schaufeli & Bakker, 2004). The items selected on face validity were: (1) “When I’m doing my work as a student, I feel bursting with energy” (vigor), (2) “I am enthusiastic about my studies” (dedication), and (3) “I am immersed in my studies” (absorption). The validity of this conceptualization of student engagement has been successfully established in several studies (e.g. Salanova et al., 2010; Schaufeli et al., 2002; Zhang, Gan, & Cham, 2007).

To date, an evaluated German version of the 9-item UWES-SF does not exist. However, in order to include student engagement as a positive form of well-being in epidemiological or university-specific surveys in Germany, we need valid and reliable measures without redundant items. Thus, the central aim of our study is to introduce and validate a German 9-item short version of the UWES-SF. Furthermore, we aim to develop and validate a German ultra-short 3-item version of the UWES-SF. Both measures, the 9-item and 3-item version shall represent reliable and valid indicators of student engagement. Therefore, we formulate the following hypotheses:

1. The three-factor solution of the UWES-9-SF (vigor, dedication, absorption) fits the data better than the one-factor solution.

2. The one-factor solution of the UWES-3-SF fits the data, and its internal consistency is good.

Moreover, we are interested in the relationship between student engagement and other measures of health and well-being as well as study-related measures. Within the Study Demands-Resources (SD-R) model (Gusy, Wörfel, & Lohmann, 2016; Mokgele & Rothmann, 2014), student engagement is predicted by study characteristics, and leads to improved or impaired health and well-being as well as academic performance over time.
Since student engagement is conceived as “the opposite positive pole of burnout” (Schaufeli et al., 2017, p. 2), we expect negative correlations between measures of student engagement and student burnout. Hence, we also hypothesize:

3. Both measures – the UWES-9-SF (and its subscales) and the UWES-3-SF – are at least moderately correlated to student burnout.

Student engagement is considered to be an active state of well-being (Bakker & Oerlemans, 2013). Therefore, negative associations should also be found for other pathogenic health indicators (physical and mental health complaints, depressive syndrome, and absenteeism). Hence, we expect negative correlations between student engagement and these pathogenic health indicators. Conversely, since subjective health and satisfaction with life are associated with pleasure, higher levels of arousal, and well-being, we expect positive correlations to these salutogenic health indicators. Therefore, we formulate the following hypotheses:

4. Both measures – the UWES-9-SF and the UWES-3-SF – are negatively related to pathogenic health indicators (i.e. physical and mental health complaints, depressive syndrome, and absenteeism).

5. Both measures – the UWES-9-SF and the UWES-3-SF – are positively related to salutogenic health indicators (i.e. subjective health, and satisfaction with life).

As stated by Crawford, LePine, and Rich (2010) and van den Broeck, de Cuyper, de Witte, and Vansteenkiste (2010), demands tend to be appraised as challenges (e.g. challenging study demands) that have the potential to stimulate personal growth – and improve student well-being. Therefore, student engagement should also be positively associated with challenging study demands:

6. Both measures – the UWES-9-SF and the UWES-3-SF – are positively related to challenging study demands.

Method

Sample and Procedures
Data was collected by means of an online questionnaire at Freie Universität Berlin in 2016. 5,558 students visited the cover page of the survey, 2,620 completed the questionnaire (47.1 %). The mean age of the respondents was 24.8 years (SD = 5.2 years), the majority was
female (71.0 %). The average study duration was 2.4 years (SD = 1.3 years). Students were enrolled in linguistics, cultural sciences, and sports (34.8 %), law, economics, and social sciences (27.2 %), mathematics, and natural sciences (26.2 %), medical study and health sciences (5.8 %), and art, and music studies (6.0 %). The measures used are presented below. The dimensionality of each instrument was confirmed. According to Nunnally and Bernstein (1994), acceptable Cronbach’s alpha coefficients should be greater than .80.

**Measures**

*The Utrecht Work Engagement Scale – Student Form (UWES-SF).* The Utrecht Work Engagement Scale – Student Form with nine items (UWES-9-SF) has three subscales (vigor, dedication, absorption) with three items each. The scale items are scored as “0 – never” and “6 – always”. For each subscale as well as for the total scale mean, (scale) scores were computed. For the ultra-short version (UWES-3-SF), one item of each dimension was selected by face validity according to Schaufeli et al. (2017). The internal consistencies of the subscales and the total scales (UWES-9-SF, UWES-3-SF) were greater than $\alpha = .84$.

To assess the convergent and criterion validity we used the following measures of health and study demands.

*Maslach Burnout Inventory – Student Form (MBI-9-SF).* This instrument adopted for students was introduced by Schaufeli et al. (2002) and translated into German, shortened and evaluated by Wörfel et al. (2015). The instrument consists of three dimensions (exhaustion, cynicism, and inefficacy) with three items each. “I feel tired when I get up in the morning and I have to face another day at the university.” is an example item for the subscale exhaustion (EX), whereas “I have become less enthusiastic about my studies.” is one for cynicism (CY). Inefficacy (IN) is measured with items such as “I believe that I don’t make meaningful contributions to the classes that I attend”. The frequency of these experiences is scored from “never” (0) to “daily” (6). The scale mean scores are computed, high scores on each dimension are indicative of burnout. The factorial validity of the abbreviated MBI-SF scales was successfully confirmed (Wörfel et al., 2015), the internal consistencies ranged from $\alpha = .76$ (IN) to $\alpha = .87$ (CY).
Subjective Health/Perceived General Health. The concept of subjective health, or perceived general health, encompasses the participants’ self-evaluation of their health. It is measured using a single item (“How is your health in general?”) that has been developed by the WHO (Bruin, Picavet, & Nossikov, 1996). The response format is “very good”, “good”, “fair”, “bad”, and “very bad”. This single-item instrument offers an efficient and internationally standardized tool for assessing overall perceived health. By keeping the wording broad, this item can potentially describe not merely physical but also social, emotional and psychological aspects of health, thus matching the WHO definition of health.

Satisfaction with Life Scale. The Satisfaction with Life Scale (SWLS) is a tool developed by Diener et al. (1985), which aims to quantify the subjects’ affective and cognitive judgement of their overall well-being. The SWLS is meant to provide a global assessment without inquiring about specific domains. The SWLS contains five items (such as “If I could live my life over, I would change almost nothing.”) to which participants can respond using a 7-point Likert scale ranging from “strongly disagree” (1) to “strongly agree” (7). The sum of the ratings comprises the overall score ranging from 5 to 35, with higher scores indicating higher levels of satisfaction. The internal consistency in our study was $\alpha = .88$.

Depressive Syndrome. In order to broadly screen for depressive symptoms, the PHQ-2 (Löwe et al., 2010) was developed as a short version of the Patient Health Questionnaire (PHQ). It does not aim to diagnose Major Depressive Disorder but rather measures the overall presence and severity of depressive symptoms regardless of whether they cross the clinical threshold. The PHQ-2 assesses two aspects of depressive affect using one item for each: “little interest or pleasure in doing things” and “feeling down, depressed, or hopeless”. The instruction ask participants to rate how often they have been bothered by these problems over the last two weeks with the response format being “not at all” (0), “several days” (1), “more than half the days” (2) and “nearly every day” (3). Calculated as the sum of the two items, the PHQ-2 total scores range from 0 to 12 with any score greater or equal 3 considered indicative of a
depressive syndrome. The correlation between the two items in our study was \( r = .64 \). However, the PHQ-2 has been validated as a unidimensional and reliable measure of depressive syndrome (Löwe et al., 2010).

**Physical and Mental Health Complaints.** To create a health complaints scale, five items from a German health behavior questionnaire (Fragebogen zur Erfassung des Gesundheitsverhaltens; Dlugosch & Krieger, 1995) were adopted, adding a sixth item for the assessment of headaches. Participants rate the frequency, ranging from “never” (1) to “every day” (7), of experienced physical symptoms: (1) cardiovascular distress; (2) gastrointestinal distress; (3) pain in limbs, shoulders, back or neck; (4) generally reduced well-being; (5) physical tension; and (6) headaches. The prevalence and frequency of these symptoms is analyzed individually, or a total score is calculated, which measures the overall occurrence of physical health complaints within a range from 6 (no symptoms) to 42 (all symptoms daily). The internal consistency in our study was \( \alpha = .82 \).

**Challenging Study Demands.** Adapted from Bakker’s Job Demands-Resources Questionnaire (2014) the challenging study demands scale aims to assess how cognitively challenging academic studies are for students. Its four items specify these demands in terms of concentration, precision, mental effort, attention, and multi-tasking. To assess multi-tasking, one item was added by the authors of this study (“Do your studies require simultaneously working on competing tasks?”). All items, for example “Do your studies require a high degree of concentration?”, are answered on a Likert scale ranging from “never” (1) to “always” (5). A high mean score indicates a higher level of (perceived) challenging study demands. The internal consistency in our study was \( \alpha = .89 \).

**Absenteeism.** The absenteeism scale is a simple one-item tool that asks students to report on how many days of the current semester they did not attend university due to illness.

**Data Analysis**

We followed the procedure of Schaufeli et al. (2017). Hence, Structural Equation Modeling (SEM) was used to test the factorial structure of the UWES-9-SF. Prior to SEM, the distribution of each UWES-9-SF item was checked for normality and multivariate outliers. All items turned out to be approximately normally distributed. Neither the skewness nor the kurtosis of any item exceeded the critical value of 1.96 (Field, 2013). The rules of thumb for model
evaluation recommended by Schermelleh-Engel, Moosbrugger, and Müller (2003) were adopted (CFI > .95; RMSEA < .08; SRMR < .1) for an acceptable fit. To test the superiority of a model, the Chi-Square Difference Test was used (Satorra & Bentler, 2010). The results of the factor analytic validation of the UWES-9-SF are presented first, followed by the psychometric properties of the (sub-) scales.

For construct validity, we investigated the UWES-SF scales intercorrelations with MBI-9-SF, SWLS, PHQ-2, the WHO item as well as with physical and mental health complaints, challenging study demands, and absenteeism. We have also used the method of Average Variance Extracted (VAE) to confirm construct validity.

Results

The Factorial Structure of the UWES-9-SF

A one-factor and a three-factor model of the UWES-9-SF items were applied to analyze the factor structure. The factor loadings of the one-factor model ranged from .63 to .82, the goodness of fit statistics for this model were not satisfactory (see Table 1; CFI = .91; SRMR = .036, RMSEA = .109). A three-factor model, representing the three subscales (vigor, dedication, absorption) shows an acceptable fit to the data (see Figure 1, see Table 1; CFI = .95; SRMR = .033, RMSEA = .088). The factor loadings range from .66 to .88. The result of the Bentler Chi Square Difference Test ($\chi^2 = 354.52; p < .01$) indicates that the three-factor model is the better solution. The dimensions are closely related, correlations between the latent variables range from .84 to .96. Hence, hypothesis 1 is confirmed: The three-factor solution fits the data better than the one-factor solution. This is in line with the results for the original scale stated by Schaufeli et al. (2002).

Table 1. Fit indices of alternative measurement models of the UWES-9-SF items (N = 2620)

<table>
<thead>
<tr>
<th>Analysis</th>
<th>$\chi^2$</th>
<th>df</th>
<th>CFI</th>
<th>TLI</th>
<th>SRMR</th>
<th>RMSEA</th>
<th>$\Delta \chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Factor</td>
<td>866.62</td>
<td>27</td>
<td>.91</td>
<td>.89</td>
<td>.036</td>
<td>.109</td>
<td></td>
</tr>
<tr>
<td>3 Factor</td>
<td>512.10</td>
<td>24</td>
<td>.95</td>
<td>.94</td>
<td>.033</td>
<td>.088</td>
<td>354.52**</td>
</tr>
</tbody>
</table>

Note. df = degrees of freedom; TLI = Tucker Lewis Index; CFI = Comparative Fit Index; SRMR = Standardized Root Mean Square Residual; RMSEA = Root Mean Square Error of Approximation; **$p < .001$. 

Internal Consistency and Correlations between UWES-9-SF and UWES-3-SF

Cronbach’s alpha for the subscales ranged from .70 (AB), .83 (DE) to .86 (VI). When all nine items were summarized to one factor the internal consistency was .86 for the UWES-9-SF. The internal consistency for the ultra-short version with 3 items – with one anchor item of each dimension – was .84. The correlations of the UWES-3-SF items ranged from .63 (VI-02-DE-02) to .64 (VI-02-AB-02). The intra-class correlation between the two measures (UWES-9-SF, UWES-3-SF) was .91 (see Table 2). In conclusion, the UWES-3-SF can be used without any significant loss of reliability, which confirms hypothesis 2.
Table 2.
Means, standard deviations, correlations and reliability estimates of the UWES-9-SF subscales (N = 2,620).

<table>
<thead>
<tr>
<th>UWES-3-SF</th>
<th>M</th>
<th>SD</th>
<th>(.86)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UWES-9-SF</td>
<td>3.13</td>
<td>1.15</td>
<td>(.86)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>UWES-9-SF Subscales</th>
<th>Vigor</th>
<th>Dedication</th>
<th>Absorption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vigor</td>
<td>2.49</td>
<td>1.22</td>
<td>(.86)</td>
</tr>
<tr>
<td>Dedication</td>
<td>3.49</td>
<td>1.37</td>
<td>.74** (.83)</td>
</tr>
<tr>
<td>Absorption</td>
<td>3.46</td>
<td>1.21</td>
<td>.74** (.70)</td>
</tr>
</tbody>
</table>

Note. Reliability estimates appear in parentheses. M = arithmetic mean; SD = standard deviation. **p < .001. 1 Intra-class correlation.

Relations with Burnout and Health Indicators

As expected, the associations between the UWES-SF (sub-) scales and burnout were moderate (-.36 < r < -.69). The strongest correlations were found with cynicism.

The associations between the UWES-SF (sub-) scales and pathogenic health indicators (physical and mental health complaints, depressive syndrome, and absenteeism) were small to moderate (-.12 < r < -.46). Correlations with both salutogenic health indicators – satisfaction with life (.30 < r < .38) and subjective health (.17 < r < .25) – were also small to moderate but in the expected positive direction (see Table 3). Since the VAE of both the UWES-3-SF and the UWES-9-SF (> .5) is larger than the correlations with pathogenic and salutogenic health indicators, criterion validity is given (see Table 3).

Relations with Challenging Study Demands

As can be seen from Table 3, student engagement correlated positively but weakly with challenging study demands (.15 < r < .28). The correlation with the UWES-9-SF was higher (r = .26) than with the UWES-3-SF (r = .15). The difference between these correlations (.11) was higher than the one reported by (Schaufeli et al., 2017) for the work context. This difference is consistent throughout the correlations with all UWES-9-SF subscales. Vigor was less associated with challenging study demands (r = .15) than dedication (r = .28) and absorption (r = .28).

In sum, the correlations of both versions and of the subscales of the UWES-9-SF were as expected, which confirms hypothesis 3, 4, 5, and 6. However, the correlations of the UWES-
3-SF with burnout, satisfaction with life, subjective health, depressive syndrome, absenteeism, health complaints, and challenging study demands were slightly lower than those of the UWES-9-SF.

Table 3. 
Correlations between MBI-9-SF scales, different health indicators, challenging study demands, and the UWES-9-SF measures.

<table>
<thead>
<tr>
<th></th>
<th>UWES-3-SF</th>
<th>UWES-9-SF</th>
<th>Vigor</th>
<th>Dedication</th>
<th>Absorption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Variance Extracted</td>
<td>.51</td>
<td>.54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exhaustion</td>
<td>-.47**</td>
<td>-.49**</td>
<td>-.56**</td>
<td>-.41**</td>
<td>-.36**</td>
</tr>
<tr>
<td>Cynism</td>
<td>-.56**</td>
<td>-.66**</td>
<td>-.56**</td>
<td>-.69**</td>
<td>-.54**</td>
</tr>
<tr>
<td>Inefficacy</td>
<td>-.42**</td>
<td>-.45**</td>
<td>-.46**</td>
<td>-.39**</td>
<td>-.38**</td>
</tr>
<tr>
<td>Satisfaction with Life Scale</td>
<td>.33**</td>
<td>.38**</td>
<td>.36**</td>
<td>.38**</td>
<td>.30**</td>
</tr>
<tr>
<td>Subjective Health</td>
<td>.20**</td>
<td>.23**</td>
<td>.25**</td>
<td>.20**</td>
<td>.17**</td>
</tr>
<tr>
<td>Depressive Syndrome</td>
<td>-.39**</td>
<td>-.45**</td>
<td>-.46**</td>
<td>-.40**</td>
<td>-.36**</td>
</tr>
<tr>
<td>Absenteeism</td>
<td>-.12**</td>
<td>-.16**</td>
<td>-.15**</td>
<td>-.14**</td>
<td>-.15**</td>
</tr>
<tr>
<td>Health Complaints</td>
<td>-.18**</td>
<td>-.20**</td>
<td>-.25**</td>
<td>-.16**</td>
<td>-.12**</td>
</tr>
<tr>
<td>Challenging Study Demands</td>
<td>.15**</td>
<td>.26**</td>
<td>.19**</td>
<td>.28**</td>
<td>.28**</td>
</tr>
</tbody>
</table>

Note: ** p < .001.

Discussion

Our study examined the psychometric structure of a short (9 items) and an ultra-short (3 items) measure of the UWES-SF in German university students. Using confirmatory factor analysis, we demonstrated that the UWES-SF can be shortened without any significant loss of information. In addition, we examined the relationships between student engagement and burnout, pathogenic as well as salutogenic health indicators, and challenging study demands.

Our data confirmed the three-factor structure of the German UWES-9-SF and the one-factor structure of the UWES-3-SF. Cronbach’s alpha for both measures was .86. The intra-class correlation between both measures was .91. Both measures showed a highly similar pattern of correlation with indicators of health and well-being, and study demands. Within the SD-R model (Gusy et al., 2016; Mokgele & Rothmann, 2014), student engagement is more strongly related to motivation and performance than to health outcomes. Hence, the correlations to subjective health as well as health complaints are relatively low. It should be noted,
however, that the correlations between the latent engagement factors of the UWES-9-SF – vigor, dedication, and absorption – are very high. Thus, researchers who are not interested in the sub-dimensions of student engagement may well use the single engagement score to assess student engagement on a general factor.

For work engagement (Schaufeli et al., 2017), the correlations are marginally lower for the ultra-short measure. We also observed this pattern: The correlations of the UWES-3-SF with burnout, satisfaction with life, subjective health, depressive syndrome, absenteeism, health complaints, and challenging study demands were slightly lower than those of the UWES-9-SF, which is a statistical consequence of shortening the measure (Schaufeli et al., 2017). Nevertheless, the ultra-short version is a reliable and valid measure to assess student engagement.

Limitations

Despite our best efforts, the present study is not free of shortcomings. First, we used a convenience sample. Since there are more female students at Freie Universität Berlin than in the average German student population, further research should verify our findings for a representative sample of German students. However, this aspect restricts the significance of our study only to a limited degree because we were mainly interested in comparing the short and the ultra-short version with each other. Second, in the current study we did not examine the UWES-3-SF independently from the UWES-9-SF. However, in a different sample of German students we examined the UWES-3-SF independently (Grützmacher, Gusy, Lesener, Sudheimer, & Willige, 2018). In that sample, Cronbach’s alpha was also high (α = .85) and thus supported the reliability of our ultra-short measure. Nevertheless, our results should be verified in other samples. Third, as Schaufeli et al. (2017) already mentioned for their ultra-short measure of work engagement, the three-dimensional structure of the original UWES-SF has been sacrificed. Using the UWES-3-SF, it is no longer possible to detect specific effects concerning vigor, dedication, and absorption. However, researchers who are interested in investigating such specific effects still have the UWES-9-SF at their disposal as a reliable and valid measure to capture all three sub-dimensions of student engagement.
Conclusion

Both measures – the UWES-9-SF and the UWES-3-SF – appear to be reliable and valid indicators of student engagement that can be used as an alternative to Schaufeli’s original version, the UWES-SF. Due to its brevity, the UWES-3-SF is most appropriate to capture student engagement in national and epidemiological surveys. For those who are interested in specific effects of student engagement, the UWES-9-SF seems to be the best choice since it is a reliable yet economic measure of vigor, dedication, and absorption.

References


General Discussion

The major aims of my research were to get detailed insights into job-related antecedents and outcomes of work engagement and its applicability in specific occupational and non-occupational contexts. Since I conducted the studies in collaboration with my colleagues, I will use the terms we and our when presenting results and implications.

Regarding the first research question (see introduction), we were able to validate the essential assumptions of the JD-R framework, which was the first theoretical model that introduced work engagement conceptually: job resources lead to work engagement even in the long run, whereas job demands do not have a substantial impact on work engagement. The JD-R framework has the crucial advantage over other occupational frameworks that it allows for both, salutogenic and pathogenic effects of job characteristics on occupational well-being. Based on our research, we concluded that the JD-R framework is an excellent theoretical basis to examine work engagement in a broad range of organizations and occupational fields.

However, we still wanted to investigate which job resources are most important for work engagement and why some job resources might be more significant than others for facilitating work engagement. Hence, we tackled the second and third research question: do job resources differentially predict work engagement, and what job resources would be the key drivers of work engagement in various occupational settings. As proposed by Nielsen et al. (2017), we differentiated job resources in organizational-level, group-level, and leader-level resources and concluded that organizational-level resources (e.g. autonomy) are the most important predictors for work engagement. This offers clarity, especially concerning where interventions to enhance occupational well-being should start. Although interventions at any of the three levels are promising, organizations are well-advised to predominantly strengthen resources at the organizational level to increase employees' work engagement.

With this in mind, we aimed to broaden the theoretical and contextual application of work engagement. To tackle the fourth research question, we integrated work engagement as a positive, salutogenic outcome into the effort-reward imbalance (ERI) model. Since the ERI model was originally designed to identify health complaints and their job-related antecedents, the integration of work engagement broadens the applicability of this model. We found that
an imbalance between effort and reward is associated with reduced work engagement. However, especially the reward dimensions esteem and security reward foster work engagement within the police context. Hence, we gained insights into which specific factors within the ERI model are most important for work engagement.

Furthermore – since work engagement is associated with various positive health- and performance-related outcomes – we applied the concept and its operationalization to the context of university students. To answer the fifth research question, we examined the psychometric structure of a short (9 items) and an ultra-short (3 items) measure of student engagement with the Utrecht Work Engagement Scale – Student Form (UWES-SF), and its relations to other pathogenic and salutogenic health indicators. We confirmed the three-factor structure of the German UWES-9-SF and the one-factor structure of the UWES-3-SF. Both measures appear to be reliable and valid indicators of student engagement that can be used as an alternative to Schaufeli’s original version, the UWES-SF (Schaufeli, Martinez, Marques Pinto, Salanova, & Bakker, 2002). Due to its brevity, the UWES-3-SF is most appropriate to capture student engagement in national and epidemiological surveys.

Our research sharpens the understanding of work engagement as a theoretical concept and its applicability in various occupational and non-occupational contexts. Based on these results, I firmly believe that work engagement – defined as a positive, fulfilling state – is one of the most beneficial states of well-being for both employees and students.

Limitations

However, despite our best efforts, several limitations of our research need to be addressed.

First, although we have highlighted the essential assumptions made by the JD-R framework and the most important predictors of work engagement longitudinally, we did not examine either the JD-R framework or the motivational process as a whole. The final outcomes of the JD-R framework are various positive health- and performance-related aspects, which we were not able to consider in our research. We did not include these assumptions because of the limited availability of well-conducted longitudinal studies that examined the associations to health- and performance-related outcomes. In addition, we were unable to realize, by ourselves, a well-conducted longitudinal study focusing on the impact of work engagement on
positive outcomes. However, future research certainly needs to address the longitudinal ef-
fects of work engagement on health- and performance-related outcomes such as in-role, or
extra-role performance.

Second, there is also evidence that not only job characteristics influence work engage-
ment, but work engagement may also be an antecedent of job resources over time (Taris &
Schaufeli, 2016). This research suggests “that high-resource workers tend to become more
engaged over time and that engaged workers tend to collect more resources in their job, which
in turn leads to even higher levels of engagement” (Taris & Schaufeli, 2016, p. 173) – the so-
called gain spirals. However, testing these reciprocal assumptions properly would have re-
quired at least 3-wave studies (Dormann, Zapf, & Perels, 2010) that include job resources and
work engagement at three or more occasions. Unfortunately, we did not find enough 3-wave
studies to test these assumptions meta-analytically, and we were also not able to conduct our
own longitudinal study that examines these gain spirals. However, I strongly agree with Taris
and Schaufeli (2016), that we need more research on this intriguing issue.

Third, several authors suggest that specific job demands may also have a specific im-
 pact on work engagement. Challenging job demands may have the potential to foster work
engagement, whereas hindering job demands may diminish work engagement over time (see
Crawford, LePine, & Rich, 2010). Although we included an imbalance between effort and re-
 ward in one of our studies, we did not take specific challenging or hindering job demands and
their impact on work engagement into account. This distinction could help to identify other
important drivers of work engagement that go beyond the impact of job resources. Since we
were not able to differentiate job demands’ impact in our research, further studies on work
engagement’s antecedents should distinguish between different types of job demands.

Fourth, although we have used the innovative methodological approach of meta-analy-
lytic structural equation modelling in two studies, we still relied exclusively on self-reported
data. We were not able to include any objective measures of job characteristics or work en-
gagement in our analyses. Therefore, common method bias (Podsakoff, MacKenzie, Lee, &
Podsakoff, 2003) due to the utilization of self-reported instruments cannot be ruled out. For
work engagement, objective measures currently do not even exist. However, objective
measures of job characteristics or work engagement should be integrated in future research.
Fifth, our research refers only to work engagement as defined by Schaufeli, Salanova, González-Romá, and Bakker (2002). As I have pointed out in the introduction, there are several other definitions of (work) engagement either as a state, a trait, or a behavior. Since we handled work engagement as a positive, work-related state of mind, we were not able to draw any conclusions on the dispositional or behavioral aspects of work engagement. However, there is no clear conceptualization and operationalization of work engagement as a trait or behavior. Future research could further integrate these dispositional and behavioral aspects to examine spillover effects from work engagement as a psychological state to personality and behavior.

Sixth, although we mainly focused on longitudinal data, we were not able to capture the issue of different time lags in our studies. In most studies, there are no comprehensive reasons for a specific time interval. Therefore, it remains unclear which interval would be ideal to prove causal effects from job characteristics to work engagement. To date, clear guidelines are missing. Our models do not take this aspect into account. Through the inclusion of all studies into one model, the impact of different time lags was averaged out. However, future research should consider the issue of different time lags in order to propose optimal time lags in occupational health psychology (OHP).

Recommendations for Future Research

As mentioned above, we were not able to address all aspects of work engagement within the JD-R framework and its motivational process in our research. Several things still need to be done both theoretically and methodologically.

*Theoretical considerations.* Although we have validated the motivational process within the JD-R framework – that job resources lead to work engagement over time – there are several more specific propositions that we were not able to address.

Future research should focus on the impact of challenging and hindering demands on work engagement as proposed by Crawford et al. (2010). At the moment, there is no clear theoretical distinction between these two categories (Taris & Schaufeli, 2016). However, as Crawford et al. (2010) and van den Broeck, de Cuyper, de Witte, and Vansteenkiste (2010) have shown, challenging demands seem to have a positive impact on work engagement, whereas hindering job demands seem to have a negative impact. This aspect needs further clarification, in the best case with clear assumptions within the JD-R framework that distinguishes between
different types of job demands. Furthermore - as we have shown with our research - job resources are differentially associated with work engagement over time. Although all job resources seem to have a positive impact on work engagement, organizational-level resources contribute much more than group-level, and leader-level resources. This could also lead to more precise hypotheses and thus to a refinement of JD-R's motivational process and its assumptions. Still, although we know that organizational-level resources are the most important predictors of work engagement, more research is needed on specific job-related antecedents, preferably across different types of jobs (Demerouti & Bakker, 2011).

We need more research on the longitudinal effects of work engagement on health- and performance-related outcomes. Strictly speaking, work engagement is not the final outcome within JD-R's motivational process, but the mediator between job- and personal-related antecedents and positive outcomes (Schaufeli, 2017). However, little research has addressed this assumption so far. I strongly advocate for more research on work engagement's outcomes, be it on physical health or job performance. This would also be a substantial improvement of the empirical evidence on the JD-R framework.

We also need more research on the issue of gain cycles within the motivational process. Although there is some evidence that work engagement also influences job resources (e.g. Llorens, Schaufeli, Bakker, & Salanova, 2007; Schaufeli, Bakker, & van Rhenen, 2009; Xanthopoulou, Bakker, Heuven, Demerouti, & Schaufeli, 2008), we need some more insight in the dynamic relationships between job resources and work engagement. More recent versions of the JD-R framework use job crafting as an explanation for gain cycles (Bakker & Demerouti, 2017). Job crafting can be defined as proactive behavior to adjust structural job characteristics (form, scope, or number of tasks), and/or as cognitive changes in the appraisal of one's job. Bakker and Demerouti (2017) assume, that employees who are engaged are likely to use job crafting behaviors, which lead to higher levels of job resources and in turn to even higher levels of work engagement. However, it is still rather unclear if these underlying complex mechanisms are evident in longitudinal research.

Currently, the JD-R framework is the only framework which contains work engagement as one of its main outcomes. In contrast, the ERI model has so far largely neglected salutogenic outcomes. The pathogenic perspective of the ERI model focusses – by definition – only
on the imbalance between effort and reward towards effort, resulting in impaired health and well-being. The model does not provide any assumptions with regard to the impact of a balanced ratio between effort and reward or even a ratio where effort is outbalanced by reward for positive health well-being. However – as we have shown – it is possible to integrate positive outcomes like work engagement into the ERI model, even though its conceptualization of job resources is limited to reward only. Future research on the ERI model should further integrate work engagement as a positive outcome and explore how an effort-reward balance or effort-reward imbalance towards reward affects employees’ work engagement.

Methodological considerations. To address the above-mentioned aspects, future research has to consider several methodological aspects, too.

I strongly argue for more well-conducted longitudinal studies on work engagement with more than two measurements. To explore the more complex assumptions of the JD-R framework (e.g. the entire motivational process with job performance as the final outcome), more studies with at least three occasions are needed. These studies should include the effects of job resources via work engagement on job performance, and back on job resources as recently proposed by the JD-R framework (Bakker & Demerouti, 2017). It would also be possible to clarify reversed and reciprocal relationships between work engagement and possible predictors and outcomes, since unidirectional causal relationships may be too simple to explain these associations (Schaufeli & Taris, 2014). Longitudinal research with three or more occasions would also address the call for more rigorous tests of causality (Bakker & Demerouti, 2017) concerning the JD-R framework.

Most studies on work engagement and the motivational process have used only self-reported measures for both job resources and work engagement (Bakker & Demerouti, 2017). In these cases, the same person provides all information, so that statistical relationships between the variables may be overestimated as a result of common method bias (Podsakoff et al., 2003). Several studies have used objective measures to capture job demands (e.g. Qin, Hom, Xu, & Ju, 2014; Wingo, Halvorsen, Beckman, Johnson, & Reed, 2016). However, almost no study has introduced objective measures for job resources and work engagement (Bakker & Demerouti, 2017). The type of measure (e.g. subjective vs. objective) of different job characteristics has a substantial impact on the reported correlations with stress and health indicators.
(Zapf, 1989). Thus, I strongly recommend a combination of different types of measures since this could reduce the risk of common method bias.

In most longitudinal studies the authors do not propose any comprehensive reason why they chose a specific time interval. We have seen this in both meta-analytic reviews. Thus, it remains unclear which interval would be optimal to prove causal effects from job characteristics to work engagement. There are currently no clear guidelines. However, Dormann and Griffin (2015, p. 8) introduced a formula\(^2\) to identify optimal time lags in OHP, which contains the stabilities of the variables as well as the assumed (reciprocal) effects. They propose pilot studies to examine the stabilities of the variables and to adapt the optimal time lag for every study. Unfortunately, I was not able to address these concerns in my research, but I strongly recommend a more comprehensible reasoning for choosing specific time lags, or using the concept of optimal time lags when conducting research in OHP.

Another aspect concerns the statistical procedure of meta-analytic structural equation modeling in metaSEM (Cheung, 2014). To date, metaSEM does not allow the integration of moderators or control variables. However, the assumptions of the JD-R framework are more complex than those associations that we were able to examine. It would be a substantial improvement of metaSEM if it would allow the testing of more complex relationships and I strongly plead for more options within the metaSEM package.

**Implications**

As we have seen, work engagement is a desirable state of occupational well-being. Considering its connections with positive organizational outcomes (Halbesleben, 2010), organizations should be interested in enhancing work engagement. But how can the JD-R framework be used to improve job characteristics and work engagement?

**The JD-R monitor.** The JD-R framework can be practically applied in organizations to structure the process of assessing and tackling employees’ health and well-being. Here, the role of the JD-R framework goes beyond analyzing the most relevant predictors of burnout

\[\omega_{opt} = \frac{\ln(\frac{\ln(\frac{d}{r})}{\frac{1}{2}(d^2 + i^2 + c^2 + r^2)})}{\ln(\frac{\ln(\frac{d}{r})}{\frac{1}{2}(d^2 + i^2 + c^2 + r^2)})},\]

with \(d\) and \(i\) as stabilities, and \(c\) and \(r\) as reciprocal effects between job characteristics and well-being.
and work engagement by playing a crucial role in the communication with all relevant stakeholders within an organization, e.g. employees, management, HR-officers (Schaufeli, 2017). In doing so, the framework constitutes a shared mindset and a common language for how all stakeholders consider the organization and their personal and professional situation (Schaufeli, 2017). Schaufeli and Taris (2014) propose the JD-R monitor as a JD-R-based online tool that consists of several steps: (1) determine aims and the project team; (2) customize the survey; (3) internal communication campaign; (4) individual feedback; (5) analyses and reporting; (6) survey feedback; (7) interventions; (8) evaluation. In the first step, the organization defines the problem, such as: How can we increase the employees' levels of work engagement. In this step, it is also important to put together the project team. In the second step, the project team tailors the online assessment to the needs of the organization and to the aims of the project. Key persons should identify the most relevant job demands, job and personal resources as well as health- and performance-related outcomes. In the third step, an internal communication campaign should stimulate employees to answer the assessment and to be proactively involved in the project. In the fourth step, the employees receive an email with an automatically generated and personalized feedback. This should contain the employee's scores compared to a benchmark and a more detailed explanation of the score's meaning. In the fifth step, a general report is given for the entire organization. This should contain the average scores for the whole organization compared to a benchmark. In the sixth step, this report is discussed throughout the organization at various levels including team or departmental levels. This critical discussion should build commitment and trust for following interventions. In the seventh step, based on the results of the JD-R monitor, interventions are discussed. These can be implemented by the employees themselves, but also at team- and organizational-level. I will discuss the issue of interventions in the next paragraph in more detail. In the last step – after the intervention – the organization should go through the seven steps again to check whether the process has been effective. The JD-R monitor could be utilized in another cycle to investigate if the aims were achieved, for instance, if work engagement has increased considerably. Schaufeli and Taris (2014) conclude that it would be ideal if the JD-R monitor could be integrated in the annual HR cycle so that job characteristics can be assessed and occupational well-being and evidence-based HR policy decision can be made.
**Work engagement interventions.** Bakker and Demerouti (2014) suggest four levels of interventions based on the JD-R framework: (1) job redesign, (2) job crafting, (3) training, and (4) strengths-based interventions. Job redesign interventions aim to change job demands and job resources at the organizational level, and usually represent top-down processes (Bakker & Demerouti, 2014). Job crafting interventions also aim to change job demands and job resources, but at the individual level, initiated by the individual employee and therefore represent bottom-up processes (Bakker & Demerouti, 2014). Trainings and development of employees aim to convey new skills, technical knowledge, and problem-solving abilities to the employees, and also take place at the organizational level (Bakker & Demerouti, 2014). However, strengths-based interventions aim to enable employees to use their strengths at work and should lead to self-efficacy, and also refer to the individual level (Bakker & Demerouti, 2014). Whereas job redesign and job crafting focus the improvement of job characteristics, training and strengths-based interventions focus the promotion of personal resources.

Ten years ago, Kang, Staniford, Dollard, and Kompier (2008) have remarked the paucity of intervention studies in OHP. This gap between research and practice still exists. To date, there are only two systematic reviews that focus interventions specifically on work engagement, both from the same authors (Knight, Patterson, & Dawson, 2017, 2019). Knight et al. (2019) reviewed the literature on work engagement interventions and found 40 studies. 20 studies had a significant positive effects, two studies had negative effects, whereas the other 18 studies did not find any effect on work engagement (Knight et al., 2019). The authors concluded, that the most effective interventions addressed either increasing job resources, personal resources, or other indicators of well-being. Building job resources and personal resources through job crafting was particularly effective, just as health promotion interventions, which utilized, for example, mindfulness-based stress reduction programs or cognitive behavioral therapy strategies (Knight et al., 2019). There were several more aspects that had a positive impact on the effectivity of the interventions: (1) Interventions that contained a group and an individual component were successful more often than interventions that included only a group or an individual component; (2) employee participation was a substantial component of successful interventions; and (3) bottom-up interventions were more successful than top-down interventions (Knight et al., 2019). However, there were also some factors that
prevented positive effects: (1) poor management support; (2) potential cross-over effects between intervention and control groups; (3) organizational restructuring; (4) and concurrent projects preventing causal conclusions (Knight et al., 2019). Hence, interventions on work engagement should follow these guidelines to ensure their effectivity. However, based on our results, I strongly recommend strengthening job resources at the organizational-level.

Conclusions

Since modern organizations expect their employees to be proactive, responsible, committed, creative, and energetic, they should also be interested in their employees’ health and well-being. Work engagement can truly make a difference for organizations as a competitive advantage (Leiter & Bakker, 2010), and can lead to better performance in the long run. My research has shown the broad applicability of the JD-R framework and more specifically the construct of work engagement in various occupational and non-occupational contexts and also the key drivers that foster employees’ work engagement. The growing number of published articles on the topic reflects the positive psychology twist in OHP, which should continue further. I am truly convinced that these results enable organizations to improve job characteristics and thereby stimulate employees’ health and well-being.

References


https://doi.org/10.4102/sajip.v37i2.974


Statement of Independence (Erklärung zur Dissertation)

Erklärung gemäß § 7 Abs. 3 Satz 4 der Promotionsordnung über den Eigenanteil an den veröf
denachtigen oder zur Veröffentlichung vorgesehenen eingereichten wissenschaftlichen
Schriften im Rahmen meiner publikationsbasierten Arbeit.

I. Name, Vorname: Lesener, Tino
   Institut: Public Health: Prävention und psychosoziale Gesundheitsforschung
   Promotionsfach: Psychologie
   Titel: Master of Science (M.Sc.)

II. Nummerierte Aufstellung der eingereichten Schriften:

   https://doi.org/10.1080/02678373.2018.1529065

2. Lesener, T., Gusy, B., Jochmann, A., & Wolter, C. The Drivers of Work Engage-
   online publication. https://doi.org/10.1080/02678373.2019.1686440

3. Wolter, C., Santa Maria, A., Georg, S., Lesener, T., Gusy, B., Kleiber, D., &
   Renneberg, B. (2019). Relationships between Effort-Reward Imbalance and Work En-
   gagement in Police Officers: Taking a Salutogenic Perspective. Journal of Public

   Work Engagement Scale – Student Form: Validation of a 9- and a 3-Item Measure of
   https://doi.org/10.1027/2512-8442/a000027
III. Darlegung des eigenen Anteils an diesen Schriften:

zu II. 1.: Konzeption (überwiegend), Literaturrecherche (überwiegend), Methodenentwicklung (überwiegend), Versuchsdesign (überwiegend), Datenerhebung (überwiegend), Datenauswertung (vollständig), Ergebnisdiskussion (überwiegend), Erstellen des Manuskriptes (vollständig), Programmierung (vollständig), Beweisführung (überwiegend)

zu II. 2.: Konzeption (überwiegend), Literaturrecherche (überwiegend), Methodenentwicklung (überwiegend), Versuchsdesign (überwiegend), Datenerhebung (überwiegend), Datenauswertung (vollständig), Ergebnisdiskussion (überwiegend), Erstellen des Manuskriptes (vollständig), Programmierung (vollständig), Beweisführung (überwiegend)


zu II. 4.: Konzeption (in Teilen), Literaturrecherche (mehrheitlich), Methodenentwicklung (mehrheitlich), Versuchsdesign (in Teilen), Datenerhebung (überwiegend), Datenauswertung (in Teilen), Ergebnisdiskussion (in Teilen), Erstellen des Manuskriptes (überwiegend), Programmierung (in Teilen), Beweisführung (mehrheitlich)

IV. Die Namen und Anschriften nebst E-Mail oder Fax der jeweiligen Mitautor_innen:

zu II. 1.: PD Dr. Dr. Burkhard Gusy; AB Public Health: Prävention und psychosoziale Gesundheitsforschung; Freie Universität Berlin; Habelschwerdter Allee 45; 14195 Berlin; E-Mail: burkhard.gusy@fu-berlin.de

Christine Wolter; AB Public Health: Prävention und psychosoziale Gesundheitsforschung; Freie Universität Berlin; Habelschwerdter Allee 45; 14195 Berlin; E-Mail: christine.wolter@fu-berlin.de

zu II. 2.: Anna Jochmann; AB Public Health: Prävention und psychosoziale Gesundheitsforschung; Freie Universität Berlin; Habelschwerdter Allee 45; 14195 Berlin; E-Mail: anna.jochmann@fu-berlin.de
zu II. 3.: Andreas Santa Maria; AB Klinische Psychologie und Psychotherapie; Freie Universität Berlin; Habelschwerdter Allee 45; 14195 Berlin; E-Mail: a.santamaria@fu-berlin.de

Stephanie Georg; AB Klinische Psychologie und Psychotherapie; Freie Universität Berlin; Habelschwerdter Allee 45; 14195 Berlin; E-Mail: st.georg@fu-berlin.de

Prof. Dr. Dieter Kleiber; AB Public Health: Prävention und psychosoziale Gesundheitsforschung; Freie Universität Berlin; Habelschwerdter Allee 45; 14195 Berlin; E-Mail: dieter.kleiber@fu-berlin.de

Prof. Dr. Babette Renneberg; AB Klinische Psychologie und Psychotherapie; Freie Universität Berlin; Habelschwerdter Allee 45; 14195 Berlin; E-Mail: b.renneberg@fu-berlin.de

zu II. 4.: siehe oben

Datum, Unterschrift des Doktoranden
Die Angaben zu III. müssen von den Mitautor_innen schriftlich bestätigt werden.

Ich bestätige die von Tino Lesener unter III. abgegebene Erklärung:

<table>
<thead>
<tr>
<th>Name</th>
<th>Unterschrift</th>
</tr>
</thead>
<tbody>
<tr>
<td>PD Dr. Dr. Burkhard Gusy</td>
<td></td>
</tr>
<tr>
<td>Christine Wolter</td>
<td></td>
</tr>
<tr>
<td>Anna Jochmann</td>
<td></td>
</tr>
<tr>
<td>Andreas Santa Maria</td>
<td></td>
</tr>
<tr>
<td>Stephanie Georg</td>
<td></td>
</tr>
<tr>
<td>Prof. Dr. Dieter Kleiber</td>
<td></td>
</tr>
<tr>
<td>Prof. Dr. Babette Renneberg</td>
<td></td>
</tr>
</tbody>
</table>