UNIVERSITÄT
DES
SAARLANDES

# Response Behavior in Work Stress Surveys:

# An Investigation Using Qualitative and Quantitative Research Methods

Dissertation

zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften der Fakultät HW

Bereich Empirische Humanwissenschaften

der Universität des Saarlandes


vorgelegt von

Berit Greulich

aus Heidelberg


Saarbrücken, 2024

| | |
|---|---|
| Dekan: | Univ.-Prof. Dr. Peter Loos |
| Berichterstatter:in 1: | Univ.-Prof. Dr. Cornelius J. König |
| Berichterstatter:in 2: | PD Dr. Dorota Reis |
| Tag der Disputation: | 20.11.2023 |

**Zusammenfassung**

Psychische Belastungen und damit verbundene Erkrankungen aufgrund belastender Arbeitsbedingungen führen zu Leiden bei Arbeitnehmer:innen und hohen Kosten für Unternehmen und Gesellschaft. Die frühzeitige Erkennung belastender Arbeitsbedingungen ist daher entscheidend, um präventive Maßnahmen ergreifen zu können. Fragebögen zur Erfassung von Stress am Arbeitsplatz bieten eine effektive Möglichkeit, sind aber auch mit Herausforderungen verbunden. Meine Dissertation zielt darauf ab, das Antwortverhalten von Arbeitnehmer:innen und Vorgesetzten in Arbeitsstressfragebögen besser zu verstehen, um mögliche Herausforderungen anzugehen und die Zuverlässigkeit und Validität der Datenerhebung zu verbessern. Zu diesem Zweck führte ich zunächst Interviews mit Arbeitnehmer:innen und Vorgesetzten aus verschiedenen Branchen durch, um ihr Vorgehen bei der Einschätzung von Stressoren und Ressourcen zu verstehen. Darauf aufbauend untersuchte ich zwei verschiedene Formen der Kontextualisierung. Erstens habe ich generische Arbeitsstressitems mit stark kontextualisierten Items verglichen, die speziell auf die Arbeitsbedingungen von medizinischem Personal zugeschnitten sind. Zweitens untersuchte ich den Einfluss eines vorgegebenen sozialen Vergleichs auf die Beantwortung von Arbeitsstressitems. Beide Formen der Kontextualisierung beeinflussen das Antwortverhalten und liefern Hinweise auf mögliche Verbesserungen der Messqualität. Darüber hinaus untersuchte ich das in den Interviews beschriebene Phänomen des "defensive biasing", bei dem Arbeitnehmer:innen Stressoren absichtlich unterschätzen und Ressourcen überschätzen, um sich vor negativen Konsequenzen seitens der Vorgesetzten zu schützen. Eine eigens entwickelte Skala ermöglicht die quantitative Erfassung dieses Phänomens, das mit einer als gering wahrgenommenen Anonymität und Neurotizismus korreliert. Zusammenfassend zeigt

meine Dissertation, dass dem Antwortverhalten in Arbeitsstressfragebögen komplexe und spezifische Prozesse zugrunde liegen, die die Bewertung beeinflussen. Schlussendlich sprechen die Ergebnisse dieser Dissertation dafür die Besonderheiten im Antwortverhalten von Arbeitsstressfragebögen weiter zu untersuchen.

## General Abstract

Mental stress and associated illnesses resulting from challenging working conditions inflict suffering upon employees and impose significant costs on both organizations and society. Timely recognition of burdensome work conditions is crucial in order to implement preventive measures. While questionnaires assessing workplace stress offer an effective means of measurement, they also present challenges. The aim of my dissertation is to gain a deeper understanding of employees' and supervisors' response behavior in work stress surveys, with the goal of addressing potential challenges and enhancing the reliability and validity of data collection. To achieve this, I conducted interviews with employees and supervisors from diverse industries to comprehend their approach in assessing stressors and resources. Building on these findings, I investigated two distinct forms of contextualization. First, I compared generic work stress items with highly contextualized items tailored specifically to the working conditions of medical personnel. Second, I examined the influence of a predefined social comparison on the responses to work stress items. Both forms of contextualization impact response behavior and provide insights into potential enhancements of measurement quality. Furthermore, I explored the phenomenon of "defensive biasing" described in the interviews, in which employees intentionally trivialize stressors and valorize resources as a protective mechanism against negative consequences from supervisors. A specifically developed scale allows for the quantitative assessment of this phenomenon, which correlates with perceived low anonymity and neuroticism. In summary, my dissertation highlights the existence of complex and specific processes underlying response behavior in work stress questionnaires, which impact the evaluation and warrant further investigation.

## Index of Publications

This publication-oriented dissertation[1] (German: publikationsorientierte Dissertation) is mainly based on two publications and two manuscripts submitted for publication. The author of this dissertation is the first author of the publications and all the manuscripts which are inserted in this dissertation in their most recent version (with slight changes due to formatting).

Study 1: Greulich, B., Debus, M. E., Kleinmann, M., & König, C. J. (2021). Response behavior in work stress surveys: A qualitative study on motivational and cognitive processes in self- and other-reports. *European Journal of Work and Organizational Psychology*, *30*(1), 40–57. https://doi.org/10.1080/1359432X.2020.1812580

Study 2: Greulich, B., König, C. J., & Meixensberger, C. V. (2023). *Contextualization of work-related stress items and its effects on criterion validity and reliability.* Manuscript submitted for publication to Journal of Personnel Psychology.

Study 3: Greulich, B., König, C. J., & Fischer, N. (in press). How relative is stress? The influence of social comparison when responding to work stress surveys. *Journal of Personnel Psychology*.

Study 4: Greulich, B., König, C. J., & Mohr, R. (2023). *When employees understate their stress: Defensive biasing in work stress surveys.* Manuscript submitted for publication to International Journal of Workplace Health Management.

[1] The layout of this dissertation is based on that of Langer (2018).

## Table of Contents

TABLE OF CONTENTS

TABLE OF CONTENTS

## Index of Figures

**Index of Tables**

# 1 General Introduction

Stress has emerged as one of the most significant health risks in the contemporary working environment (O'Connor et al., 2021), rendering it a central focus within the field of occupational health psychology (e.g., Quick & Tetrick, 2011). Various work-related demands, such as the digitization of the workplace, continuous multitasking, time pressure, monotony, and the blurring of professional and personal life boundaries (e.g., through remote work arrangements), pose challenges for numerous employees (Rigó et al., 2021). When these demands are not appropriately balanced and persistently exceed an individual's coping capacity, they can become stressors (Demerouti et al., 2001). Consequently, employees experience strain, which can result in diminished performance, motivation, and, over time, serious health issues such as cardiovascular diseases (e.g., Gerhardt et al., 2021; Taibi et al., 2021) and mental illnesses (e.g., Hu et al., 2017; Schaufeli et al., 2009). Furthermore, work stress imposes substantial financial burdens on organizations and society as a whole, including costs associated with long-term sick leave, disability benefits, reduced productivity, and healthcare expenses (Hassard et al., 2018). Hence, it is imperative for organizations to recognize potential work stress and it's underlying causes in order to establish a healthy work environment that safeguards employee well-being. In Germany, for instance, legislation has acknowledged this necessity, requiring organizations not only to protect employees from accidents but also to address psychosocial work stress (§5 Absatz 3 Satz 6 ArbschG). Understanding the impact of work stress and its detrimental consequences is crucial for organizations to implement effective measures and interventions that promote employee health and mitigate the negative effects of stress in the workplace. By doing so, organizations can

foster employee well-being, enhance performance, and contribute to a healthier and more sustainable work environment (van Wingerden et al., 2017).

In order to fulfill the requirement of providing employees with health-preserving working conditions, organizations must possess a comprehensive understanding of the sources and causes of work stress, encompassing both the identification of potential stressors and the recognition of resources that can help mitigate them (Erazo-Chamorro et al., 2022). Therefore, it is crucial to employ valid and reliable methods of data collection to accurately measure work stress. The obtained results should offer a realistic representation of as many employees within an organization as possible, enabling the development and implementation of targeted and effective interventions to foster health-preserving conditions. In Germany, the assessment of work stress has traditionally been approached through job analysis, exemplified by instruments like the Instrument for Stress-Oriented Task Analysis (ISTA) developed by Semmer et al. (1999). Job analysis typically involves examining and documenting the tasks, responsibilities, skills, and tools required for specific job roles (Morgeson et al., 2019), often through self-report measures. Self-reports have the advantage of being widely applicable, easy to administer, and directly reflecting employees' perspectives (Spector & Eatough, 2013). However, self-report measures are also susceptible to various biases stemming from respondents' cognitive and motivational processes during the response process (e.g., Morgeson & Campion, 1997, 2012). Measuring work stress presents an additional challenge due to its highly individual and subjective nature, as evidenced by stress research highlighting the significance of the evaluation of stressors in shaping the stress experience (Lazarus & Folkman, 1987; Perrewé & Zellars, 1999). Therefore, it can be assumed that assessing work stress is significantly more complex for respondents than simply describing

tasks or identifying the need for tools, as typically done in traditional job analyses. Consequently, it can be expected that distinct cognitive and motivational processes are at play, exerting considerable influence on response behavior. To design instruments that yield reliable data for measuring work stress, it is crucial to gain a deep understanding of how employees engage in the assessment process, encompassing all the cognitive and motivational processes involved in this highly individual and subjective evaluation. By unraveling these processes, measurement tools can be developed that effectively capture employees' perceptions of work stress, ultimately leading to more accurate and reliable data.

Therefore, the overall goal of my dissertation is to deepen our understanding of response behavior in work stress surveys, with a particular focus on cognitive and motivational processes and other potential influences, in order to improve the quality of such surveys. To accomplish this aim, Study 1 employed a qualitative approach to develop a conceptual model of response behavior, encompassing cognitive and motivational processes, as well as other influences within work stress surveys. Additionally, the study explored similarities and differences in response patterns among self, supervisor, and coworker assessments of work stress survey items. Building upon the qualitative findings from the first study, the subsequent three studies aimed to quantitatively investigate and map the identified processes. Study 2 examined the impact of high-level contextualization on response behavior among medical assistants, drawing inspiration from the positive effects observed in personality questionnaires (Holtrop et al., 2014). Study 3 explored the influence of social comparison processes, specifically the tagged form of contextualization known as the reference group effect (Credé et al., 2010; Heine et al., 2008), on response behavior within work stress surveys. In the interviews conducted during the first

study, participants also mentioned motivational processes that arise when answering work stress surveys, which exhibit a high degree of specificity and warrant more comprehensive examination. Therefore, Study 4 sought to investigate the phenomenon of trivializing stressors and valorizing resources due to the fear of negative consequences associated with honest responses, referred to as defensive biasing.

By addressing these research objectives, my dissertation aims to advance our understanding of the complexities surrounding response behavior in work stress surveys. It is expected that the findings will contribute to the refinement and improvement of survey instruments, ultimately enhancing their reliability and validity in capturing the experiences and perceptions of work stress among employees.

# 2 General Theoretical Background

In the following text, I will first provide an overview of the definition of stress and its conceptualization within the field of occupational health psychology. I will then highlight the importance of measuring work stress and the need for organizations to be aware of its impact. Finally, I will discuss the potential challenges and unique aspects associated with assessing work stress that have shaped the formulation of three key research questions. In the following chapter, I will explain the connection between these research questions and the studies conducted in my dissertation, and then present the four individual studies.

## 2.1 A Definition of Stress in the Context of Occupational Health Psychology

The concept of stress was primarily shaped by the physiological and biomedical research of Selye (1976), who initially defined it as the adaptive response of the body to any harmful stimulus. This perspective aligns with reaction-oriented stress model, wherein stress is regarded as a dependent variable, representing a stress reaction. Within such biologically oriented models, stress is often equated with arousal (Koolhaas et al., 2011). Subsequently, the concept was further refined by differentiating between the "stress response" and the "stressor". Therefore, stimulus-oriented models have emerged, which conceptualize stress as an independent variable characterized by stress-inducing factors within the environment (e.g., Wheaton, 1999). These factors can include individual life circumstances, critical life events, or occupational demands that impact an individual and lead to strain (Rom & Reznick, 2016). To illustrate the distinction between reaction- and stimulus-oriented approaches, two common everyday statements can be cited: the expression "I am stressed" typically refers to personal states (reactions), such as irritability and inner restlessness, while the phrase "I have stress" usually refers to external circumstances

(stimuli), such as time pressure. While both approaches provide a basic understanding of stress, they do not take into account interindividual differences in the appraisal and management of stressors.

One of the primary theoretical frameworks explaining interindividual differences in stress management is Lazarus and Folkman's (1987) transactional stress model, which is based on a multistage process of appraisal and coping. According to this model, stress results from an individual's appraisal of the transaction between him or herself and the environment, specifically through primary and secondary appraisals. Primary appraisal involves perceiving a situation as challenging or threatening, while secondary appraisal focuses on evaluating the resources available to cope with the situation. The model also suggests that the coping strategies chosen by an individual significantly influence the outcome of the stressor, resulting in either successful adaptation or further distress. To effectively manage the situation, individuals can employ problem-focused coping, involving concrete efforts to change the situation, or emotion-focused coping, which centers on regulating the associated emotions. Following the implementation of a coping strategy, a reassessment of the situation occurs (reappraisal). If necessary, the situation is evaluated differently; otherwise, adjustments are made to the coping behavior if the stress persists. This model highlights the significance of an individual's cognitive evaluation of the stressor and their coping strategies in understanding the stress response, thus serving as a crucial framework for the development of work-related stress models that integrate organizational and individual stressors, the individual's appraisal of these stressors, and their responses to stress encompassing both short-term and long-term voluntary as well as involuntary reactions (Schuler, 1982).

Systematic research on work stress gained momentum in the late 1970s with the introduction of the job strain model, also known as the job demand-control model (JD-C model) by Karasek (1979). According to the JD-C model, an individual's experience of job strain is influenced by the interaction between job demands and job control. Job demands encompass the physical and psychological requirements of the job, such as workload and time pressure, while job control refers to the level of autonomy and decision-making latitude an individual has in their work. The model suggests that high job demands coupled with low job control can result in elevated levels of stress, whereas high job control can act as a buffer against the negative effects of high job demands (Karasek, 1979; Karasek & Theorell, 1990). An alternative model, the effort-reward imbalance model (ERI model), was introduced by Siegrist in 1996. In contrast to the JD-C model, the ERI model focuses on the rewards derived from work rather than the control aspect. The model posits that a mismatch between an individual's effort and the rewards they receive can lead to feelings of frustration, dissatisfaction, and diminished job satisfaction, ultimately leading to detrimental consequences such as stress and burnout over time. This imbalance can be exacerbated by a lack of control or autonomy over one's work and a lack of social support. It is important to note that rewards in this context include not only financial compensation, but also non-monetary aspects such as recognition, autonomy, and opportunities for growth and development. The presence of a high effort-low reward combination in the workplace has been recognized as a risk factor for various health issues, including cardiovascular problems, minor psychiatric disorders, and burnout (e.g., van Vegchel et al., 2005). In comparison to the JD-C model, the ERI model incorporates an additional personal component by defining overcommitment as excessive striving and a strong desire for approval, which may influence the

relationship between effort-reward imbalance and employee well-being (de Jonge et al., 2000). Therefore, the ERI model takes into account both structural and personal factors when examining the association between work stress and health outcomes. Consequently, both the JD-C model and the ERI model assert that job strain arises when there is a dearth of job resources, such as autonomy and rewards, in the face of high job demands. However, subsequent research has also found that social support from colleagues and supervisors is a critical resource (e.g., Viswesvaran et al., 1999). Although attempts have been made to incorporate social support into the JD-C model (job demand-control-support model, Johnson & Hall, 1988), the individual characteristics inherent to diverse job contexts are not adequately accounted for, as these models lack the necessary flexibility for comprehensive integration (Xanthopoulou et al., 2007). However, given the intricacies of the modern work environment and the diversity of job roles, both models appear overly simplistic in their approach (de Jonge et al., 2010).

One solution to address the limitations of previous work stress models in comprehensively capturing individual job characteristics is offered by the job demand-resource model (JD-R model). Introduced by Demerouti and colleagues (2001), the JD-R model has undergone extensive development (Bakker et al., 2023). The model proposes that all job characteristics, regardless of industry or organization, can be categorized into two groups: job demands and job resources. Job demands refer to the aspects of a job that require sustained physical, cognitive, and/or emotional effort and are associated with physiological and/or psychological costs, such as physical, psychological, social, or organizational demands. On the contrary, job resources encompass the facets of a job that possess motivational potential, contribute to the attainment of work objectives, moderate the impact of job demands,

and facilitate learning and personal development. These resources can be categorized as physical, psychological, social, or organizational in nature (Bakker et al., 2007). The JD-R model posits two distinct processes initiated by job demands and resources. The health impairment process suggests that high frequency and severity of job demands deplete employees' physical, emotional, and cognitive resources, leading to exhaustion and health problems. In turn, the motivational process suggests that job resources satisfy individuals' basic psychological needs, fostering work engagement and subsequently enhancing performance and creativity. Moreover, the interaction between job demands and resources has a synergistic effect on employee well-being. According to the buffer hypothesis, job resources act as a protective factor that mitigates the negative impact of job demands on strain. Conversely, the boost hypothesis posits that job demands enhance the positive influence of job resources on work engagement and motivation. Furthermore, personal resources, including traits like optimism and self-efficacy, have been found to play a significant role in mitigating the effects of job demands on employee well-being. Employees possess the capacity to actively shape their job demands and resources through job crafting, which can foster a positive cycle of work engagement and performance. However, when individuals experience job strain, it can trigger maladaptive cognitive processes and behaviors related to self-regulation, leading to a detrimental cycle of self-undermining actions and an escalation of job demands (Bakker et al., 2023). The JD-R model has prompted significant research in the field of occupational health psychology, including systematic reviews and meta-analyses, validating its multifaceted approach and proposed connections between job characteristics, well-being, health, and job performance (e.g., Alarcon, 2011; Crawford et al., 2010; Nahrgang et al., 2011; Rattrie & Kittler, 2014). Moreover, a meta-analysis focusing on

longitudinal studies confirmed the JD-R model's causal assumptions, demonstrating that stress and burnout increase when demands are high or resources are lacking, and strong resources can buffer the negative effects of demands on burnout (Lesener et al., 2019).

Another solution to comprehensively account for the diversity of individual job characteristics has emerged from German-language occupational health research. This approach employs action theory to construct a framework for job characteristics based on the hierarchical-sequential model of action regulation (Hacker, 1973, 1986; Volpert, 1982). This model provides insight into the hierarchical structure of goals and plans that govern the progression of actions, leading to the identification of three distinct categories of work characteristics: regulation requirements, regulation possibilities, and regulation problems. Regulation requirements, such as workplace complexity and variability, and regulation possibilities, such as task control and time control, are linked to the overarching concept of resources, while regulation problems, such as work interruptions and time pressure, are viewed as stressors derived from action theory (Irmer et al., 2019; Semmer et al., 1999). While Demerouti et al. (2001) intentionally employed the term "demands" in their JD-R model, preferring to reserve the term "stressor" for external factors with the potential to negatively influence most people in most situations, job demands can indeed transform into job stressors when they require high effort and individuals lack sufficient recovery opportunities (Meijman & Mulder, 1998). Additionally, Demerouti et al. (2001) note that job demands align with stressors identified in other work stress models, as supported by meta-analytic evidence (Irmer et al., 2019). Therefore, it can be assumed that job demands align with the concept of work stressors. As the first study in this dissertation is based on a work stress survey developed from the action theory

approach (Semmer et al., 1999), the term "stressor" will be consistently employed throughout this dissertation to denote the concept of demands in the JD-R model. Consequently, the assessment of work stress will encompass the measurement of both stressors and resources.

### 2.2. The Importance of Measuring Work Stress

Measuring work stress plays a crucial role in the field of occupational health and safety (Erazo-Chamorro et al., 2022). Given the significance of work in people's lives, both in terms of providing structure and meaning, the workplace holds great potential for promoting health and well-being (e.g., Blustein, 2006; Parker, 2014; Spector, 2021). Stress-related job characteristics pose significant occupational hazards, leading to various adverse health outcomes, including cardiovascular diseases and mental illnesses, making the assessment of work stress of utmost importance (Niedhammer et al., 2021). Among these health outcomes, mental illnesses are particularly noteworthy. In Germany, sick days attributed to mental illnesses, primarily depression, have increased by 41.0% since 2011 (DAK Gesundheit, 2022). Depression, anxiety, and adjustment disorders not only contribute to long-term absenteeism but also have far-reaching effects on medical care utilization, personal functioning, and productivity (OECD, 2015). The economic impact of mental health issues in EU countries amounts to approximately EUR 600 billion annually, encompassing healthcare and social security expenses, as well as decreased worker productivity (OECD, 2018). Thus, measuring work-related stressors and resources enables the early identification of individuals at risk and may help mitigate the potential for negative health outcomes.

Furthermore, there are several compelling reasons why organizations need to be aware of work stress, including the frequency and severity of stressors and the

availability of resources. First, work stress can significantly impact an employee's ability to concentrate and work efficiently, resulting in lower quality work and decreased overall performance (e.g., Ford et al., 2011). Second, a stressful work environment can have a detrimental effect on employees' job satisfaction, leading to lower levels of commitment and morale, and potentially encouraging counterproductive work behaviors (Fox et al., 2001). Third, work stress is strongly linked to higher rates of sick leave and absenteeism. Last, stress may also contribute to higher turnover rates, as employees may seek employment in less stressful environments (Dewe et al., 2012). Collectively, all of these factors create significant disadvantages for organizations, resulting in high costs associated with absenteeism, productivity losses, and continued investment in recruitment and training efforts for new employees (e.g., Kocakulah et al., 2016; Mitchell & Bates, 2011). Therefore, it is crucial for organizations to identify and address work stress in order to minimize its negative consequences.

Moreover, the importance of addressing work stress has been recognized by legislators, who have placed legal responsibilities on employers to provide a safe working environment. In several European countries, such as Austria, the Netherlands, and France, the assessment of mental stress risk is legally mandated (Leka et al., 2017). In Germany, the assessment of work-related stressors and strain is anchored in the Occupational Safety and Health Act (§5 Absatz 3 Satz 6 ArbschG). According to this regulation, employers are obligated to identify and record stress-related job characteristics if there are indications that they could jeopardize the health of employees. The risk assessment of work stress is a crucial aspect of occupational health and safety, aiming to identify and mitigate potential risks at an early stage, thereby protecting employees' mental health. While the legislation does not specify

the specific methods to be employed by employers for the risk assessment of work stress, it is imperative to utilize high-quality instruments to obtain reliable results.

### 2.3 Challenges in Measuring Work Stress

Concurrently with the development of the aforementioned work stress models, instruments for assessing stress-related job characteristics have emerged within the framework of job analysis (e.g., Work Design Questionnaire, Morgeson & Humphrey, 2006; Instrument for Stress-Oriented Task Analysis, Semmer et al., 1999). Job analysis involves various activities aimed at exploring, understanding, and describing the nature of work tasks performed by individuals (Morgeson et al., 2019). These activities include describing and enumerating specific work steps within a job, as well as identifying the necessary skills, abilities, and tools required. Such descriptive content can be collected through direct observation by an analyst.

However, assessing work stress is more challenging because research on stress has shown that subjective appraisals of stressors and resources play a central role in shaping the experience of stress (Lazarus & Folkman, 1987; Perrewé & Zellars, 1999). Consequently, the common approach to assessing stress-related job characteristics, including stressors and resources, is through self-report measures. Self-reports are highly cost-effective, easily to administer, and can accommodate the participation of a large numbers of employees. Additionally, the individuals performing the work are considered experts on their own work and the related working conditions (Spector & Eatough, 2013). However, it is important to acknowledge that self-reports are based solely on subjective assessments, and, as such, have inherent challenges that must be considered when using them (Paulhus & Vazire, 2007).

### *2.3.1 Cognitive and Motivational Processes in Survey Response Behavior*

A substantial body of psychological research has demonstrated that self-report measures are susceptible to bias stemming from cognitive and motivational sources of inaccuracy (e.g., van de Mortel, 2008). Morgeson and Campion (1997, 2012) have highlighted various sources of bias and emphasized the importance of mitigating or minimizing their potential negative effects to ensure data quality (Morgeson & Dierdorff, 2011). In addition, studies have shown that individuals may experience difficulties when completing self-report surveys (Bradburn et al., 1987; Hill et al., 2019; Lievens et al., 2008). It is important to recognize that responding to survey items involves a complex cognitive process with multiple steps for the respondent (Schwarz, 2007; Tourangeau, 1984). For example, describing the frequency of an event or condition requires various cognitive processes that depend on the relative frequency of the event (Blair & Burton, 1987). Furthermore, features of the survey instrument, such as format, question wording, context, and response scale, play a significant role in influencing self-report responses (Schwarz, 1999).

In addition to cognitive processes during survey completion, motivations can also have a significant impact on response behavior. Different motivations can lead individuals to respond in different ways, potentially affecting the accuracy and validity of the information collected (e.g., Donaldson & Grant-Vallone, 2002). For instance, individuals motivated by a desire to be perceived positively may engage in social desirability bias, in which they are inclined to respond in a socially acceptable or positively viewed manner. This may result in overreporting of positive characteristics and underreporting of negative characteristics, thereby distorting the true representation of the individual (Krumpal, 2013). Similarly, individuals who are motivated to impress or gain an advantage may employ impression management

techniques, manipulating information to create a desired image. This can involve exaggeration or manipulation of self-reported information (Bolino et al., 2016). For instance, in the context of personnel selection, applicants may engage in impression management and faking to enhance their chances of securing a job (Ingold et al., 2015). Motivations can vary depending on the situation or context, suggesting that specific motivational processes may underlie employees' response behavior in work stress surveys and thus influence the outcomes.

The literature on survey methodology in general offers various suggestions on how to address these issues (e.g., de Leeuw et al., 2008). While there are numerous recommendations for survey usage in organizational contexts pertaining to content, language, format, measurement, and administration, it becomes apparent that applying these recommendations to the context of work stress is not a straightforward task. For instance, Morrel-Samuels (2002) proposes asking for observable behavior to minimize bias, but the content queried in work stress surveys often pertains to aspects that are not directly observable (cf., Debus et al., 2015). Moreover, as stress research has shown, the subjective appraisal of stressors and resources is crucial to the experience of stress (Lazarus & Folkman, 1987; Perrewé & Zellars, 1999). Consequently, it can be assumed that response behavior in work stress surveys is influenced by a multitude of cognitive and motivational processes that require deeper understanding to effectively address the issue of bias.

### *2.3.2. The Influence of Contextualization on Survey Response Behavior*

Survey items are often worded in general terms to facilitate wide applicability. However, this generality creates room for interpretation, leading respondents feeling uncertain about how to interpret and respond to items (e.g., Hill et al., 2019). It has been suggested that respondents tend to assign their own personal meaning to these

generic items, resulting in varied interpretations and posing challenges for precise measurement (Credé et al., 2010; Davison & Bing, 2009; Holtz et al., 2005; Lievens et al., 2008; Schmit et al., 1995). To address this issue, the idea of contextualizing items by incorporating a specific context into a context-free scale emerged. This phenomenon, referred to as the frame-of-reference effect, has traditionally been investigated in the context of personality measures (e.g., Schmit et al., 1995). Evidence suggests that incorporating elements of situational context into personality assessments improves predictive validity (Shaffer & Postlethwaite, 2012; Swift & Peterson, 2019) and increases reliability. The positive effects of the frame of reference on reliability are attributed to the reduction of intraindividual variance, whereas the effects on validity are attributed to the reduction of interindividual variance. These effects are influenced by the number of participants using the correct frame of reference during testing and the number of items that are rated using the correct frame of reference during item response (Lievens et al., 2008).

There are several ways to use a frame of reference to contextualize an item. In the realm of personality measures, a situational frame of reference (e.g., at work or at school) is often employed. General items are typically tagged with a situational context, such as "at work." In work stress surveys, the reference to work is already given, so the challenge lies in determining which situational frame leads to greater consistency in respondents' response behavior. Another approach is the use of a temporal frame, where items are provided with a specific time period (e.g., the last six months), prompting participants to consider that timeframe when responding. Initial evidence suggests that a temporal frame of reference can be beneficial in work stress surveys, as it influences the assessment of stressors like job insecurity (Debus et al., 2019). Another contextualization method, borrowed from comparative cultural

research, is the reference-group effect. Respondents in personality tests are asked to compare themselves with individuals of the same age or gender, resulting in improved predictability (Credé et al., 2010). Furthermore, the extent of contextualization can be varied, ranging from providing a context element in the instruction (low) to tagging every item with a context (medium) or redesigning every item to incorporate the context (high). Given the positive results, particularly in terms of predictive validity, achieved through contextualization, it is worth considering this approach to enhance the quality of the scale.

In conclusion, work stress is a significant concern that organizations must address by identifying and measuring stressors and resources, allowing for the implementation of effective interventions. Self-report methods offer an economical means of measuring stress and enable employees to provide their own perspectives as experts on their stressor and resource levels. However, as stress is a subjective experience, self-report methods may be susceptible to bias. Drawing insights from other fields, such as personality measures, can offer valuable knowledge to enhance the measurement of data through self-reports. However, it remains uncertain whether all of this knowledge can be seamlessly transferred to the context of work stress surveys.

## 2.4 Open Research Questions

My dissertation aims to improve the quality of work stress surveys in order to obtain reliable and meaningful data on work-related stressors and resources. To achieve this, I will address three key research questions using a sequential mixed-methods approach that combines qualitative and quantitative research methods.

First, what is the response behavior of employees in work stress surveys via self- and other-reports and what cognitive and motivational processes, as well as other

influences, can be described? Second, what are the effects of contextualizing scales used to measure work stress and can this improve the quality of the scales? Third, is there a way to measure defensive biasing as a form of motivational process that influences response behavior in work stress surveys, and what factors might trigger this motivational process?

By addressing these research questions, I aim to gain a deeper understanding of the complexities underlying work stress surveys and provide insights to improve their quality. The sequential mixed-methods approach allows for a comprehensive exploration of the topic, combining qualitative insights with quantitative data analysis to provide a nuanced understanding of the research questions at hand.

# 3 Studies

The following chapter of this dissertation contains four studies aimed at addressing the aforementioned research questions. Study 1 focuses on the first research question, aiming to gain a deeper understanding of the response behavior in work stress surveys in general and possible differences between self- and other-report. Studies 2, 3, and 4 build on the qualitative findings of Study 1 and aim to examine them quantitatively. Studies 2 and 3 address the second research question by examining the effects of contextualization on work stress survey scales in terms of means, reliability, and criterion validity.

Given the limited knowledge on this topic, a qualitative research approach using grounded theory (Glaser & Strauss, 1967) was chosen for Study 1. Through interviews guided by a constantly evolving interview guide, I investigated the response behavior in relation to cognitive and motivational processes and other potential influences. Additionally, I explored potential differences between self-report and other-report responses in work stress surveys. Based on the interview analyses, I developed a conceptual model of response behavior, focusing on the cognitive processes of respondents (interpretation, information retrieval, judgment formatting, response formatting) and the influences of situational framing on motivational processes, including intentions (e.g., protecting against negative consequences) and actions (e.g., trivialization of stressors). Furthermore, I present the similarities and differences in response behavior among self-report, supervisor-report, and coworker-report perspectives.

Study 2 focuses on the difference between a generic version and a highly contextualized version of work stress survey items. Expert-selected generic items from commonly used work stress surveys (Copenhagen Psychosocial Questionnaire,

COPSOQ, Kristensen et al., 2005; Nübling et al., 2005) were redesigned based on the work context of medical assistants in a doctor's office, following Holtrop et al.'s (2014) approach. Participants completed a paper-pencil survey that included both the generic and contextualized versions of the items, with the order randomized. Means and reliabilities of the two conditions were compared, and psychological well-being served as a criterion to assess whether the contextualized version had higher criterion validity than the generic version.

Similarly, Study 3 examines the contextualization effects of a forced social comparison process on responses to work-related stressors and resources items. Selected items from the ISTA (Semmer et al., 1999) and COPSOQ (Kristensen et al., 2005; Nübling et al., 2005) were tagged with the context of "in comparison with your direct colleague," and their impact on means, reliability, and criterion validity was investigated. Participants were randomly assigned to either a comparison group or a non-comparison group. Additionally, the study investigated whether employees tend to engage in upward or downward social comparisons during the social comparison processes.

Study 4 focuses on a motivational process that emerged from the qualitative interviews in Study 1, namely, the understatement of work stress due to fear of negative consequences. Drawing on the concept of defensive silence (Brinsfield et al., 2009), I labeled this phenomenon "defensive biasing" in the context of responding to work stress surveys. As this is a novel area of study, a scale to measure defensive biasing had to be developed. With the assistance of subject matter experts, four items were created to assess defensive biasing in work stress surveys. Furthermore, based on qualitative statements from Study 1 and a literature review, it was hypothesized that factors such as anonymity, neuroticism, job insecurity, and trust in supervisor

may influence the occurrence of defensive biasing. Therefore, Study 4 quantitatively examines the existence of defensive biasing and its potential influencing factors.

In summary, this dissertation provides in-depth insights into the response behavior in work stress surveys, considering cognitive and motivational processes and other influencing factors. It demonstrates that job-related contextualization of survey items and the inclusion of contextual elements through a tagged social comparison process can impact mean scores and measurement quality of stressor and resource scales. Additionally, it uncovers a form of biasing termed "defensive biasing," which involves trivializing work-related stress due to fear of negative consequences. The dissertation also explores potential factors that may promote the occurrence of this motivational process.

## 3.1 Study 1

# Response Behavior in Work Stress Surveys: A Qualitative Study on Motivational and Cognitive Processes in Self- and Other-Reports

Berit Greulich[1], Maike E. Debus[2], Martin Kleinmann[3],

and Cornelius J. König[1]

[1]Universität des Saarlandes, Germany

[2]Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

[3]Universität Zürich, Switzerland

## Author Note

Berit Greulich https://orcid.org/0000-0002-5955-7423 and Cornelius J. König https://orcid.org/0000-0003-0477-8293, Industrial and Organizational Psychology, Saarland University, Germany. Maike E. Debus https://orcid.org/0000-0003-1040-540X, Institute of Labor Market and Socioeconomics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany and Martin Kleinmann https://orcid.org/0000-0002-0939-1349, Work and Organisational Psychology, University of Zurich, Switzerland.

This research was supported by the German Research Foundation (DFG) grant KO4244/4-1 to the last author.

Correspondence concerning this article should be addressed to Berit Greulich, Saarland University, Campus A1 3, 66123 Saarbrücken, Germany. E-mail: berit.greulich@gmail.com

**Abstract**

Work stressors have major consequences for employees' health and performance. Although organizations often ask employees to fill out work stress surveys regarding stressors and resources, the literature on survey responding offers only limited advice on how to formulate work stress surveys. Furthermore, self-, supervisor-, and coworker-reports show only low convergence. To deepen our understanding of motivational and cognitive processes when individuals respond to work stress surveys, we used a qualitative, grounded theory approach. We interviewed employees after they responded to representative items, asking them about their thoughts, motivational processes, potential factors that might have biased their responses, and the contexts they considered when responding. Since organizations are often also interested in other-reports of stress at work, we also interviewed supervisors and coworkers. We reached theoretical saturation after 31 interviews. A multi-stage coding-process with three raters resulted in new theoretical findings regarding motivational processes, comparisons, and differences between self- and other-reports. For example, employees sometimes deliberately distort answers for fear of consequences. Furthermore, employees, supervisors, and coworkers undergo different comparison processes. The findings of this study suggest that more specific and context-rich wording of items may lead to a more reliable and comparable assessment of stressors and resources at work.

*Keywords*: work stressors, survey, biasing factors, contextualization, frame of reference, self-report, other-report, qualitative study

**Response Behavior in Work Stress Surveys: A Qualitative Study on Motivational and Cognitive Processes in Self- and Other-Reports**

Work-related stress is a ubiquitous phenomenon, costing Western Europe, Australia, Canada, and the United States a combined US$ 221.13 million to US$ 187 billion per year (Hassard et al., 2018). By definition, work stressors refer to the degree to which the work environment contains stimuli that require sustained cognitive, emotional, or physical effort. Research has linked work stressors to a number of detrimental outcomes, such as impaired psychological health and wellbeing (e.g., anxiety, depression), impaired attitudes (e.g., Powell, Enright, & Enright, 2015; Semmer, Grebner, & Elfering, 2003; Spector, Chen, & O'Connell, 2000; Theorell et al., 2015), and aggressive or counterproductive work behavior (e.g., Berry, Ones, & Sackett, 2007; Fida et al., 2015). These effects of stressors can be buffered by work resources, broadly referring to a kind of energy reservoir that a person taps into (e.g., de Jonge & Dormann, 2006).

In light of the largely negative outcomes of work stressors, organizations are typically interested in monitoring stressors and resources in the workplace and in developing interventions to address them. One of the most prevalent methods to monitor and assess work stressors and resources are work stress surveys; this method directly reflects employees' perspectives (Perrewé & Zellars, 1999; Spector & Eatough, 2013) and constitutes a particularly economical way to collect data. However, filling out such surveys is typically a difficult task for respondents, because such surveys often consist of fairly generically formulated items − thus rendering it unlikely that all persons would interpret such items in the same way (e.g., Blaszczynski, Dumlao, & Lange, 1997; Credé, Bashshur, & Niehorster, 2010; Heine, Lehman, Peng, & Greenholtz, 2002). Furthermore, researchers have suggested that

when employees fill out such surveys, they often anticipate consequences – for example, they might believe that a particular outcome of the survey may cause the organization to try to change working conditions (e.g., Morgeson & Campion, 1997; Singer, 2001). Thus, when work stress surveys are initiated by employers, they might be affected by motivational processes, unlike surveys that are filled only for research purposes. In addition, stress at work is oftentimes also assessed by other-reports (e.g., supervisors and coworkers), but there is already evidence that these self-reports and other-reports do not converge that well (Debus et al., 2015; Meier & Semmer, 2018; Semmer et al., 2003), and it is unclear which motivational and cognitive processes may lead to such differences.

In the present qualitative interview study, we aim to gain a deeper understanding of the different processes that occur when employees answer work stress surveys. In doing so, this study contributes to the literature in several ways. First, we develop new theoretical arguments about answering stress surveys by identifying and analyzing a broad range of possible motivational and cognitive processes, particularly their impact on response behavior. Second, we deepen our understanding of item contextualization in work-related stress surveys. In particular, we provide information on contextualization mechanisms through different time frames, comparison processes, and situational frames. Finally, this study contributes to a better understanding of similarities and differences among the three perspectives: self-, supervisor-, and coworker-report. These findings can be used to complement and modify measurement methods to assess work stressors and resources and thus obtain more reliable and valid results.

## Background

Many organizations have realized that work stress matters because it impairs employee job performance (Gilboa et al., 2008) and is related to higher absence rates (e.g., Schaufeli, Bakker, & Rhenen, 2009), which might cause additional workload for the remaining colleagues. Furthermore, work stress may also make employees engage in negative oral communication to potential customers and applicants (see Wetzer, Zeelenberg, & Pieters, 2007). In some countries, organizations are even legally required to assess work stressors and take preventative action (e.g., in Germany due to its "Arbeitsschutzgesetz" [German Occupational Safety and Health Act], Treier, 2019). Some organizations might also genuinely care about work stress just because the top management feels personally responsible for their employees and their well-being (Nielsen et al., 2017).

If an organization intends to better know their employees' working conditions in order to reduce stressors or increase resources, the organization will need to assess those working conditions. Thus far, several scales have been developed that assess employees' perceptions of occupational stressors and resources (e.g., Karasek et al., 1998; Kristensen, Borritz, Villadsen, & Christensen, 2005; Morgeson & Humphrey, 2006; Semmer et al., 1999; Van Katwyk et al., 2000). Commonly used response formats in such work-related stress surveys are frequency scales or rating scales. Sometimes, respondents are asked to evaluate whether they are more like person A or person B, with both A and B experiencing (diverging) work situations (see, e.g., Semmer et al., 1999).

If organizations plan to use stress survey data to initiate changes in employees' working conditions, they need to collect data that is unbiased and realistic. Otherwise, organizations might, for example, restructure a process such that the workload lessens

for the wrong people (i.e., for employees who just claim to have too much to do, e.g., Bolino, Kacmar, Turnley, & Gilstrap, 2008). Although collecting survey data might appear like an easy task, initial evidence from autobiographical questions (Bradburn et al., 1987) and personality research (Lievens et al., 2008) suggests that employees might find filling out such surveys fairly difficult. There is also evidence from the field of general workplace surveys that reports problems regarding content, format, language measurement, and administration (Morrel-Samuels, 2002). More precisely, items are often formulated in such a way that they allow a relatively large room for interpretation (e.g., what is "often" in a question like "How often are you under time pressure?"), which could then allow different biases to come into play (e.g., Menon, 1994; Morgeson & Campion, 1997, 2012; Schwarz, 1999, 2007).

To obtain data that give a realistic picture of working conditions, organizations could consult the literature on the survey method in general (e.g., Moser & Kalton, 2017; Sudman, Bradburn, & Schwarz, 1996). Here, they will find different suggestions for creating surveys that minimize such risks – each of them, however, being associated with some pitfalls as well. First, practitioners should be aware that respondents describing the frequency of events use a variety of cognitive processes depending on the relative frequency of the event (Blair & Burton, 1987): It seems easier to recall and count every instance of an infrequent behavior, whereas it becomes more difficult to recall and count frequent behaviors, meaning that respondents tend to use rate-based estimations (e.g., "three times within the last week"). It is, however, unclear what the implications of such findings are.

Second, the general literature on the survey method also advises that surveys should specify the context ("contextualization") by, for instance, adding "at work" to items. For example, Schmit et al. (1995) tagged each item of a conscientiousness scale

with a situational context ("I am courteous to everyone I meet *at work*"), which changed the means, reliability, and validity of the scale (see also, e.g., Credé et al., 2010; Holtrop et al., 2014). However, questionnaires assessing work stress are typically already contextualized to work because they aim to assess *work* stressors and *work* resources (e.g., "How often are there moments *at work* which require your highest level of concentration for a short time?" from the Instrument for Stress-Oriented Task Analysis; Semmer et al., 1999).

Related to this, the literature on cross-cultural survey research suggests the addition of another contextualization: a specific comparison target (e.g., compared to people in your country; Heine, Buchtel, & Norenzayan, 2008). Until now, many work-related stress surveys do not use a format that requires social comparisons. However, even if an organization follows this advice, it is currently unclear which comparison target might be most appropriate (e.g., "compared to your direct colleagues" or "compared to all colleagues in your organization"?).

Applying previous knowledge about survey responding to work stress surveys becomes even more complicated for two reasons. First, although these surveys are usually answered by the job incumbents themselves (because they are considered to be experts for their workplace), they are also used for other-reports such as supervisors and coworkers (Debus et al., 2015; Meier & Semmer, 2018). Studies thus far only demonstrate low convergence between self- and other-reports (Debus et al., 2015; Meier & Semmer, 2018), perhaps because the third party cannot fully observe the elements being evaluated (Vazire, 2010), or because of differences in motivational and cognitive processes between reporting about oneself and reporting about others (Berry et al., 2012). Second, filling out a questionnaire for a research project likely differs from filling out a work stress survey, because respondents might expect that

organizations will attempt to change working conditions based on the summary of the survey results (Pond et al., 1984; Ryzin, 2013). Thus, motivational processes like the hope for change might also change the way respondents answer a work stress survey – so much so that employees might even intentionally distort their answers.

The aim of this study is therefore to go one step back and to gain a better understanding on what is in people's minds when filling out stress surveys. To achieve this aim, we use a qualitative, grounded theory approach (Glaser & Strauss, 1967). Qualitative research offers a more nuanced understanding of these motivational and cognitive processes, and can open up new aspects of biasing factors that have not yet been considered (e.g., Flick, 2018). When respondents are interviewed while completing such surveys (e.g., employees, supervisors), it is possible to inquire about characteristics and variations of their responses and the motivation behind them. Based on our elaborations above, this qualitative study thus focused on the following three research questions:

*Research Question 1: How can respondents' motivational and cognitive processes when responding to work-related stress questionnaires be characterized, and how do these processes influence their responses?*

*Research Question 2: Which contexts do individuals have in mind while responding to work-related stress questionnaires and how do these contexts influence individuals' answers?*

*Research Question 3: Are there any differences between self- and other-reports via supervisors and coworkers regarding the contexts they considered and motivational and cognitive processes? What do these differences look like?*

# Methods

## Qualitative Approach

Although studies on work-related stress using quantitative methods have been important to the field, quantitative research approaches have their limitations. They can only use pre-existing knowledge, which makes it difficult to provide a deeper understanding of behavior and processes about which not much is known. Qualitative research, however, aims to subject certain phenomena to a deeper and more differentiated analysis; the procedure is mostly inductive and generates hypotheses and/or theories (unlike quantitative approaches). Subjective realities and subjective constructions of meaning and theories of everyday life are often also examined, and life worlds described "from within" – individual views and opinions or motives – are analyzed, with the aim of not only describing them in detail, but also of understanding them (Misoch, 2019). Representativeness is not characterized in a statistical sense but in terms of content, and the data are collected in interaction by means of communication (Misoch, 2019).

In our case, qualitative research deepens existing results from the work stress and survey literature and provides new knowledge, connections, and variables by detailing the personal experiences of people (Mazzola et al., 2011). The experience and behaviors are processual and can vary inter-individually, and personal interviews are therefore a key means of probing respondents' subjective experiences (Wimpenny & Gass, 2000), which make interviews a commonly employed qualitative method, also in the research field of work-related stress (Kinman & Jones, 2005). In such an interview, the interviewee is given space for his or her own views and the researcher has at the same time the opportunity to focus on particular aspects by asking questions during the interview process.

Grounded theory, the approach used in this study, is a qualitative method that is especially useful when there is only limited knowledge about a phenomenon but also when the aim is to better understand existing knowledge and to add new aspects to it (Glaser & Strauss, 1967). This method is becoming increasingly popular in industrial and organizational psychology (Bluhm et al., 2011), likely because it can help researchers to understand complex motivational and cognitive processes (Little et al., 2014; Wilhelmy et al., 2016). We considered grounded theory to be suitable for our study because little is known about other-reporting in the field of occupational stress and to deepen the knowledge about how participants respond to surveys, especially in the context of occupational stress. What makes grounded theory special is that the analysis of data does not follow data collection as in quantitative research; instead, data collection and analysis occurs simultaneously. This means that data collection iteratively influences the data analysis and vice versa. Although such iterative processes likely make qualitative studies more difficult to replicate than quantitative studies, there is some literature (Aguinis & Solarino, 2019) offering advice on how to achieve transparent and replicable qualitative studies, which we followed for this study.

**Sample**

Following grounded theory principles (Glaser & Strauss, 1967), we iteratively developed ideas regarding whom we could interview next (and which questions we could ask) based on statements from earlier interviews (i.e., theoretical sampling, Byrne, 2001). To get a deeper understanding of response behavior in work-related stress surveys, we interviewed (a) employees, who were asked to rate the stressors and resources in their own job (using a measure described below). To gain insights into other-reporting of work-related stress, we collected data from (b) supervisors,

who were asked to rate the work stressors and resources of an employee's job using the same measure. The simultaneous data evaluation revealed that participants thought that supervisors and coworkers in other-reporting would probably rate stressors and resources differently. We therefore also interviewed (c) coworkers, who were asked to rate the psychological stressors and resources of another employee's job using the same measure.

According to grounded theory, data saturation is reached when there is enough information to replicate the study, when the ability to obtain additional new information has been attained, and when further coding is no longer feasible (Fusch & Ness, 2015). To reach the data saturation of each group we followed the suggestions of Nelson (2017), who suggested iteratively evaluating the following five criteria for saturation: (a) range (i.e., Are multiple examples of concepts in the data provided?), (b) complexity (i.e., Is there a rich network of concepts and themes?), (c) subtlety (i.e., Have instances of same codes been compared to assess how they are similar or different?), (d) resonance (i.e., Can the emerging concepts be connected to the existing literature), and (e) validity (i.e., Do the concepts seem useful for others?). We applied these five criteria after each block of interviews (see below). Saturation for motivational and cognitive processes was reached after 21 and after ten interviews more for the differences between self- and other-report. Overall, we stopped data collection after 31 interviews. The demographic variables of the three subsamples are presented in Table 1.

All interviews were conducted in the German language by the same interviewer. We recruited the interviewees through our own personal networks and through suggestions from participating interviewees. If the interviewer was acquainted with the participant, the interview typically took place in private

surroundings. In the case of non-acquaintance, the interview usually occurred at the participant's workplace with the permission of the supervisor or management.

**Table 1**

*Demographic Variables of the Three Subsamples*

| | Employees ($n$ = 10) | Supervisors ($n$ = 10) | Coworkers ($n$ = 11) |
|---|---|---|---|
| $M_{age}$ ($SD$) | 32.4 (6.9) | 40.9 (9.0) | 40.8 (12.0) |
| Gender | 60% female | 50% female | 36% female |
| Experience with work stress surveys | 30% | 30% | 36% |
| Occupational sector (count) | Information and communication (2)<br>Manufacturing (2)<br>Education (1)<br>Human health and social work activities (4)<br>Administrative and support service activities (1) | Manufacturing (2)<br>Administrative and support service activities (2)<br>Professional, scientific and technical activities (2)<br>Public administration and defense (1)<br>Wholesale and retail trade (1)<br>Financial and insurance activities (1)<br>Human health and social work activities (1) | Administrative and support service activities (3)<br>Professional, scientific and technical activities (1)<br>Public administration and defense<br>Wholesale and retail trade (1)<br>Arts, entertainment and recreation (1)<br>Information and communication (2)<br>Wholesale and retail trade (1)<br>Human health and social work activities (2) |

## Data Collection and Data Analysis

In qualitative research, data analysis and data collection do not take place separately (as in quantitative studies), but rather occur iteratively. This means that after a first data collection block with about three to four interviews, a first data analysis block took place, the results of which were taken into account in the second data collection block. After the second block of interviews had also been evaluated, a third data collection block followed, followed by a third evaluation block, and so on. Data collection and data analysis thus influenced each other. The respective procedures in the individual survey and evaluation blocks were identical in each block and are described below.

### *Data Collection*

From the very first contact, we ensured that we wanted to investigate how individuals respond to work-related stress questionnaires. Therefore, participants were first shown example items of work-related stressors and resources so that they had a clear idea of what kind of stress the study is about. As suggested by grounded theory (Glaser & Strauss, 1967), we used an interview guide (see Table 2), which we amended over the course of the study. We initially interviewed participants immediately after each item from the work stress survey described below. However, we later realized that interviewees reported too many details about their work life instead of providing enough insights into their cognitive processes. Therefore, we changed this procedure after the first five interviewees, and from interview six onwards, employees were interviewed after they completed the whole survey. This revised approach made it easier for participants to report on and talk about how they answered the work stress survey, which is why we followed this approach for all the other interviews. In line with Bluhm et al. (2011), we complemented the interview by

making memos following each in-depth interview and also during the coding process. These memos were later used to document ideas for data interpretation.

**Table 2**

*Interview Guideline at the Start of the Study*

| |
|---|
| 1. How did you go about answering the questions? What were your thoughts while you were answering the questions? |
| 2. Did you have a particular time frame in mind? If so, which? Did this vary from question to question? |
| 3. To answer some of the questions, did you draw any comparisons, e.g., with coworkers, supervisors, friends, family, other professions or industries? |
| 4. Did you have difficulties understanding some of the questions? |
| 5. Are there any factors which might influence your answers? Would you answer differently under certain circumstances? |
| 6. What are important stressors at your workplace? |
| 7. What expectations do you think the employees would have if a mental health risk assessment were conducted with surveys? |
| 8. Did you notice anything, or would you like to say anything we haven't yet discussed? |

### *Work Stress Survey*

When designing the study, we compared several work-stress instruments and found that these instruments greatly overlapped with regard to the stressors and resources that were covered. We finally decided to use a selection of items taken from the Instrument for Stress-Oriented Task Analysis (ISTA, Semmer et al., 1999), one of the most commonly used and popular work stress surveys in Germany that covers the most important stressors and resources (Irmer et al., 2019). The ISTA has also been employed in internationally published research (e.g., Debus, Sonnentag, Deutsch, & Nussbeck, 2014; Frese, Kring, Soose, & Zempel, 1996; Schmitt, Ohly, & Kleespies,

2015; Semmer et al., 2015; Sonnentag, Binnewies, & Mojza, 2010). It comprises 19 scales that assess typical stressors and resources (e.g., autonomy, work interruptions, variability, complexity, etc.). Response formats are diverse, ranging from comparisons (e.g. “Are you more like A or B?”) to frequency scales with a 5- or 6-point format (e.g., very little, fairly little, somewhat, quite a lot, very much). For item selection, we defined two core criteria: (a) every scale from the ISTA and (b) each response format (e.g., frequency, intensity, examples of behavior) should be represented once. A selection based on item loading was not possible due to lack of data. In total, we therefore used 19 items, which are displayed in Table 3.

**Table 3**

*Example Items from the ISTA (Instrument for Stress-Oriented Task Analysis, Semmer et al., 1999)*

| Scale | Item | Response |
|---|---|---|
| Qualification requirements | 1. If you consider your job as a whole – how much qualification does it require? | Very little<br>Fairly little<br>Some<br>Quite a lot<br>Very much |
| Complexity | 2. Coworker A works on tasks which s/he needs to first plan precisely before s/he can carry them out. Coworker B works on tasks for which no planning is necessary. Which of the two jobs is most similar to your job? | Exactly like that of A<br>Similar to that of A<br>Between A and B<br>Similar to that of B<br>Exactly like that of B |
| Scope for action | 3. If you consider your work as a whole, how much possibility to make your own decisions does your work offer? | Very little<br>Fairly little<br>Some<br>Quite a lot<br>Very much |
| Variability | 4. A has to do a very large amount of routine tasks in his/her work. In B's work, tasks rarely repeat themselves. Which of the two jobs is most similar to your job? | Exactly like that of A<br>Similar to that of A<br>Between A and B<br>Similar to that of B<br>Exactly like that of B |
| Participation | 5. In the following, we would like to know how much influence you have on how your breaks are organized. | I have no influence at all<br>I'm just informed<br>I'm able to make suggestions<br>I'm involved in the decision<br>I have a lot of influence on the decision |
| Time scope | 6. To what extent can you determine your work speed yourself? | Very little<br>Fairly little<br>Somewhat<br>Quite a lot<br>Very much |

| | | | |
|---|---|---|---|
| Uncertainty | 7. | How often does it happen in your work that you have to make decisions without sufficient information being available? | Very rarely/never<br>Rarely<br>Occasionally<br>Often<br>Very often/constantly |
| Risk of accidents | 8. | If you aren't very careful here at work, an accident can happen easily. | Not true<br>Barely true<br>Somewhat true<br>Mostly true<br>Completely true |
| Problems with the organization of work | 9. | A has to spend a lot of time procuring information, materials or tools to carry on working. B always has the necessary information, materials or tools available. Which of the two jobs is most similar to your own? | Exactly like that of A<br>Similar to that of A<br>Between A and B<br>Similar to that of B<br>Exactly like that of B |
| One-sided burden | 10. | To what extent is your work physically varied? | Very little<br>Fairly little<br>Somewhat<br>Quite a lot<br>Very much |
| Work interruptions | 11. | How often are you interrupted by your supervisors/other coworkers/employees/clients while you are working? | Very rarely/never<br>Rarely (about once a week)<br>Occasionally (about once a day)<br>Often (several times a day)<br>Very often (several times an hour) |
| Environmental burden | 12. | To what degree is your workplace characterized by the environmental factor "noise"? | Very low/not at all<br>Low<br>Moderate<br>High<br>Very high |
| Work interruptions | 13. | Does it happen that you have to interrupt your current work because something important comes up? | Very rarely/never<br>Rarely (about once a week)<br>Occasionally (about once a day)<br>Often (several times a day)<br>Very often (several times an hour) |

| | | |
|---|---|---|
| Concentration demands | 14. How often there are moments at work which for a short time require your highest level of concentration? | Very rarely/never<br>Rarely (about once a week)<br>Occasionally (about once a day)<br>Often (several times a day)<br>Very often (several times an hour) |
| Scope for cooperation | 15. How often are you under time pressure? | Very rarely/never<br>Rarely (about once a week)<br>Occasionally (about once a day)<br>Often (several times a day)<br>Very often (several times an hour) |
| Communication possibilities | 16. During your work, can you talk to coworkers about things that have got nothing to do with work? | Yes, without difficulty<br>Yes, with little difficulty<br>With some difficulty<br>With a fair amount of difficulty<br>The work doesn't permit it in practical terms |
| Scope for cooperation | 17. A can generally decide which coworkers he or she works with. For B it is always precisely specified with whom he or she works. Which of the two jobs is most similar to your own? | Exactly like that of A<br>Similar to that of A<br>Between A and B<br>Similar to that of B<br>Exactly like that of B |
| Closeness of cooperation | 18. In your work, how strongly are you dependent on how well or how poorly your coworkers are working? | Not at all<br>A little<br>Somewhat<br>A lot<br>Very much |
| Cooperation requirements | 19. To what extent does your work require joint planning with your coworkers? | Very strongly<br>Strongly<br>Moderately<br>A little<br>Very little/not at all |

### *Interview Guide*

Following an orienting theoretical perspective (Locke, 2001), the interview guide was derived from insights from the existing literature on survey methodology and frames of reference. Table 2 shows the four different aspects that were covered by this guide: (a) chosen time frames and comparisons for responding to the items, (b) other biases while responding to the items, (c) typical work stressors, and (d) expected consequences/outcomes of survey results. All interviews began with an open question ("How did you proceed in responding to the items?") and continued depending on participants' answers. A core characteristic of grounded theory is that data analysis influences the strategy of data collection and vice versa (Glaser & Strauss, 1967). Therefore, depending on the insights we gained, the interview questions were continuously adapted during the data collection process; that is, questions asked at the beginning of the research process were different to those asked later on. For instance, some later questions were used to verify ideas that emerged from earlier interviews (e.g., "Which expectations do you think employees would have of psychological risk assessment with surveys?"). When interviewing the coworkers and supervisors, we added questions about the change of perspective while rating work stressors (e.g., "Are there aspects that are difficult for you as a supervisor to judge because you are not directly in the job to be assessed?").

At the beginning of every interview, the interviewer introduced herself and explained the procedure regarding anonymity, confidentiality, and the processing of the data. In the course of data processing, each interview was assigned a number and no company names or names of persons were written down during the transcription (only abbreviations). The participants gave their consent for the audio-recording of the interview and for the use of their data for scientific purposes after our

confirmation that we would anonymize their data (i.e., only use number codes to distinguish them when we report their codes). At the end of the interview, participants completed a short demographic survey (i.e., gender, age, highest educational level, profession, work experience, position in the organization, and experience with work-related stress surveys). The in-depth interviews lasted on average 16 minutes (*SD* = 9). Each interview was transcribed verbatim immediately afterwards, with all transcribed interviews together amounting to 434 double-spaced pages.

***Content Analysis***

As already mentioned, this data analysis always took place in alternation with data collection (see above). Following Kreiner et al. (2009) and Wilhelmy et al. (2016), we used a multi-step coding system to analyze each interview. The whole coding process was also conducted in German. Multiple codes were iteratively derived from words, sentences, or passages of the written interviews, which the coders subsequently agreed upon. Codes are descriptive terms or shorthand sentences (e.g., “expectation,” “peak phases,” “comparison with coworker”) used to summarize and categorize the contents of the interviews (Lewis, 2015). During this process, emerging codes were documented in a dictionary that developed over time. There were three coders: the first author, who conducted the interviews, and two advanced psychology students (research assistants), one of whom also transcribed the interviews. Before the coding process, the first author trained the other two coders in how to code and use the coding software MAXQDA (VERBI Software, 2017).

Two raters (the first author who conducted the interviews and the research assistant who transcribed the interviews) independently inspected the interviews of each block (see above) sentence by sentence and provided passages with codes using MAXQDA before the further data collection started. A total of about 1200 codes were

assigned. As suggested by Strauss and Corbin (1990), two coders allow for multiple perspectives on the data, which should decrease personal biases. In an iterative process, the two raters agreed upon the coded text passages and their terminology. This resulted in a coding dictionary consisting of eleven superordinate categories with around 170 codes. In Table 4, we present the superordinate categories with frequencies of how often their subordinate codes were assigned to the interviews. However, it should be noted that data collection and analysis ensued reciprocally – depending on what has been said, questions were changed, dropped, or added. Therefore, the numbers should be interpreted with caution.

During the entire evaluation process, we regularly tested the plausibility of the terminology of codes and their assignment to the text passages by giving the dictionary to a third person (also a research assistant), who re-coded all of the transcripts with the codes from the dictionary. Subsequently, discussion among all three raters occurred until consensus about the terminology and coded passages was reached. The categories are not fixed in this step of analysis, which makes calculating agreement impossible.

When the coded text passages and their terminology was tested by the third person, we moved from the final codes in the dictionary to abstract categories (e.g., cognitive and motivational processes and contexts considered while responding to a work stress survey). The raters identified connections between the codes and developed subordinate categories and themes through ongoing comparisons of emerging ideas from the data. This process was repeated at several joint meetings. Any new categories and changes in categories were documented and equipped with descriptions and examples. This process resulted in a model that shows processes, contexts, and differences to other-reports.

**Table 4**

*Superordinate Categories with Frequencies from the Coding Dictionary*

| Category | Examples |
|---|---|
| Cognitive processes (187) | Referring to the activity or person, adaptation to job, visualization, imagination |
| Motivational processes (56) | Concerns about anonymity, worries about negative consequences, reducing frustration, protecting self-esteem |
| Expectation of change (34) | Stresses should be taken seriously, results are not taken seriously |
| Opinion about work-related stress questionnaire (26) | Skepticism, positive |
| Situational frames (126) | Economic situation, current stressors in the company / private |
| Time frames (163) | Short, long, special event, during project |
| Comparisons (93) | With colleagues, other tasks/jobs |
| Questionnaire/item construction (69) | Unspecific items, broad scope for interpretation, items do not fit with the job |
| Rating perspective (30) | Objective vs. subjective |

### *Interrater Agreement*

Following Kreiner and colleagues (2009), we gave the final categories and a representative sample of text passages from the interviews (10%, following Bluhm et al., 2011) to two additional research assistants (who were unfamiliar with the study). These research assistants were instructed to read each passage and assign it to the category to which they believe it fits best. The overall percentage of agreement between the two coders was 91%, and Cohens' $\kappa$ was .81, suggesting almost exact agreement (Fleiss & Cohen, 1973). For overall agreement between the two research assistants and the original coding, the percentage of agreement between the first coder

and original coding was 89% and for the second coder 93%. Cohen's κ was .85 and .87, respectively.

## Results

### Overview

For Research Question 1, we were interested in the characteristics of respondents' motivational and cognitive processes when responding to work-related stress questionnaires, as well as how these characteristics influence their responses. Respondents reported motivational processes that were specific to the context of stress at work and which influenced the answers in various ways; we also found several cognitive processes that were used by all three interviewed groups. Research Question 2 specifically addressed contexts. Respondents reported using different time frames, social comparisons, and situational conditions for contextualization. For Research Question 3, we asked about differences between internal and other-reporting in work-related stress questionnaires. The other-reporters (i.e., coworkers and supervisors) named problems with the perspective they were supposed to take as well as further motivational processes that influenced the answers. According to the intention of grounded theory, we set each of the findings in a model in relation to each other and put them into an overall context (see Figure 1). In the following sections, we first explain the motivational and cognitive processes and their influence on the answers, followed by the contexts considered and their impacts, and subsequently the differences between self- and other-reporting (see Figure 2). We exemplify all results with quotes from the interviews that were translated from German into English by a professional translator. All translated passages were re-read by the interviewer to ensure that the content remained the same.

## Research Question 1: Characteristics of Motivational and Cognitive Processes While Responding to Work Stress Items

### *Motivational Processes*

Participants reported on motivational processes that had influenced their response behavior in different ways. First, the general motivation or willingness to participate in the survey can be *positive* (i.e., employees have a positive expectation for the survey participation) or *negative* (i.e., employees are more likely to fear consequences if they participate). A positive outcome for the respondent could be the reduction of frustration, and the desire to change stress-producing factors. For example, if survey respondents were dissatisfied with their workplace, then such an employee survey could be a good outlet to voice this, as one respondent commented:

> I think, that's, um, a method that would help the employees to release a bit of their frustration, so that which they would otherwise not dare to say in conversation or with their leaders, they'd maybe write it down here, or tick the box. (#12)

Furthermore, respondents also mentioned their hopes that management would take the results seriously and take measures for improvement, as demonstrated by the following quote (#17, a nurse): "the expectation is there, when someone asks us about it, that a change will also take place, so because of us, not just because of the patients, but that the concern is also with us." However, this desire for change and the reduction of frustration can also lead to an exaggeration of the conditions in the workplace. For instance, #14 reported: "[…] but maybe even a bit, um, yes, a bit more strongly should I say? To make it clear that the coworkers aren't doing well." This is a motivational bias which aims to actively increase the possibility that change will follow. A decisive factor seems to be *anonymity*: "If you do it anonymously […] a

picture will emerge that everything, everyone is stressed and it can't go on like that and so on, because the opportunity is there for people to express themselves anonymously." (#11). If, however, respondents do not feel secure because of a lack of anonymity, they might fear negative consequences:

> If I can express it like this, I might also be expressing weaknesses, that you say, ok, I'm not coping with the stress, I can't, I can't manage the tasks, that you construe yourself negatively by expressing that. Or just, whether something will be inferred from it or not, to what extent it really would be anonymous. So I think that you would perhaps fear consequences. (#20)

Motivation might also be negative, meaning that *grievances are mitigated* (i.e., understated). Employees might avoid telling the truth in order to protect themselves from negative consequences, as exemplified by this respondent's experiences with a previous survey:

> I didn't tell the truth much because of fear of losing my job, of being recognized after all, even though it's anonymous, but this fear in the back of my mind, if I give a few data, male, married, age from, to, and you know your coworkers, can be aware of who might that have been. That's why no real evaluation occurred. (#15)

Thus, concerns about anonymity seem to play an important role in whether existing grievances are expressed accurately or are exaggerated, or whether respondents mitigate them for fear of suffering consequences. Furthermore, whether or not the rater thinks that participating in the survey will lead to good outcomes contributes to their ratings.

### *Cognitive Processes*

Participants described different cognitive processes: interpretation, information retrieval, judgment formation, and response formatting. We only briefly present these processes here because we found them to be similar to processes identified in previous studies about response behavior (e.g., Sudman, Bradburn, & Schwarz, 1996). In principle, individuals imagine their activities and environmental conditions with mental representations and generalize them for response formatting. For example, employee #09 stated: "I imagined a normal kind of daily routine. What happens at our place? How things go and according to this I tried to go through how the day proceeds on average." Others reported recalling particular situations. For instance, #07 stated: "I often went into situations that were very stressful," and one participant even mentally went through his Outlook calendar:

> Um (.) To be honest, at the moment I just went through my Outlook calendar in my head, I thought about when the last difficult assignment was, and it was just about in February and at the beginning of January. Not yet in March. That might still come. So therefore, I would say about once. Once, twice a month. Can't [unintelligible] really say. (#04)

After imagining, the interviewees tried to generalize their activities in order to come up with a response. For example, #17 described: "Then you try to form a kind of average from it" or considered daily situations, such as #07: "I actually thought about what kind of situations I know from work and then put myself in them again. So, I, um, then really purposefully searched for the situation and then tried to consider whether it is frequent or not."

## Research Question 2: Context Effects

### *Effects of Different Time Frames*

The data analysis revealed two different approaches to using a time frame to respond to work stress items. One approach is *choosing a specific period of time*, from a time in the past up to the present. How far respondents think into the past varies considerably – from one day to their whole working life. For example, #20 reported a very short time period: "I then thought about what there was now, for example today, for some questions," whereas others thought about whole weeks (e.g., #12), months (e.g., #01) or longer, such as #07: "I'd say rather six months to a year". Others mentioned that they had their entire working life in mind when answering the survey, like #02: "I'd go on the basis of my whole career, now really of the 20 years." The time frame seems to depend on how often something happens, with the period tending to be somewhat shorter for regular events, as #12 reported: "Often it was a short time period, the last two to four weeks, for questions where I was very certain, because I notice that it happens often or regularly." For rare events, the periods are longer (as mentioned by #15). The chosen time periods seem to vary both inter-individually and intra-individually. For instance, respondent #11 stated:

> Um, as I said, there are questions here that refer to, to hours, or per day, that's a really good help to be able to estimate it, but at the end of the day you have to consider after all that the employee doesn't do the same work every day, and therefore it's always the case, zoom in and zoom out, over, I estimate, a period of a few weeks. (#11 describing the stress of another employee)

The second approach is to *go back in time to a certain event in the past* and to evaluate work life until then. For instance, #05 stated: "Yes, maybe the time around

the last six months, because we had a change of CEO, so in September last year we had a new CEO […].” Similarly, #04 evaluated the time since he joined the company: “Um, so I’ve been here with my current employer since 2011. And I actually always think there over the whole time.” In any case, the use of different time frames influences the question asked, like #29 stated:

> So if it was defined as “Please now consider the last month”, then of course I’d also answer differently. Because in the last month, also the last 2 months, 3, we had quite a high workload. If the concern was with considering the whole period of time I’ve been doing this, then the response behavior is of course a bit different.

***Effects of Comparisons***

Around two thirds of the respondents made comparisons while assessing stressors and resources in their jobs. These comparisons can be differentiated into three different groups: The first group compared themselves with their *direct coworkers* with the same job, the second group made comparisons with coworkers *in their company but with another job*, and the third group made comparisons with *people who had another profession*. Some members of the first group already reported on this when they explained their general approach, like #02: “Yes, so I’ve also got more freedom than, um, (.) my coworkers who are at the workbench. I’ve just got more freedom than them, haven’t I?” #05 also reported a possible reason for the comparisons: If employees predominantly have coworkers with a lot of experience, they might compare themselves with those. However, this reasoning is questionable because the comparison can also be downward:

> The coworker who I’m thinking about, she’s a great coworker, so she’s got a lot of knowledge and ability, but she’s also very hectic and she

> quickly reaches a certain limit, so let's say she decompensates slightly. Which doesn't really happen at all to me for example, I really tackle things more calmly, she's just a bit (…), she makes a mountain out of a molehill at times, and that of course also makes things less calm sometimes, but it does settle down again, and I don't have that for example. (#17)

As the next quote nicely illustrates, the choice of comparison target is important for the assessment of stressors. Here, the respondent talks about a survey item that gives the respondent the choice between a fictitious coworker A and B:

> So just for example here with this, with the second question, um, I imagined how I am with my assistant, I always have an assistant with me, and he is then more B, and when I'm out with another coworker who has a bit more experience, then I am more B. But most of the time I'm more A. So I had to switch it a bit and then just make a rough estimate of it. (#09)

Comparisons are also made with *people who do the same job but work in another organization.* For instance, #08 (an ambulance driver) talked about swapping shifts and explained:

> But I also know of other stations where that's not possible at all, where you really only work according to the roster, and you also don't have the possibility to change it. Or you just have to give written notice. For us it's completely different for example. So it's completely different depending on where you work I think. So here in the city, they have much, much more call-outs than somewhere in the countryside. (#08)

The comparison targets reported above are close to the respondents because they work with them or belong to the same profession. Other respondents made

comparisons with individuals who perform other jobs, but who are still close to them, in particular individuals from the same company but from other departments. For instance, #19, a lawyer in a municipality, compared himself "with lawyer coworkers, solicitors, judges, other lawyers, what they do."

The third group made comparisons with individuals with *completely different professions* in their assessment. In particular, respondents with office jobs compared themselves with those with typical industrial jobs, where employees are more likely to work on an assembly line or work on a piecework basis. For example, #02 reported: "So I also thought for example about my mother-in-law, well, she works in production." In addition, #07 compared the noise at her workplace with occupational groups that she believes have higher or lower levels of noise pollution, such as stonemasons. We also observed that some respondents used comparisons only for some questions (e.g., #18). In general, it seems that comparisons with other occupations are made when it comes to typical characteristics of a particular occupation that are mentally easy to assess even for someone who is not working in this profession. In addition to the self-selected comparisons, there were also forced comparisons using the A-B formats (e.g., A has to spend a lot of time procuring information, materials, or tools to carry on working. B always has the necessary information, materials or tools available. Which of the two jobs is most similar to your own?). Nearly half of respondents evaluated this format as appropriate and reported that they were able to imagine the two fictitious coworkers well and could place themselves accordingly. Only three interviewees reported difficulties and #24 expressed uncertainty as to whether person A and B remained the same for all A-B questions or should be re-elected each time.

### *Situational Contexts*

Respondents also talked about some general situational context factors; in particular, concern about the *economic situation* of their organization and the associated consequences, such as staff reductions or restructuring. This tense situation in turn biases the response to work stress surveys according to #15, because it can lead to a general atmosphere of fear in an organization:

> At the moment here in our company – we've got a somewhat tense situation, so regarding staff cuts, personnel restructuring etc. And (.) you've definitely always got such things in the back of your mind, haven't you? For the other questions, and, um, I think if we didn't have this situation, it might of course be that I would have tended to answer a bit differently. (#15)

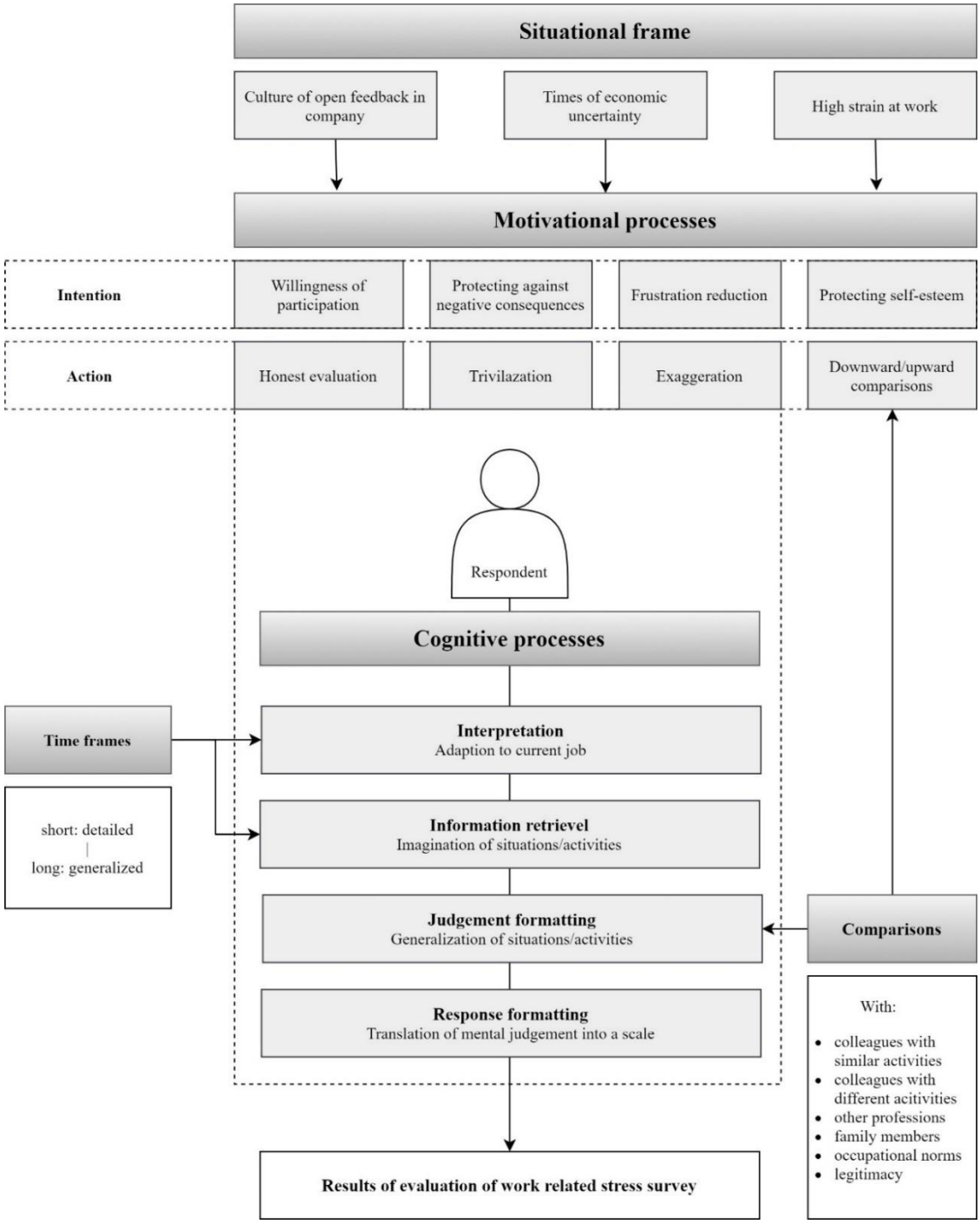
In addition to the current situation of the company, the *respondent's daily mood* might also bias the assessment of stress: "I find to a small percentage it's also dependent on daily mood. So if I had to give a percentage, let's say 5-10%. Because you then simply see certain things differently. I now ticked 3, but 2 or 4, I can imagine the fluctuation." (#15). Furthermore, one's own current stress level in one's private life or one's state of health could also bias how one responds to the questions, as described in this quote:

> […] if I'd just come from work, was completely stressed, then I think the test would have turned out differently, I'm quite sure. Um, if I'd been sick for a long time, something else, I think it would have again been different. (#07)

Furthermore, a *recent conversation* about the stress level in another person's job might also have influenced respondents because they might have wondered more about how stressful the other person's job really is.

**Figure 1**

*Conceptual Model of Response Behavior in Work Stress Surveys*



**Research Question 3: Differences Between Self- and Other-Reporting**

Since stressors and resources are assessed very idiosyncratically, the idea evolved during data collection and analysis that other-report might be different. In fact, we found several differences between self- and other-report regarding motivational and cognitive processes and contexts considered, which are illustrated in Figure 2.

***Motivational Processes***

Both the coworkers and the self-reporters believed that there is a high degree of agreement between them. The supervisors, on the other hand, are of the opinion that employees usually exaggerate and assess their activities as more stressful than they actually are, like #11: "If it is done anonymously, a picture will emerge that everyone is stressed and things can't continue this way". Only about half of supervisors considered the stress assessment to be useful at all. They saw their motivation in relativizing employees' assessments and in making a more objective and therefore, in the supervisors' eyes, more correct assessment, "...because the employee always evaluates him/herself as fundamentally better than he/she actually is, and always deems the situation to be more critical than it actually is" (#11). On the other hand, coworkers were more concerned that they might put their coworker in a negative light, like #29:

> Yes, so I'd find it very hard to answer about a colleague in general. Especially if it also involves personality-related judgments. So if I actually had to write who the rating is about, I'd find it hard, whether I had to respond truthfully depending on who I'm filling the questionnaire in about […] It smacks a bit of somehow confronting this person with it.

In the end, this likely leads in both cases to an underestimation of the stress assessment, but – as highlighted above – due to different reasons.

***Cognitive Processes***

Nearly all third-party assessors reported that tried to be as objective as possible in their assessments. For example, #32 said: "I […] try in any case to answer the things as objectively as possible and independently of the question", and #26: "… I know that I also tried to answer objectively" Moreover, when asked about their approach, other-reporters often stated that they mentally represented the activity (instead of representing the person), like #13:"… I visualized what the coworker's workplace looks like and the tasks he's instructed to do, and how his daily routine proceeds." The possibility of referring more to the activity was not considered by the self-evaluators, as #10 makes clear: "… I only evaluated myself" – self-evaluators just focused on themselves. According to #32, "describing the activity and less so the person" might well help to obtain a more objective assessment.

Coworkers and supervisors reached their responses in different ways. More precisely, coworkers mostly based their assessments on knowledge because they themselves carried out the activity to be assessed. When we asked #17 whether this person was able to assess the stressors and resources of this person's coworker, #17 replied "…yes, because I actually do exactly the same thing". That the assessment can be useful for coworkers working closely together was also confirmed by #29: "So because the work is relatively similar for us overall, I think it makes most sense to consider someone from my department". The fact, however, that the survey is about a coworker helped them to leave out their own subjective distortions, as #26 stated: "I think that it's […] distorting if you describe yourself and therefore in that regard I would, I think actually see it as the best option to describe colleagues in the

immediate environment, because you're in turn best able to assess that". Even in the case of assessments of stressors that were not directly measurable or visible, coworkers stated that they had sufficient knowledge because they were in close contact with their coworkers and regularly exchanged information, like #12: "…because we talk to each other about work", and #25: "…I [have] discussed work with him very often and intensively". Supervisors, in contrast, often arrived at their assessment by observing and interpreting behavior, facial expressions, gestures, and the content of verbal utterances. In a best-case scenario, they used to hold the job themselves, but they often did not perform the same job, which might cause biases (Conway & Lance, 2010). This was also often criticized by the employees themselves – that their supervisors did not have much of an idea how much stress they experienced, like #16: "…because I think he judges us differently as employees because he's the superior and just doesn't have many insights (…)". For example, #14 also assumed that the supervisor considered the job to be "much less stressful or burdensome". Two thirds of the surveyed employees therefore rejected an other-report by their supervisors because, as #29 explained:

> […] of course, he primarily sees work results and not the work processes and the questionnaire, very strongly targets the actual processes and procedures, in the daily work. That is, and this isn't discussed as much, unless there are process changes or something similar. Therefore I think that the superior would guess at a lower level than the persons themselves […]

Uncertainty due to lack of specification of the items also arose in the choice of the perspective to be adopted in the other-report (another link to Research Question 2). For example, coworkers and supervisors were not always sure whether they should

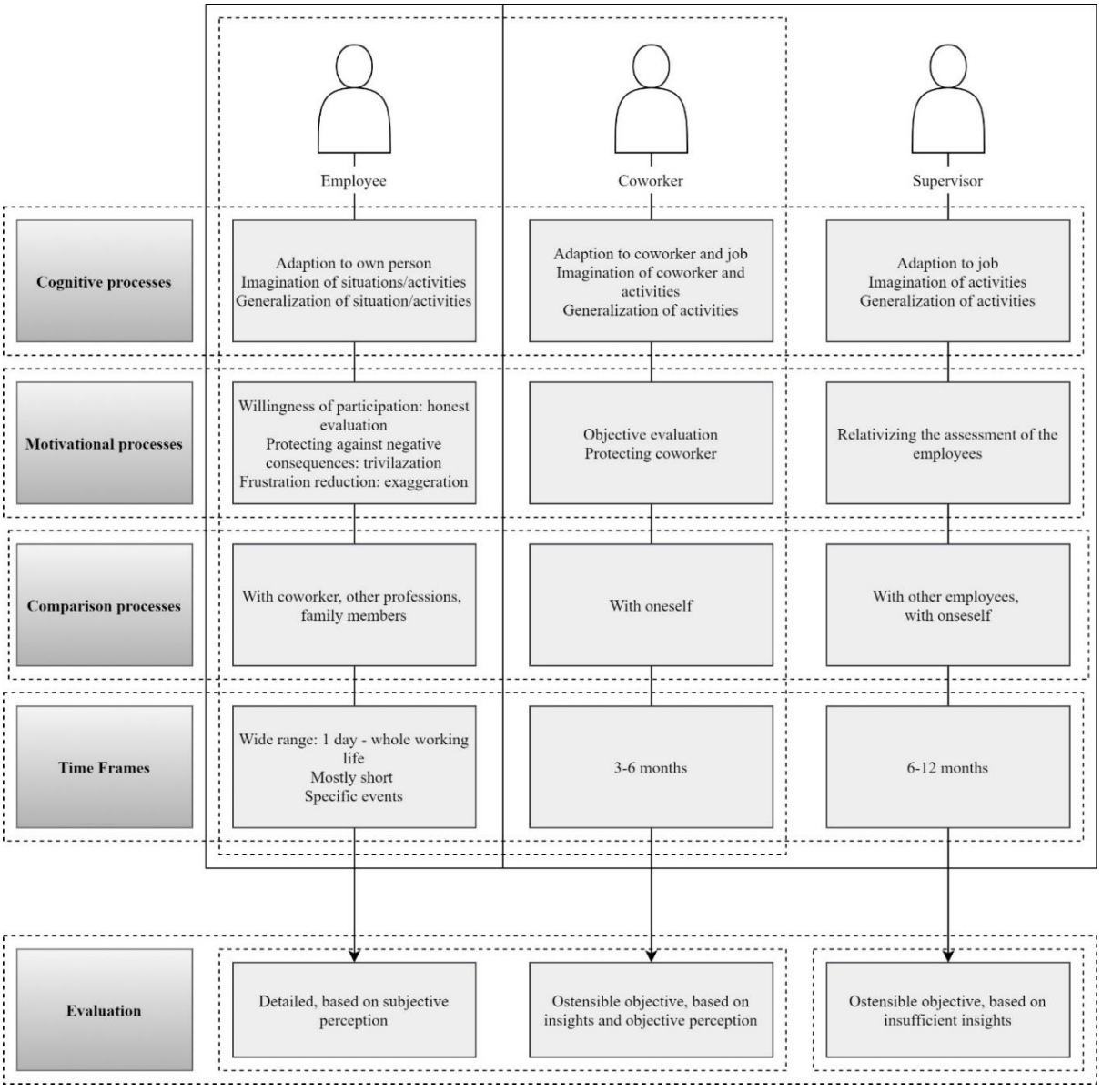
use their own perception or take the subjective view of the person being assessed, such as #24: “That’s also actually what I asked myself when I was answering some of the questions, whether I should assess how the person sees it or how it actually is.” For example, an other-reporter might not evaluate the prevailing noise pollution as stressful, but she or he knows that her or his coworker does. #29 therefore attempted “to see these things from his perspective, and that’s how I then answered it”.

***Effects of Comparisons***

Comparison processes were used more frequently in the other-report condition (about 80%) than in the self-report condition (30%), which links answers to this Research Question to answers to Research Question 2. Coworkers often compared the focal person with themselves; for example, when asked whether they drew comparisons when completing the survey, #29 answered: “Yes, of course, regularly with myself, how I would assess it”. The supervisors, on the one hand, tended to compare with themselves when they had the same job in the past, but mainly compared the employees with each other, like #24: “I did wonder whether maybe the same task area would bring about a different feeling or a different perception in another person”. In part, supervisors chose comparisons, such as #28, “with colleagues who work in the same office….”, but also with employees who are not so close to the person being assessed, such as #33, “between other employees on the hierarchical level above and below.”

**Figure 2**

*Similarities and Differences When Responding to Work Stress Survey Items Among the Three Perspectives of Self-, Supervisor-, and Coworker-Reports*



## Discussion

The aim of this study was to contribute to the understanding of the processes raters use while answering work-related stress surveys and to identify differences between self- and other-reports. We found cognitive processes like interpretation (mental adaptations to current job), information retrieval (imagination of situations/activities), judgment formatting (generalization of situations/activities), and response formatting (translation of the mental assessment into a given answer on a scale). Superordinate to these processes are motivational processes with the intention of willingness to participate, protecting against negative consequences, frustration reduction, and protecting self-esteem. These in turn were found to be influenced by situational frames such as companies' open-feedback cultures, instances of economic uncertainty, and high strain at work. Furthermore, our interviewees reported the use of different time frames and different comparison processes. Other-reporters highlighted their desire to give highly objective ratings. Our participants also assumed that the agreement between colleagues' reports and self-assessments is higher than the agreement between supervisors' reports and self-assessments, likely because colleagues based their assessment on knowledge about the job and the person.

### Theoretical Implications

The cognitive processes reported by the subjects form the core of the developed model. Although these processes partly replicated those of previous studies (e.g., Menon, 1994; Tourangeau, 1984), our rich qualitative data allowed us to develop several new theoretical links and arguments (see Figure 1). First, our study contributes to the understanding of the motivational processes when work stressor surveys are filled out. In our model (Figure 1), the motivational processes are shown as higher priority because they are upstream of, and seem to have a considerable

influence on, the cognitive processes. We subdivided them into a motivational intention followed by a certain behavior (action). Some participants reported that they feared negative consequences from their supervisors; therefore, they had the intention to protect themselves against these negative consequences by deliberately *understating* or *trivializing* their responses. Others reported the motivation to reduce their frustration and therefore *exaggerated* their responses. Employees seem to anticipate their supervisors' or management's reactions in advance and adjust their response behavior accordingly in order to create the right impression. This suggests that impression management theory (Leary & Kowalski, 1990), which is, for instance, used to explain faking in personnel selection processes when applicants try to present themselves as better than they really are (e.g., Birkeland, Manson, Kisamore, Brannick, & Smith, 2006), might also be relevant for stress surveys. Some survey participants might hope to avoid negative consequences and others might hope to achieve a positive outcome, namely better working conditions. In this regard, perceived anonymity (Dunnette & Heneman, 1956) seems to be a crucial factor, but if there is no confidence in the trustworthiness of supervisors or top management, participants might believe that exaggeration represents the only means to express their dissatisfaction.

Second, our data expand knowledge on the importance of contexts for answering survey items. To date, contextualization has mainly been investigated in the context of personality surveys (e.g., Lievens, de Corte, & Schollaert, 2008; Schmit, Ryan, Stierwalt, & Powell, 1995). Our analyses provide further indications that other contexts matter for work stressor surveys, in particular time frames and social comparisons (see Figure 1). The range of time frames was rather high: from one day to the rater's whole work life. Presumably, the use of time frames plays a role mainly

in the interpretation of items and in information retrieval. The shorter the time frame, the more detailed the memories and the resulting evaluation might be. With broader time frames, it seems to be easier for respondents to think in a more generalized way when interpreting the items and retrieving information. Therefore, respondents' answers might reflect cumulative effects of stressors instead of the effects of particular events (see Igic et al., 2017). This contextualization of time might not only change ratings between participants, but also within an individual participant. As such, the variation of time periods used by participants when thinking about their jobs would be a form of contextualization that may bias the response process and therefore lead to unreliable results (Debus et al., 2019).

Another form of contextualization is using a social context: Our interviewees reported that they had compared themselves with, and made their assessments depending on, other people or groups. Other-reporters seemed to particularly use comparisons for judgment and answer formatting. Comparisons are used to better assess and evaluate previously remembered situations and activities. Coworkers often used themselves for comparison and supervisors frequently compared their employees. Comparisons with others might thus help respondents to assess the extent to which they have the resources necessary to perform a task. Despite the importance of these comparison processes in our study, such comparison processes are rarely taken into account by stress theories. In fact, most stress theories (e.g., Lazarus & Folkman, 1987) assume that respondents' assessments are largely independent of the level of stressors of others. The present results (and some others, e.g., Ford & Jin, 2015; Greenberg, Ashton-James, & Ashkanasy, 2007) suggest that these theories need to be supplemented by comparison processes. For example, if reactions to a stressor depend on the legitimacy of this stressor (as the "stress-as-offence-to-self" perspective

argues; Semmer, Jacobshagen, Meier, & Elfering, 2007), then this legitimacy perception is likely also influenced by comparing oneself to colleagues. To understand these comparison processes, it might also be helpful to combine stress theories with fairness theories, because comparison processes play an import role in fairness theories (Greenberg et al., 2007).

Such comparisons also have motivational effects because social comparisons can be used for a downward comparison: If people like to enhance their self-esteem by showing how well they can do something compared to others, they engage in downward comparisons. In the context of stress this means that employees compare themselves to others who experience, for instance, more time pressure, and such downward comparisons will likely have a protective effect for employees. However, upward comparison processes can also occur – if employees compare themselves to others whose working conditions they consider worse, they might take stressors more lightly. Thus, depending on the context in which stressors occur, stressors might be assessed as less or more severe. Unfortunately, it appears to be difficult to predict whether respondents report higher or lower stress levels when they compare themselves to others.

Third, there are situational frames that influence motivational processes. Especially if an organization is in an economically bad situation, a positive evaluation by others is a strong motive for most people (e.g., Sedikides & Strube, 1997): They strive to protect their self-worth, which refers to the regard or respect received from others (e.g., Leary & Baumeister, 2000). Additionally, other situational frames seem to have an influence on motivational processes, such as an open-feedback culture, times of economic uncertainty, and high strain at work. Interviewees reported that the existence of an open-feedback culture in their company raises the willingness for

participation – participating in such a survey has meaning and can also change something for the better, resulting in an honest assessment of the work-related stress conditions. This is different in times of economic uncertainty. Employees are more likely to be afraid of losing their jobs if they honestly describe how stressed and strained they are. To protect themselves from these negative consequences, they might trivialize the stressful working conditions and their own well-being. If employees feel safe and have high stress at work, interviewees described that they would vent their anger with such a survey and exaggerate the working conditions by describing them as more stressful and burdensome than they actually are.

What is more, our results also shed light on the processes that are involved in other-reports of work-related stress (supervisors and coworkers, see Figure 2). Although the assessment of stress is very idiosyncratic due to the assessment of available resources and stressors (Lazarus & Folkman, 1987; Semmer & Meier, 2009), finding a more objective measurement thus seems attractive (e.g., through other-reports), but incumbents and their supervisors seem to differ considerably in their judgments (Debus et al., 2015; Meier & Semmer, 2018). The low convergence of these different perspectives might be at least partially explained by our findings, which reveal how different the processes are in the minds of supervisors, coworkers, and employees who describe their own working conditions. Our results indicate (as interviewees also suspected) that supervisors tend to estimate the levels of stress as lower than the employees themselves because supervisors typically do not have in-depth understanding of the employee's tasks and processes.

Furthermore, our results seem consistent with a role-theoretic explanation (see Sluss, van Dick, & Thompson, 2011). Role theory is reflected and thus supported by

two elements found in the answers of our interviewees: the scope of the supervisors' role responsibilities and the role distance between supervisors and their subordinates.

At the same time, our data reveal that coworker reports seem to function differently. In fact, our findings suggest a higher convergence between coworker reports and self-reports than between supervisor reports and self-reports. Coworkers are often close to each other because they have the same job or at least similar jobs. Unlike supervisors, coworkers often have more knowledge about a workplace and potentially also about the feelings of the focal person. Coworkers also have a smaller area of responsibility than supervisors and might thus more easily concentrate on one person when filling out such a stress survey. Finally, they do not have to fear negative consequences for themselves (their reports might deviate from the truth only if they are motivated to protect coworkers from negative consequences). According to role theory (e.g., Sluss, van Dick, & Thompson, 2011), coworkers and the focal employee occupy the same or at least similar roles, which should lead to a similar view, in contrast to the role of the supervisor.

Supervisors and coworkers sometimes wondered whether they should base their answers on their own perception or the perception of the person being assessed. For example, other-reporters might not find the prevailing noise stressful, but they might know that their coworker finds this stressful. Although the adoption of one particular perspective could increase the agreement, it remains to be seen whether this results in an objectively more correct assessment of the workplace situation. This raises the general question of what exactly is the goal of the stress survey – whether it is about assessing (a) the individual, subjective level of stress of an employee, or assessing (b) objective working conditions (Semmer, Zapf, & Greif, 1996).

**Practical Implications**

The main implication of this study is that researchers and practitioners should be aware of the complexity of respondents' motivational and cognitive processes, as well as the variability in the context that raters consider while completing work stress surveys. In particular, they should recognize the differences between self- and other-reports, and within other-reports, the differences between reports made by supervisors or by coworkers. Practitioners (and researchers) should also be aware that stress items are often formulated too generally and therefore leave respondents with too much scope for interpretation, which means, as our results show, that respondents may use it in a way that increases inter- and intra-individual variability and thus could distort results. Therefore, we recommend that practitioners contextualize items by specifying the temporal frames of reference. So far, some surveys already offer a timescale to frame the responses, and this use of contextualization should be further pursued and expanded. In particular, we recommend contextualization by specifying comparisons to unify response behavior.

Furthermore, items in work-related stress surveys should be formulated in a way that clarifies if it is the activity in general that is to be evaluated or if items are about individual perceptions of stressors and resources. This could also be specified in the instructions. In addition, each item could be tagged with a corresponding instruction to calculate an average or an individual value (e.g., "in general" or "in this moment/today"). Potentially, instructions could instruct respondents not to include extraordinary events to avoid distorted ratings. In the case of an other-report, it also seems advisable to explain the perspective that should be adopted: whether other-reporters should give the view of the focal employees or their own assessment.

In addition, practitioners should be aware of the influences by motivational processes. Our results imply that organizations should inform the respondents about goals and intentions before the survey as much as possible. A great deal of effort should be made to ensure that respondents perceive anonymity and trust the persons who work with the data (and who ideally do not work within organizations and who certainly should not be the supervisors themselves).

**Avenues for Future Research**

Due to the qualitative setting, this study is based on a small number of people. Therefore, quantitative research will be needed to consolidate the findings of this study. Furthermore, this study opens several new avenues for future research. First, future research could focus on the biasing effect of different time frames and social comparisons. In particular, experimental manipulations with several conditions are conceivable. Potentially, respondents remember more details and include current events in their assessment, drawing on shorter periods of time, which might therefore not describe how their work typically is. Furthermore, participants reported that they compared themselves with different individuals including coworkers, other people in the same profession, or people in other professions. These comparison processes could also be investigated using experiments in which researchers vary the targets of comparison and study the effects on reliability and validity. Moreover, previous studies suggest that people make downward or upward comparisons depending on whether they want to motivate themselves or protect their self-esteem (Major et al., 1991; Taylor & Lobel, 1989). In order to understand such upward and downward comparisons in the context of work-related stress, it seems important to understand whether experiencing work stressors such as time pressure is considered as a sign of success or of failure (e.g., of having poor time management).

Second, the ISTA also includes some questions in A-B format, in which respondents are presented with two hypothetical people (A and B) and asked to state similarities with either one (see Table 3). The advantages and disadvantages of this format are unknown – perhaps such a forced comparison facilitates the responding processes. Maybe some respondents find it difficult to assign themselves to either of the two options, which implies that such items do not contribute much to a meaningful assessment. Thus, a closer examination of the relationship between certain types of questions and the chosen comparison processes could be a worthwhile idea for future qualitative and quantitative research.

Third, future studies could take a closer look at motivational processes: distortions due to the desire for change (i.e., exaggeration) vs. due to the fear of negative consequences (i.e., understatement). It seems relevant to determine whether employees' frustration leads to strong exaggeration, maybe because this allows them to vent their anger and to draw an organization's attention to the need for change. Alternatively, exaggeration might also be an expression of revenge because high stress levels in a team may be seen as a poor result for the team manager and thus might have negative consequences for the manager (then this could be considered as a new form of counterproductive work behavior; Mercado, Dilchert, Giordano, & Ones, 2018). In addition, anonymity seems to play an important role in distortion due to fear. Even if anonymity is officially assured, participants' perception of it might differ because they might feel easily identifiable and may therefore not respond honestly to issues that need improvement.

Fourth, there are several ways in which future research may build upon the different biases identified in this study. Although our results highlight the existence of different biases, we do not yet know about the co-occurrences of such biases. Put

differently, there might be subpopulations of participants that differ in the combinations of biases that affect their response behavior. Applying such a person-centered approach (Wang et al., 2013; Wang & Bodner, 2007) that identifies subpopulations that differ in the configuration of biases appears to be a helpful next step in order to gain a better understanding of individuals' response behavior. Such an approach would also allow predictors (e.g., demographics, personality traits, occupational groups) of such response behavior profiles to be identified, thus allowing researchers and practitioners to take preventative steps against the occurrence of certain biases when designing questionnaires. Furthermore, we focused on the cross-sectional survey in our research because this has traditionally been the most prominent method for data collection in stress research (e.g., Chan, 2009), but with the advancement of measurement-intensive methods (e.g., diary methods and event-sampling approaches; e.g., Gabriel et al., 2018), future research may also elucidate the generalizability of our findings to these methods. More specifically, situational biases (such as through a person's current affect) and biases through comparisons may play a greater role when people are repeatedly asked about their work situation. People might also be more likely to apply intra-individual comparisons due to the fact that they are repeatedly asked about their situation.

**Potential Limitations**

As with all studies, the present study has its limitations. First, all items were taken from one instrument, namely the ISTA (Semmer et al., 1999), which might limit the generalizability of our results. We had originally chosen this instrument because the ISTA is a well-validated instrument that is used very often in German-speaking countries (e.g., Silla & Gamero, 2014; Sonnentag & Zijlstra, 2006). Moreover, many ISTA items show great overlap with other instruments that measure stressors and

resources at work (e.g., the Copenhagen Psychosocial Questionnaire [COPSOQ; Kristensen et al., 2005], the Work Design Questionnaire [WDQ; Morgeson & Humphrey, 2006]). Nevertheless, we cannot exclude the possibility that the use of other instruments might have resulted in additional insights. Second, it should be kept in mind that this study focused on general stress surveys (i.e., on surveys that inquire how stressful a particular job is in general) and not about stress surveys that can be found in diary and event-sampling studies (see, e.g., Gabriel et al., 2019). Although the latter already contain time frames (e.g., "Today, I was required to work fast at work" in a bedtime survey; e.g., Mojza & Sonnentag, 2010), which make the interpretation of such items easier, answering might still be perceived as difficult. Third, our interviewees came only from one country (i.e., Germany). Although stress in the workplace is a global phenomenon (Spector et al., 2002), reactions to stressors vary between countries (e.g., Debus, Probst, König, & Kleinmann, 2012). This highlights the need for replication studies in other countries, ideally countries that are not WEIRD (Western, Educated, Industrialized, Rich, and Democratic; Henrich, Heine, & Norenzayan, 2010). Fourth, grounded theory (Glaser & Strauss, 1967) approaches have the ability to adjust the data collection according to previous results, and readers should be aware of the advantages and disadvantages of this. The advantage of such an approach lies in the flexibility that it offers to researchers (e.g., Holloway & Todres, 2003), but the disadvantage lies in the fact that changing the procedure during the data collection not only stands in contrast to most quantitative approaches, but also makes it more difficult to replicate qualitative studies.

**Conclusion**

In summary, this qualitative study contributes to a better understanding of response behavior when administering work stress surveys. Not only did our analyses

reveal a general approach when responding to such items, but they also uncovered several biases that can change the assessment of work-related stressors and resources (e.g., different comparison targets, different time frames). Furthermore, we also found initial indications of motivational and situational influencing factors that might potentially bias survey items. We hope that future research will provide further confirmation and refinement of our results, improving the theoretical understanding of the response process and ultimately helping practitioners to obtain reliable survey results.

**References**

Aguinis, H., & Solarino, A. M. (2019). Transparency and replicability in qualitative research: The case of interviews with elite informants. *Strategic Management Journal*, *40*(8), 1291–1315. https://doi.org/10.1002/smj.3015

Berry, C. M., Carpenter, N. C., & Barratt, C. L. (2012). Do other-reports of counterproductive work behavior provide an incremental contribution over self-reports? A meta-analytic comparison. *Journal of Applied Psychology*, *97*(3), 613–636. https://doi.org/10.1037/a0026739

Berry, C. M., Ones, D. S., & Sackett, P. R. (2007). Interpersonal deviance, organizational deviance, and their common correlates: A review and meta-analysis. *Journal of Applied Psychology*, *92*(2), 410–424. https://doi.org/10.1037/0021-9010.92.2.410

Birkeland, S. A., Manson, T. M., Kisamore, J. L., Brannick, M. T., & Smith, M. A. (2006). A meta-analytic investigation of job applicant faking on personality measures. *International Journal of Selection and Assessment*, *14*(4), 317–335. https://doi.org/10.1111/j.1468-2389.2006.00354.x

Blair, E., & Burton, S. (1987). Cognitive processes used by survey respondents to answer behavioral frequency questions. *Journal of Consumer Research*, *14*(2), 280–288. https://doi.org/10.1086/209112

Blaszczynski, A., Dumlao, V., & Lange, M. (1997). "How much do you spend gambling?" Ambiguities in survey questionnaire items. *Journal of Gambling Studies*, *13*, 237–252. https://doi.org/10.1023/A:1024931316358

Bluhm, D. J., Harman, W., Lee, T. W., & Mitchell, T. R. (2011). Qualitative research in management: A decade of progress. *Journal of Management Studies*, *48*(8), 1866–1891. https://doi.org/10.1111/j.1467-6486.2010.00972.x

Bolino, M. C., Kacmar, K. M., Turnley, W. H., & Gilstrap, J. B. (2008). A multi-level review of impression management motives and behaviors. *Journal of Management*, *34*(6), 1080–1109. https://doi.org/10.1177/0149206308324325

Bradburn, N., Rips, L., & Shevell, S. (1987). Answering autobiographical questions: The impact of memory and inference on surveys. *Science*, *236*(4798), 157–161. https://doi.org/10.1126/science.3563494

Byrne, M. (2001). Sampling for qualitative research. *AORN Journal*, *73*(2), 494–498. https://doi.org/10.1016/S0001-2092(06)61990-X

Chan, D. (2009). So why ask me? Are self-report data really that bad? In C. E. Lance & R. Vandenberg (Eds.), *Statistical and methodological myths and urban legends: Doctrine, verity and fable in the organizational and social sciences* (pp. 309–336). Routledge.

Conway, J. M., & Lance, C. E. (2010). What reviewers should expect from authors regarding common method bias in organizational research. *Journal of Business and Psychology*, *25*(3), 325–334. https://doi.org/10.1007/s10869-010-9181-6

Credé, M., Bashshur, M., & Niehorster, S. (2010). Reference group effects in the measurement of personality and attitudes. *Journal of Personality Assessment*, *92*(5), 390–399. https://doi.org/10.1080/00223891.2010.497393

de Jonge, J., & Dormann, C. (2006). Stressors, resources, and strain at work: A longitudinal test of the triple-match principle. *Journal of Applied Psychology*, *91*(6), 1359–1374. https://doi.org/10.1037/0021-9010.91.5.1359

Debus, M. E., Greulich, B., König, C. J., & Kleinmann, M. (2019). Insecure about how to rate your job insecurity? A two-study investigation into time frames

applied to job insecurity measures. *Occupational Health Science*, *3*(4), 421–435. https://doi.org/10.1007/s41542-019-00049-x

Debus, M. E., König, C. J., Kleinmann, M., & Werner, C. S. (2015). Examining the effects of negative affectivity on self- and supervisor ratings of job stressors: The role of stressor observability. *Work & Stress*, *29*(4), 341–361. https://doi.org/10.1080/02678373.2015.1075233

Debus, M. E., Probst, T. M., König, C. J., & Kleinmann, M. (2012). Catch me if I fall! Enacted uncertainty avoidance and the social safety net as country-level moderators in the job insecurity–job attitudes link. *Journal of Applied Psychology*, *97*(3), 690–698. https://doi.org/10.1037/a0027832

Debus, M. E., Sonnentag, S., Deutsch, W., & Nussbeck, F. W. (2014). Making flow happen: The effects of being recovered on work-related flow between and within days. *Journal of Applied Psychology*, *99*(4), 713–722. https://doi.org/10.1037/a0035881

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, *130*(3), 355–391. https://doi.org/10.1037/0033-2909.130.3.355

Dunnette, M. D., & Heneman, H. G., Jr. (1956). Influence of scale administrator on employee attitude responses. *Journal of Applied Psychology*, *40*(2), 73–77. https://doi.org/10.1037/h0040428

Fida, R., Paciello, M., Tramontano, C., Fontaine, R. G., Barbaranelli, C., & Farnese, M. L. (2015). An integrative approach to understanding counterproductive work behavior: The roles of stressors, negative emotions, and moral disengagement. *Journal of Business Ethics*, *130*(1), 131–144. https://doi.org/10.1007/s10551-014-2209-5

Fleiss, J. L., & Cohen, J. (1973). The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement*, *33*(3), 613–619. https://doi.org/10.1177/001316447303300309

Flick, U. (2018). *Designing Qualitative Research*. SAGE.

Ford, M. T., & Jin, J. (2015). Incongruence between workload and occupational norms for time pressure predicts depressive symptoms. *European Journal of Work and Organizational Psychology*, *24*(1), 88–100. https://doi.org/10.1080/1359432X.2013.858701

Frese, M., Kring, W., Soose, A., & Zempel, J. (1996). Personal initiative at work: Differences between East and West Germany. *Academy of Management Journal*, *39*(1), 37–63. https://doi.org/10.5465/256630

Fusch, P., & Ness, L. (2015). Are we there yet? Data saturation in qualitative research. *The Qualitative Report*, *20*(9), 1408–1416. https://doi.org/10.46743/2160-3715/2015.2281

Gabriel, A. S., Podsakoff, N. P., Beal, D. J., Scott, B. A., Sonnentag, S., Trougakos, J. P., & Butts, M. M. (2018). Experience sampling methods: A discussion of critical trends and considerations for scholarly advancement. *Organizational Research Methods*, *22*, 969–1006. https://doi.org/10.1177/1094428118802626

Gilboa, S., Shirom, A., Fried, Y., & Cooper, C. (2008). A meta-analysis of work demand stressors and job performance: Examining main and moderating effects. *Personnel Psychology*, *61*(2), 227–271. https://doi.org/10.1111/j.1744-6570.2008.00113.x

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Aldine.

Greenberg, J., Ashton-James, C. E., & Ashkanasy, N. M. (2007). Social comparison processes in organizations. *Organizational Behavior and Human Decision Processes*, *102*(1), 22–41. https://doi.org/10.1016/j.obhdp.2006.09.006

Hassard, J., Teoh, K. R. H., Visockaite, G., Dewe, P., & Cox, T. (2018). The cost of work-related stress to society: A systematic review. *Journal of Occupational Health Psychology*, *23*(1), 1–17. https://doi.org/10.1037/ocp0000069

Heine, S. J., Buchtel, E. E., & Norenzayan, A. (2008). What do cross-national comparisons of personality traits tell us? The case of conscientiousness. *Psychological Science*, *19*(4), 309–313. https://doi.org/10.1111/j.1467-9280.2008.02085.x

Heine, S. J., Lehman, D. R., Peng, K., & Greenholtz, J. (2002). What's wrong with cross-cultural comparisons of subjective Likert scales? The reference-group effect. *Journal of Personality and Social Psychology*, *82*(6), 903–918. https://doi.org/10.1037//0022-3514.82.6.903

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Holloway, I., & Todres, L. (2003). The status of method: Flexibility, consistency and coherence. *Qualitative Research*, *3*(3), 345–357. https://doi.org/10.1177/1468794103033004

Holtrop, D., Born, M. Ph., de Vries, A., & de Vries, R. E. (2014). A matter of context: A comparison of two types of contextualized personality measures. *Personality and Individual Differences*, *68*, 234–240. https://doi.org/10.1016/j.paid.2014.04.029

Igic, I., Keller, A. C., Elfering, A., Tschan, F., Kälin, W., & Semmer, N. K. (2017). Ten-year trajectories of stressors and resources at work: Cumulative and chronic effects on health and well-being. *Journal of Applied Psychology*, *102*(9), 1317–1343. https://doi.org/10.1037/apl0000225

Irmer, J. P., Kern, M., Schermelleh-Engel, K., Semmer, N. K., & Zapf, D. (2019). The instrument for stress-oriented task analysis (ISTA): A meta-analysis. *Zeitschrift für Arbeits- und Organisationspsychologie*, *63*(4), 217–237. https://doi.org/10.1026/0932-4089/a000312

Karasek, R., Brisson, C., Kawakami, N., Houtman, I., Bongers, P., & Amick, B. (1998). The job content questionnaire (JCQ): An instrument for internationally comparative assessments of psychosocial job characteristics. *Journal of Occupational Health Psychology*, *3*(4), 322–355. https://doi.org/10.1037/1076-8998.3.4.322

Kreiner, G. E., Hollensbe, E. C., & Sheep, M. L. (2009). Balancing borders and bridges: Negotiating the work-home interface via boundary work tactics. *Academy of Management Journal*, *52*(4), 704–730. https://doi.org/10.5465/amj.2009.43669916

Kristensen, T. S., Borritz, M., Villadsen, E., & Christensen, K. B. (2005). The Copenhagen Burnout Inventory: A new tool for the assessment of burnout. *Work & Stress*, *19*(3), 192–207. https://doi.org/10.1080/02678370500297720

Lazarus, R. S., & Folkman, S. (1987). Transactional theory and research on emotions and coping. *European Journal of Personality*, *1*(3), 141–169. https://doi.org/10.1002/per.2410010304

Leary, M. R., & Baumeister, R. F. (2000). The nature and function of self-esteem: Sociometer theory. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 32, pp. 1–62). Academic Press.

Leary, M. R., & Kowalski, R. M. (1990). Impression management: A literature review and two-component model. *Psychological Bulletin*, *107*(1), 34–47. https://doi.org/10.1037/0033-2909.107.1.34

Lewis, S. (2015). Qualitative inquiry and research design: Choosing among five approaches. *Health Promotion Practice*, *16*(4), 473–475. https://doi.org/10.1177/1524839915580941

Lievens, F., de Corte, W., & Schollaert, E. (2008). A closer look at the frame-of-reference effect in personality scale scores and validity. *Journal of Applied Psychology*, *93*(2), 268–279. https://doi.org/10.1037/0021-9010.93.2.268

Little, L. M., Major, V. S., Hinojosa, A. S., & Nelson, D. L. (2014). Professional image maintenance: How women navigate pregnancy in the workplace. *Academy of Management Journal*, *58*(1), 8–37. https://doi.org/10.5465/amj.2013.0599

Locke, K. (2001). *Grounded theory in management research*. Sage.

Major, B., Testa, M., & Blysma, W. H. (1991). Responses to upward and downward social comparisons: The impact of esteem-relevance and perceived control. In J. Suls & T. A. Wills (Eds.), *Social comparison: Contemporary theory and research* (pp. 237–260). Erlbaum.

Mazzola, J. J., Schonfeld, I. S., & Spector, P. E. (2011). What qualitative research has taught us about occupational stress. *Stress and Health*, *27*(2), 93–110. https://doi.org/10.1002/smi.1386

Meier, L. L., & Semmer, N. K. (2018). Illegitimate tasks as assessed by incumbents and supervisors: Converging only modestly but predicting strain as assessed by incumbents, supervisors, and partners. *European Journal of Work and Organizational Psychology*, *27*(6), 764–776. https://doi.org/10.1080/1359432X.2018.1526785

Menon, G. (1994). Judgments of behavioral frequencies: Memory search and retrieval strategies. In N. Schwarz & S. Sudman (Eds.), *Autobiographical memory and the validity of retrospective reports* (pp. 161–172). Springer.

Mercado, B. K., Dilchert, S., Giordano, C., & Ones, D. S. (2018). Counterproductive work behaviors. In D. S. Ones, N. Anderson, C. Viswesvaran, & H. K. Sinangil (Eds.), *The SAGE handbook of industrial, work & organizational psychology: Personnel psychology and employee performance* (2nd ed., Vol. 1, pp. 109–211). Sage Reference.

Misoch, S. (2019). *Qualitative Interviews [Qualitative interviews]*. De Gruyter Oldenburg. https://doi.org/10.1515/9783110545982

Mojza, E. J., & Sonnentag, S. (2010). Does volunteer work during leisure time buffer negative effects of job stressors? A diary study. *European Journal of Work and Organizational Psychology*, *19*(2), 231–252. https://doi.org/10.1080/13594320902986097

Morgeson, F. P., & Campion, M. A. (1997). Social and cognitive sources of potential inaccuracy in job analysis. *Journal of Applied Psychology*, *82*(5), 627–655. https://doi.org/10.1037//0021-9010.82.5.627

Morgeson, F. P., & Campion, M. A. (2012). A framework of potential sources of inaccuracy in job analysis. In M. A. Wilson (Ed.), *The handbook of work*

*analysis: The methods, systems, applications and science of work measurement in organizations* (pp. 593–601). Routledge.

Morgeson, F. P., & Humphrey, S. E. (2006). The Work Design Questionnaire (WDQ): Developing and validating a comprehensive measure for assessing job design and the nature of work. *Journal of Applied Psychology*, *91*(6), 1321–1339. https://doi.org/10.1037/0021-9010.91.6.1321

Morrel-Samuels, P. (2002, February). Getting the truth into workplace surveys. *Harvard Business Review*, *80*(2), 111–118.

Moser, C. A., & Kalton, G. (2017). *Survey methods in social investigation* (2nd ed.). Routledge. https://doi.org/10.4324/9781315241999

Nelson, J. (2017). Using conceptual depth criteria: Addressing the challenge of reaching saturation in qualitative research. *Qualitative Research*, *17*(5), 554–570. https://doi.org/10.1177/1468794116679873

Nielsen, K., Nielsen, M. B., Ogbonnaya, C., Känsälä, M., Saari, E., & Isaksson, K. (2017). Workplace resources to improve both employee well-being and performance: A systematic review and meta-analysis. *Work & Stress*, *31*(2), 101–120. https://doi.org/10.1080/02678373.2017.1304463

Perrewé, P. L., & Zellars, K. L. (1999). An examination of attributions and emotions in the transactional approach to the organizational stress process. *Journal of Organizational Behavior*, *20*(5), 739–752. https://doi.org/10.1002/(SICI)1099-1379(199909)20:5<739::AID-JOB1949>3.0.CO;2-C

Pond, S. B., Armenakis, A. A., & Green, S. B. (1984). The importance of employee expectations in organizational diagnosis. *The Journal of Applied Behavioral Science*, *20*(2), 167–180. https://doi.org/10.1177/002188638402000207

Powell, T. J., Enright, S. J., & Enright, S. J. (2015). *Anxiety and stress management*. Routledge.

Ryzin, G. G. V. (2013). An experimental test of the expectancy-disconfirmation theory of citizen satisfaction. *Journal of Policy Analysis and Management*, *32*(3), 597–614. https://doi.org/10.1002/pam.21702

Schaufeli, W. B., Bakker, A. B., & Rhenen, W. V. (2009). How changes in job demands and resources predict burnout, work engagement, and sickness absenteeism. *Journal of Organizational Behavior*, *30*(7), 893–917. https://doi.org/10.1002/job.595

Schmit, M. J., Ryan, A. M., Stierwalt, S. L., & Powell, A. B. (1995). Frame-of-reference effects on personality scale scores and criterion-related validity. *Journal of Applied Psychology*, *80*(5), 607–620. https://doi.org/10.1037/0021-9010.80.5.607

Schmitt, A., Ohly, S., & Kleespies, N. (2015). Time pressure promotes work engagement: Test of illegitimate tasks as boundary condition. *Journal of Personnel Psychology*, *14*(1), 28–36. https://doi.org/10.1027/1866-5888/a000119

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, *54*, 93–105. https://doi.org/10.1037/0003-066X.54.2.93

Schwarz, N. (2007). Cognitive aspects of survey methodology. *Applied Cognitive Psychology*, *21*(2), 277–287. https://doi.org/10.1002/acp.1340

Sedikides, C., & Strube, M. J. (1997). Self evaluation: To thine own self be good, to thine own self be sure, to thine own self be true, and to thine own self be better. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology*

(Vol. 29, pp. 209–269). Academic Press. https://doi.org/10.1016/S0065-2601(08)60018-0

Semmer, N. K., Grebner, S., & Elfering, A. (2003). Beyond self-report: Using observational, physiological, and situation-based measures in research on occupational stress. In P. L. Perrewé & D. C. Ganster (Eds.), *Research in occupational stress and well-being* (Vol. 3, pp. 205–263). Emerald.

Semmer, N. K., Jacobshagen, N., Meier, L. L., & Elfering, A. (2007). Occupational stress research: The "stress-as-offence-to-self" perspective. In J. Houdmont & S. McIntyre (Eds.), *Occupational health psychology: European perspectives on research, and practices* (Vol. 2, pp. 42–66). ISMAI.

Semmer, N. K., Jacobshagen, N., Meier, L. L., Elfering, A., Beehr, T. A., Kälin, W., & Tschan, F. (2015). Illegitimate tasks as a source of work stress. *Work & Stress*, *29*(1), 32–56. https://doi.org/10.1080/02678373.2014.1003996

Semmer, N. K., & Meier, L. L. (2009). Individual differences, work stress and health. In M. J. Schabracq, J. A. M. Winnubst, & C. L. Cooper (Eds.), *International handbook of work and health psychology* (3rd ed., pp. 99–121). Wiley-Blackwell.

Semmer, N. K., Zapf, D., & Dunckel, H. (1999). Instrument zur stressbezogenen Tätigkeitsanalyse ISTA [Instrument for Stress-Oriented Task Analysis ISTA]. In H. Dunckel (Ed.), *Handbuch psychologischer Arbeitsanalyseverfahren* (pp. 179–204). vdf Hochschulverlag.

Semmer, N. K., Zapf, D., & Greif, S. (1996). 'Shared job strain': A new approach for assessing the validity of job stress measurements. *Journal of Occupational and Organizational Psychology*, *69*(3), 293–310. https://doi.org/10.1111/j.2044-8325.1996.tb00616.x

Silla, I., & Gamero, N. (2014). Shared time pressure at work and its health-related outcomes: Job satisfaction as a mediator. *European Journal of Work and Organizational Psychology*, *23*(3), 405–418. https://doi.org/10.1080/1359432X.2012.752898

Singer, E. (2001). The use of incentives to reduce nonresponse in household surveys. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. A. Little (Eds.), *Survey Nonresponse* (p. 520). Wiley & Sons Ltd.

Sluss, D. M., van Dick, R., & Thompson, B. S. (2011). Role theory in organizations: A relational perspective. In S. Zedeck (Ed.), *APA handbook of industrial and organizational psychology, Vol 1: Building and developing the organization* (pp. 505–534). American Psychological Association.

Sonnentag, S., Binnewies, C., & Mojza, E. J. (2010). Staying well and engaged when demands are high: The role of psychological detachment. *Journal of Applied Psychology*, *95*(5), 965–976. https://doi.org/10.1037/a0020032

Sonnentag, S., & Zijlstra, F. R. H. (2006). Job characteristics and off-job activities as predictors of need for recovery, well-being, and fatigue. *Journal of Applied Psychology*, *91*(2), 330–350. https://doi.org/10.1037/0021-9010.91.2.330

Spector, P. E., Chen, P. Y., & O'Connell, B. J. (2000). A longitudinal study of relations between job stressors and job strains while controlling for prior negative affectivity and strains. *Journal of Applied Psychology*, *85*(2), 211–218. https://doi.org/10.1037/0021-9010.85.2.211

Spector, P. E., Cooper, C. L., Lu, L., Siu, O. L., Spector, P. E., Cooper, C. L., Sanchez, J. I., Sparks, K., Siu, O., Stora, J. B., Etudes, H., & Groupe, C. (2002). Locus of control and well-being at work: How generalizable are western findings. *Academy of Management Journal*, 453–466. https://doi.org/10.5465/3069359

Spector, P. E., & Eatough, E. (2013). Quantitative self-report methods in occupational health psychology research. In M. Wang, S. Sinclair, & L. E. Tetrick (Eds.), *Research methods in occupational health psychology* (pp. 248–267). Taylor & Francis.

Strauss, A. L., & Corbin, J. (1990). *Basisc of qualitative research: Techniques and procedures for devoloping grounded theory*. Sage.

Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking about answers: The application of cognitive processes to survey methodology*. Jossey-Bass.

Taylor, S. E., & Lobel, M. (1989). Social comparison activity under threat: Downward evaluation and upward contacts. *Psychological Review*, *96*(4), 569–575. https://doi.org/10.1037/0033-295X.96.4.569

Theorell, T., Hammarström, A., Aronsson, G., Träskman Bendz, L., Grape, T., Hogstedt, C., Marteinsdottir, I., Skoog, I., & Hall, C. (2015). A systematic review including meta-analysis of work environment and depressive symptoms. *BMC Public Health*, *15*(1), 738. https://doi.org/10.1186/s12889-015-1954-4

Tourangeau, R. (1984). Cognitive science survey methods: A cognitive perspective. In T. Jabine, M. Straf, J. Tanur, & R. Tourangeau (Eds.), *Cognitive aspects of survey methodology: Building a bridge between disciplines* (pp. 73–100). National Academy Press.

Treier, M. (2019). *Gefährdungsbeurteilung psychischer Belastungen: Begründung, Instrumente, Umsetzung [Risk assessment of mental stress: Justification, instruments, implementation]*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-23293-1

Van Katwyk, P. T., Fox, S., Spector, P. E., & Kelloway, E. K. (2000). Using the Job-Related Affective Well-Being Scale (JAWS) to investigate affective responses to work stressors. *Journal of Occupational Health Psychology*, *5*(2), 219–230. https://doi.org/10.1037/1076-8998.5.2.219

Vazire, S. (2010). Who knows what about a person? The self–other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology*, *98*(2), 281–300. https://doi.org/10.1037/a0017908

VERBI Software. (2017). *MAXQDA 2018*. VERBI Software. https://www.maxqda.com

Wang, M., & Bodner, T. E. (2007). Growth mixture modeling: Identifying and predicting unobserved subpopulations with longitudinal data. *Organizational Research Methods*. https://doi.org/10.1177/1094428106289397

Wang, M., Sinclair, R. R., Zhou, L., & Sears, L. E. (2013). Person-centered analysis: Methods, applications, and implications for occupational health psychology. In R. R. Sinclair, M. Wang, & L. E. Tetrick (Eds.), *Research methods in occupational health psychology: Measurement, design, and data analysis* (pp. 349–373). Routledge.

Wetzer, I. M., Zeelenberg, M., & Pieters, R. (2007). "Never eat in that restaurant, I did!": Exploring why people engage in negative word-of-mouth communication. *Psychology & Marketing*, *24*(8), 661–680. https://doi.org/10.1002/mar.20178

Wilhelmy, A., Kleinmann, M., König, C. J., Melchers, K. G., & Truxillo, D. M. (2016). How and why do interviewers try to make impressions on applicants? A qualitative study. *Journal of Applied Psychology*, *101*(3), 313–332. https://doi.org/10.1037/apl0000046

Wimpenny, P., & Gass, J. (2000). Interviewing in phenomenology and grounded theory: Is there a difference? *Journal of Advanced Nursing*, *31*(6), 1485–1492. https://doi.org/10.1046/j.1365-2648.2000.01431.x

## 3.2 Study 2

# Contextualization of Work-Related Stress Items and its Effects on Criterion Validity and Reliability

Berit Greulich, Cornelius J. König, and Carla V. Meixensberger

Saarland University, Germany

### Author Note

Berit Greulich https://orcid.org/0000-0002-5955-7423, Cornelius J. König https://orcid.org/0000-0003-0477-8293, and Carla V. Meixensberger, Industrial and Organizational Psychology, Saarland University, Saarland, Germany.

Data available on request due to privacy/ethical restrictions. We have no conflicts of interest to disclose.

We gratefully acknowledge funding from the German Research Foundation (to Cornelius J. König), Grant KO4244/4-1.

Correspondence concerning this article should be addressed to Berit Greulich, Saarland University, Campus A1 3, 66123 Saarbrücken, Germany. E-mail: berit.greulich@gmail.com

**Abstract**

Employee surveys offer a cost-effective means of assessing work-related stress, but their generic formulation often allows for broad interpretation. Contextualizing survey items can enhance reliability and validity by reducing this interpretive scope. This study explores the effects of contextualizing work stress survey items on reliability and validity in medical assistants (N = 323). Items from the Copenhagen Psychosocial Questionnaire were adapted to the medical practice context with expert input. Participants completed generic and contextualized versions of the items in a counterbalanced order. Psychological strain served as the criterion. Despite lower reliability, contextualized items showed higher validity. Participants perceived greater predictive validity for job performance with the contextualized version. This study highlights the potential of contextualization to enhance work stress measurement validity.

*Keywords:* work-related stress, contextualization, work stress surveys, reliability, validity

## Contextualization of Work-Related Stress Items and its Effects on Criterion Validity and Reliability

Work stress is associated with physical and mental health affections (Ganster & Rosen, 2013). Additionally, it places a financial burden on organizations and society at large due to long-term sick leave, disability benefits, productivity loss, and healthcare costs (Hassard et al., 2018). In order to counteract these consequences at an early stage, management needs to know the extent of stressors and resources in their organization to successfully implement interventions (Havermans et al., 2018). A common practice to assess work stress is via self-report in employee surveys, which usually ask about the frequency and intensity of different stressors (e.g., time pressure, work interruptions) and resources (e.g., leadership quality, communication possibility) at the workplace. This method directly reflects employees' perspectives and has the advantage of economical conduction and simple implementation (Spector & Eatough, 2013). Nonetheless, the scope for interpretation of the items poses a notable challenge: Employee survey items are generally formulated in a generic and non-contextualized manner, which can lead to variations in subjects' interpretations of the items, and this variability in interpretation needs to be addressed to ensure accurate and unbiased measurement of constructs (Greulich et al., 2021).

To counteract the challenge of generic items, researchers interested in the assessment of personality have advocated for adding relevant context to questionnaires (Holtz et al., 2005). This contextualization creates a consistent frame-of-reference (FoR) that should have positive effects on criterion validity and reliability (Holtrop et al., 2014) due to the improvement of two quality criteria of questionnaires: increased within-person consistency due to responders rating items with a consistent frame-of-reference, and reduced between-person variability due to a

high number of respondents using a consistent frame-of-reference, Lievens et al., 2008). The aim of the present study is to transfer the positive effects of contextualization from the personality field to work stress survey items.

## Theoretical Background

Researchers have identified several work-related risk factors linked to increased rates of common mental health problems (e.g., low job control, high job demands, bullying, and role conflicts, Harvey et al., 2017). According to occupational health psychology models such as the Job Demands-Control model (Karasek, 1979) and the Job Demands-Resources model (Bakker et al., 2023), the frequency and severity of job demands (e.g., time pressure) lead to increased effort, which depletes employees' physical, emotional, and cognitive resources, leading to exhaustion and stress among employees and, over time, to serious health problems. Job resources (e.g., task control) serve as a kind of buffer and therefore weaken the impact of job demands on strain (e.g., Lesener et al., 2019). In fact, many studies have investigated the effects of work-related stress on employees' well-being and health, finding links to conditions such as musculoskeletal disorders (e.g., Taibi et al., 2021) and mental disorders (e.g., Gerhardt et al., 2021). In addition to the suffering that these illnesses represent for the individual employee, they are associated with considerable impact on organizational productivity: The economic costs of work-related stress are rather high (Hassard et al., 2018).

Due to the negative consequences of work-related stress, organizations have an interest in identifying the nature and extent of potential stressors and resources. This creates the basis for the development and implementation of measures to prevent negative outcomes. Work-related stressors and resources are usually assessed via self-reports (e.g., Kristensen et al., 2005; Morgeson & Humphrey, 2006). These are work

stress surveys that can be completed by any employee as part of an employee survey by rating stressors and resources in terms of their frequency and intensity. Typical stressors measured with such questionnaires include time pressure, concentration demands, and one-sided physical demands; typical resources surveyed include task control, support from colleagues or supervisors, and development opportunities.

Self-reports are popular in this context because they are convenient and economical for organizations, and employees likely know their workplace better than any other potential rater (Semmer et al., 1995). In addition, stress research has shown that the subjective appraisal of stressors and resources is crucial for the perception of stress (Lazarus & Folkman, 1987). For these surveys to be cost-effective, they must be easily applicable to many different workplaces. Therefore, these questionnaires often consist of generically formulated items that leave a wide scope for interpretation. Answering survey items is a complex cognitive process that is susceptible to biases (Greulich et al., 2021; Tourangeau, 1987). It has been argued that respondents are likely to interpret generic items differently and give each of them an individual context (Credé et al., 2010; Holtrop et al., 2014; Holtz et al., 2005; Lievens et al., 2008; Mount et al., 1994; Schmit et al., 1995). The large scope of interpretation in turn poses a problem for measurement accuracy. Interindividual differences in interpretation, for example, can lead to losses in criterion validity (Lievens et al., 2008). If measures are inaccurate, this might influence managements' implementation of prevention measures.

Researchers have begun to study contextualization on the assumption that it has a positive effect on criterion validity (e.g., Lievens et al., 2008; Voss et al., 2023). Contextualization means that survey items are embedded in a detailed and relevant context (sometimes also called 'adding FoR', both labels are used synonymously here;

Schmit et al., 1995). Different types of contextualization can be found in the literature, varying in their extent (i.e., low, medium, high; Holtrop et al., 2014; Lievens & Sackett, 2017). A *low* degree of contextualization occurs using an instruction. For example, subjects are asked to think of a situation while completing a survey. *Medium* means tagged contextualization, which assigns a tag to each item that can refer to an environment or to a time period. Subjects are thus given a context to refer to when answering the items (Debus et al., 2019; Holtrop et al., 2014). When an item is completely rewritten to match a specific context, a *high* extent of contextualization is achieved (Holtrop et al., 2014). Giving self-report survey items a FoR limits the individual interpretation of items, and interpreting items might be more complicated for respondents than envisioned by survey designers (Greulich et al., 2021). If FoR reduces the scope of interpretation, it should also improve the internal consistency (i.e., the reliability) of measures, and some studies have indeed shown that contextualization of personality survey items enhances reliability (e.g., Lievens et al., 2008; Swift & Peterson, 2019). Contextualization has also been found to enhance criterion-related validity of personality scales in comparison to standard non-contextualized inventories by reducing error variance (for meta-analytic evidence see Shaffer & Postlethwaite, 2012).

In the context of work stress surveys, responding items has been shown to require different cognitive and motivational processes, each influenced by individual contextual factors due to uncertainty about item interpretation (Greulich et al., 2021). Therefore, standardizing work stress survey items by adding a context could go along with the positive effects of a FoR, as it enables employees to more effectively evaluate their work-related stressors and resources. Indeed, it has already been shown that adding time contexts to a job insecurity scale can influence response behavior

(Debus et al., 2019). Furthermore, the act of contextualizing work stress survey items by incorporating a well-defined social comparison has been demonstrated to affect the means of stressor and resource scales, as well as enhance the criterion validity of a stress scale (Greulich et al., 2023). It is reasonable to postulate that the process of contextualization through the redesign of survey items exerts a similar influence on scales employed for assessing work-related stress. In contrast to standard personality survey items, workplace-related stress items already have an environmental context. For example, items like “Do you have to work fast?” obviously refer to work, making an exploration of the effects of complete contextualization more relevant than tagged contextualization. Furthermore, complete contextualization was reported to lead to an additional criterion validity increase compared to tagged contextualization (Holtrop et al., 2014). Thus, this study tests the effect of complete contextualization with the expectation that there should be positive effects on reliability and criterion validity (with psychological strain being the criterion in this study). More formally, we hypothesize:

*Hypothesis 1a: Complete contextualized work stress survey items have a higher internal consistency than the generic work-related stress items.*

*Hypothesis 1b: Complete contextualized work stress survey items have a higher criterion validity than the generic work stress survey items.*

Participants may also react differently to contextualized items. According to Gilliland (1993), contextualized items were preferred by participants because they make a stronger connection to their situation than do the generic items. Indeed, Holtrop and colleagues (2014) found increased perceived predictive and face validity for contextualized items (although participants unexpectedly liked generic items

better than contextualized ones). Given that our participants are employees, contextualizing work stress items should give them a greater understanding of their current situation, potentially inducing a sense of fairness when completing the items (Gilliland, 1993). Thus, following Holtrop et al. (2014), we investigated three participant reactions: liking of the different item versions, face validity (i.e., perceived job relevance), and perceived predictive validity (for their working performance). We expect more positive participant reactions for the complete contextualized work-related stress items. Thus, we hypothesize:

*Hypothesis 2: Complete contextualized items will achieve a higher approval rating for participants' reactions (liking, face validity, and perceived predictive validity) than the generic items.*

## Method

### Sample

The sample consists of 323 German medical assistants in doctor's offices of different specialties (e.g., family medicine, gynecology, cardiology etc.). To recruit them, we contacted randomly selected doctor's offices in two towns, and additional data were collected in a training center for medical assistants. Informed consent was obtained from all participants. Nearly all participants were female (99.6%) and aged between 18 and 63 years ($M$ = 36.49, $SD$ = 11.62). On average, the participants had been employed for 15.17 years ($SD$ = 11.28) and 66.5 % were working full-time.

### Procedure

The study was conducted via a paper-and-pencil questionnaire, which consisted of demographic data, a generic version of work stress items and participants' reaction to them, a contextualized version of work stress items and participants reaction to them, as well as a psychological strain scale. The order of the

generic and contextualized versions was counterbalanced. The questionnaires were distributed to doctors' offices by a research assistant. After completion of the questionnaire, the participants placed them in a ballot box that was picked up after approximately two weeks.

## Measures

### *Work Stress Survey*

With the help of a physician and a medical assistant, relevant stressor and resource scale items were selected from the Copenhagen Psychosocial Questionnaire (Kristensen et al., 2005; Nübling et al., 2005), a commonly used work stress survey in Germany. A total of 13 items were selected from the German version of the questionnaire for their specific relevance to a medical assistant's everyday work and their potential to transform them into a highly contextualized version. The items can be assigned to five different scales. Three scales were used to assess stressors: quantitative demands, emotional demands, and hiding emotions; Two scales were used to assess resources at work: sense of community and quality of leadership. All items are assessed in terms of frequency on a five-point Likert scale ranging from 1 (never) to 5 (always).

For the generic version, the selected items remained unchanged. To create the complete contextualized version (see Table 1), the procedure of complete item contextualization was adopted from Holtrop et al. (2014). First, two research assistants independently developed a complete contextualized version of the previously selected generic items. Subsequently, they discussed and agreed on a contextualized version for each item. Finally, a medical assistant tried to assign the contextualized items to the respective generic ones. If this was successful, the contextualized item was retained. Otherwise, the process started again. The aim of the

contextualization was to establish a clear reference to the everyday work of a medical assistant by illustrating stressors and resources in a specific behavioral context and by using job-specific terms (e.g., “patient”). Correlations between the same stressors and resources scales in the generic and contextualized version (see Table 2) provide evidence that the underlying constructs are measured despite modifications.

### *Participants’ Reactions*

Reactions were measured as in Holtrop et al. (2014). Directly after completing each version of the work stress survey, participants reactions to the particular version (generic/contextualized) were measured via three items. For liking, an item based on Wiechmann and Ryan (2003) was used: *I did not enjoy completing this questionnaire* (reverse coded). For face validity and perceived predictive validity, items from Smither et al. (1993) were used: *The content of this questionnaire is clearly related to my work* and *With the results of this questionnaire my work performance can be predicted.* Participants reacted on a five-point Likert scale, ranging from 1 (completely disagree) to 5 (completely agree). Holtrop et al. (2014) analyzed on the item-level, and we followed their example.

### *Psychological Strain*

To test the effect of complete contextualization on criterion validity, we used psychological strain as a criterion, using the General Health Questionnaire with twelve items (GHQ-12; Goldberg & Hillier, 1979; Linden et al., 1996) as the measure. The GHQ-12 has a four-point response scale, corresponding to symptom present from 0 (not at all) to 3 (much more than usual), resulting in a possible range of 0 to 36 and with higher scores representing higher level of psychological strain.

**Table 1**

*Generic and Contextualized Work-Related Stress Items, Chosen From the Copenhagen Psychosocial Questionnaire (Kristensen et al., 2005; Nübling et al., 2005)*

| Scale | Generic version | Contextualized version |
|---|---|---|
| Quantitative demands | Do you have to work very fast? | In the everyday work of a medical assistant, there are situations in which many things have to be done as quickly as possible. For example, careless mistakes can happen when coordinating appointments due to time pressure. How often do you have to hurry when processing tasks? |
| | How often do you not have time to complete all your work tasks? | In the everyday work of a medical assistant, it can happen that a lot of tasks have to be handled. How often does it happen, for example, that you do not have enough time for administrative tasks because you also have to take care of the patients' needs on site, or vice versa? |
| | Do you have to do overtime? | In the everyday work of a medical assistant, it can happen that patients who do not have an appointment come to the practice shortly before closing time. This means that working hours are extended. How often does this happen? |
| Emotional demands | Do you have to deal with other people’s personal problems as part of your work? | In the everyday work of a medical assistant, it is part of the job to be in contact with patients and relatives. How often do you have to deal with patients' personal problems? |
| | Is your work emotionally demanding? | Confronting the suffering of patients is part of the everyday work of a medical assistant. How often do you find that their fate is very close to you and places an emotional burden on you? |
| Hiding emotions | Does your work require that you hide your feelings? | The everyday work of a medical assistant involves many emotional situations. Feelings such as sadness or anger in contact with patients are not uncommon. How often do you have to suppress your feelings in order to appear "neutral" to others? |

| | | |
|---|---|---|
| | Does your work require that you do not state your opinion? | The everyday work of a medical assistant involves many conflictual situations. Nevertheless, you must always maintain your professionalism. How often does it happen that you have to show an opinion to others in a situation that is not in line with your actual opinion? |
| Sense of community | Is there a good atmosphere between you and your colleagues? | In the everyday work of a medical assistant teamwork plays a major role. Good teamwork has a positive effect on morale. How often do you feel that the cooperation between you and your colleagues is harmonious? |
| | Is there good co-operation between the colleagues at work? | In the everyday work of a medical assistant, patient care and administrative tasks should be understood as teamwork. How well does co-operation between colleagues work? |
| Quality of leadership | To what extent would you say that your immediate superior makes sure that the members of staff have good development opportunities? | In the everyday work of a medical assistant, it is important to constantly develop both social and professional skills. How often does your immediate supervisor attach importance to training and other opportunities for further development? |
| | To what extent would you say that your immediate superior gives high priority to job satisfaction? | Job satisfaction is very important in the everyday work of a medical assistant as it influences everyday contact with patients and colleagues, for example. How often does your immediate supervisor take time to talk to you about how you are feeling? |
| | To what extent would you say that your immediate superior is good at work planning? | Good planning and organization are essential to ensure that practice processes run smoothly. How well is your immediate supervisor structured at work and does it allow for good planning of procedures in the doctor's office? |
| | To what extent would you say that your immediate superior is good at solving conflicts? | In the everyday work of a medical assistant problems can arise due to conflicts in the team or too high workload. How often does your immediate supervisor give you the opportunity to voice these problems and offer possible solutions? |

*Note.* All items originally in German.

**Table 2**

*Correlations for Predictor Variables and the Criterion Variable*

| | | *M* | *SD* | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Predictor | | | | | | | | | | | | | |
| 1 | (Generic) Quantitative demands | 3.38 | 0.69 | | | | | | | | | | |
| 2 | (Generic) Emotional demands | 3.53 | 0.87 | .34** | | | | | | | | | |
| 3 | (Generic) Hiding emotions | 3.03 | 1.01 | .18** | .39** | | | | | | | | |
| 4 | (Generic) Sense of community | 4.13 | 0.75 | -.12* | -.02 | -.10 | | | | | | | |
| 5 | (Generic) Quality of leadership | 3.30 | 0.99 | -.25** | -.09 | -.18** | .33** | | | | | | |
| 6 | (Contextualized) Quantitative demands | 3.48 | 0.72 | .62** | .43** | .23** | -.12* | -.19** | | | | | |
| 7 | (Contextualized) Emotional demands | 3.11 | 0.77 | .21** | .55** | .23** | -.03 | .04 | .33** | | | | |
| 8 | (Contextualized) Hiding emotions | 3.04 | 0.76 | .11* | .26** | .35** | -.12 | -.13* | .22** | .32** | | | |
| 9 | (Contextualized) Sense of community | 3.88 | 0.79 | -.16** | -.04 | -.10 | .73** | .37** | -.15** | -.01 | -.15** | | |
| 10 | (Contextualized) Quality of leadership | 3.21 | 0.90 | -.20** | -.11* | -.21** | .32** | .79** | -.15* | .04 | -.19** | .33** | |
| Criterion | | | | | | | | | | | | | |
| 11 | GHQ score | | | .33** | .16** | .22** | -.28** | -.32** | .32** | .16** | .25** | -.34** | -.29** |

*Note*. $N = 323$. GHQ = General Health Questionnaire.

*$p < .05$, ** $p < .01$

## Results

For reliability analysis, Cronbach's alpha was calculated for each scale (reported in Table 3). The item analysis revealed that the fourth item of the contextualized scale 'quality of leadership' had to be excluded due to insufficient discriminatory power ($r < .3$). Most of the generic stressor and resource scales used in this study showed acceptable to high reliability (.65 to .88); the generic scale 'quantitative demands' had a rather low Cronbach's alpha (.59). To analyze changes in the internal consistency, we compared them using the computer program Alpha Test, a program for testing hypotheses about coefficient alpha using chi-square (Lautenschlager & Meade, 2008). The internal consistency of the stressor scales 'quantitative demands' and 'emotional demands' did not differ between the generic and context versions, $\chi^2(1) = 1.51$, $p = .219$ (for 'quantitative demands'), $\chi^2(1) = 2.10$, $p = .148$ (for 'emotional demands'). All other scales decreased significantly in internal consistency due to item contextualization, 'hiding emotions': $\chi^2(1) = 6.02$, $p = .014$; 'sense of community': $\chi^2(1) = 12.12$, $p < .001$; 'quality of leadership': $\chi^2(1) = 50.47$, $p < .001$. These results therefore did not support H1a.

Table 4 reports the results regarding H1b (i.e., the effects on criterion validity). Hierarchical regression analyses showed that the completely contextualized version of the scales explained additional variance in the GHQ-scores over and above the generic version for the scales "quantitative demands", "hiding emotions", and "sense of community", as indicated by a significant increase in explained variance (i.e., $\Delta R^2$). Such an increase in $\Delta R^2$ was not found for the scales 'emotional demands' and 'quality of leadership'. These findings offer some support for H1b.

**Table 3**

*Means, Standard Deviations, and Internal Consistency*

| Variable | | Generic | Contextualized |
|---|---|---|---|
| Stressors | | | |
| Quantitative demands | *M* | 3.38 | 3.48 |
| | *(SD)* | (0.69) | (0.72) |
| | α | .59 | .65 |
| Emotional demands | *M* | 3.53 | 3.11 |
| | *(SD)* | (0.87) | (0.77) |
| | α | .65 | .56 |
| Hiding emotions | *M* | 3.03 | 3.04 |
| | *(SD)* | (1.01) | (0.76) |
| | α | .73 | .58 |
| Resources | | | |
| Sense of community | *M* | 4.13 | 3.88 |
| | *(SD)* | (0.75) | (0.79) |
| | α | .88 | .83 |
| Quality of leadership | *M* | 3.30 | 3.21 |
| | *(SD)* | (0.99) | (0.90) |
| | α | .85 | .69 |

*Note.* α = Cronbach's Alpha. *N* = 323.

**Table 4**

*Results of Hierarchical Regression Analyses for Psychological Strain*

| Variable | Model | $R$ | $R^2$ | $R^2_{adj}$ | $SE$ | $\Delta R^2$ | $F_{change}$ | $df_1$ | $df_2$ | $p$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Quantitative demands | 1[a] | .33 | .11 | .11 | .50 | .11 | 39.46 | 1 | 321 | < .001 |
| | 2[b] | .36 | .13 | .13 | .50 | .02 | 7.54 | 1 | 320 | .006 |
| Emotional demands | 1[a] | .16 | .03 | .02 | .53 | .03 | 8.63 | 1 | 320 | .004 |
| | 2[b] | .19 | .03 | .03 | .53 | .01 | 2.63 | 1 | 319 | .15 |
| Hiding demands | 1[a] | .22 | .05 | .05 | .52 | .05 | 16.97 | 1 | 321 | < .001 |
| | 2[b] | .29 | .09 | .08 | .51 | .03 | 12.03 | 1 | 320 | < .001 |
| Sense of community | 1[a] | .28 | .08 | .07 | .51 | .08 | 26.63 | 1 | 319 | < .001 |
| | 2[b] | .34 | .12 | .11 | .50 | .04 | 13.51 | 1 | 318 | < .001 |
| Quality of leadership | 1[a] | .32 | .10 | .10 | .51 | .10 | 35.78 | 1 | 321 | < .001 |
| | 2[b] | .32 | .11 | .10 | .51 | .00 | 1.60 | 1 | 320 | .206 |

*Note.* [a] Predictor: Generic version; [b] Predictors: Generic version and contextualized version. $N$ = 323.

In the last step, we tested if contextualization affected participants' reactions. Overall, participants' reactions were rather neutral (see Table 5). Table 5 also shows the *t*-values for paired tests as well as effect sizes: Liking, and face validity did not significantly differ but perceived predictive validity differed between the generic and the contextualized versions of the stressor and resource scales, $t(312) = 2.39$, $p < .05$, $d = 0.14$. Therefore, H2 was partly supported.

**Table 5**

*Descriptive Statistics and Paired* t*-tests for Participant Reactions*

| Reaction Variable | $M\ (SD)_{generic}$ | $M\ (SD)_{contextualized}$ | *t* | *p* | $d_{within}$ |
|---|---|---|---|---|---|
| Liking | 3.14 (0.90) | 3.21 (0.92) | -1.44 | .15 | 0.08 |
| Face validity | 3.54 (0.83) | 3.54 (0.80) | -0.14 | .89 | 0.01 |
| Perceived predictive Validity | 3.09 (0.81) | 3.20 (0.81) | -2.39 | .02* | 0.14 |

*Note.* $N = 323$.

*$p < .05$.

## Discussion

Our results show that contextualization can have an impact on criterion validity even in work stress survey items. The complete contextualized version of three out of the five scales explained more variance in criterion (i.e., psychological strain) than generic versions. However, contextualization did not result in improved reliability. Participant reactions also did not differ between the generic and contextualized versions.

The contextualization effect has been studied in more detail mainly for personality questionnaires (e.g., Lievens et al., 2008; Swift & Peterson, 2019). Preliminary evidence from the field of occupational health psychology suggests that

this effect can be transferred (Debus et al., 2019; Greulich et al., 2023), and the current study serves as an additional indicator in support of this notion As the interviewees in the qualitative study by Greulich and colleagues (2021) have indicated, it can be quite challenging to answer stressor and resource scales in work-related stress surveys because several interpretations of items are often possible. Contextualization stress items thus might offer a way to reduce the scope of interpretation (Greulich et al., 2023). At the same time, it should be noted that compared to other studies (e.g., Bing et al., 2004; Holtrop et al., 2014; Reddock et al., 2011), the increase in criterion validity was a bit smaller, which was likely due to the fact that the comparison version with the generic items already include some kind of contextualization because all items are work-related, whereas previous studies compared a contextualized version with a context-free one. The fact that the positive effect did not show up for all scales fits with the results of Voss and colleagues (2022) who concluded that the effects of contextualization might be scale-specific.

Contrary to our expectation triggered by arguments by Lievens et al. (2008), no increase in reliability was shown by the complete contextualization of the items. Our results seem to suggest that the completely contextualized items were handled inconsistently by the individual responder. A reason might be that we used the complete contextualization, which means that items are redesigned to fit within a particular context (in our study: by establishing a reference to the everyday activities of medical assistants), whereas complete contextualization in the personality questionnaires means, for example, redesigning a generic item such as ‘I see myself as someone who likes to cooperate with others’ as ‘I see myself as someone who is a team player at sports’ to illustrate a sport context. Thus, our type of contextualization was much more specific, and this also meant that items were significantly more

extensive and sometimes contained several sentences. Maybe, the items were so specific that they were no longer answered a coherent way.

Last, participants' reactions were rather similar to the generic and the contextualized version of the questionnaire. There were no differences in liking and face validity. An explanation could be that the work-related stressor and resource items already contain the context of work, and an even stronger contextualization to the specific workplace might thus not offer added value for the participants' response. Nevertheless, it might also be good news for practitioners that the longer version did not produce negative reactions despite being considerably longer. Interestingly, participants reported a higher perceived predictive validity for work performance with the contextualized items. This suggests that employees believe that inferences can be drawn about their job performance based on how they perceive stressors and resources. For instance, if employees perceive the stressors and resources as highly burdensome, they may fear being viewed as less resilient and poor performers. These findings are consistent with the results of Greulich et al. (2021), whose participants expressed concerns about the potential attribution of their responses and the resulting negative consequences. On the other hand, employees may also leverage this perception to their advantage by evaluating their work environment and tasks as minimally burdensome, thereby creating a positive impression of their job performance. Overall, this is an intriguing finding that warrants further investigation.

Like all studies, this study has its limitation – in particular the small numbers of items selected from the Copenhagen Psychosocial Questionnaire (Kristensen et al., 2005; Nübling et al., 2005), resulting in low internal consistency for some scales. The main reason for this was that not all items were suitable for complete contextualization for the everyday work of a medical assistant. Furthermore, we

wanted to avoid excessive response burden, because the participants had to answer both versions of the items due to the within-subject design and the contextualized items required considerably more reading. This resulted in three scales consisting of only two items each, which limits the reliability of the scale from the outset. However, given that we were interested in the reliability differences between conditions and not absolute reliability level, this seemed acceptable.

Further research is warranted to enhance our understanding of item contextualization. As previously mentioned, there is a possibility that items may have been overly specific, resulting in incoherent responses. Thus, it is crucial to determine the appropriate level of contextualization. Specifically, researchers could investigate the effectiveness of utilizing tagged contextualization for work stress items. Employing a moderate level of contextualization, linked to specific job settings (e.g., "in your doctor's office") or positions (e.g., "in the lab"), might suffice to establish a shared context among employees. Additionally, future research should aim to elucidate the underlying reasons for the increased criterion validity observed in the contextualized version of work-related stress surveys. Lievens and colleagues (2008) suggested that using contextualized items leads to decreased within-person variability. Our data, results from Lievens and colleagues' (2008) Study 1 and results from Reddock et al. (2011) do not offer strong support for this argument. Thus, other reasons need to be tested.

Taken together, our results illustrate that contextualizing work-related stressor and resource scales can have an impact on criterion validity of these scale. Thus, contextualizing might be a way for organizations to obtain better data about the mental health of their employees, which should then lead to more appropriate interventions.

## References

Bakker, A. B., Demerouti, E., & Sanz-Vergel, A. (2023). Job demands-resources theory: Ten years later. *Annual Review of Organizational Psychology and Organizational Behavior*, *10*(1), 25–53. https://doi.org/10.1146/annurev-orgpsych-120920-053933

Bing, M. N., Whanger, J. C., Davison, H. K., & VanHook, J. B. (2004). Incremental validity of the frame-of-reference effect in personality scale scores: A replication and extension. *Journal of Applied Psychology*, *89*(1), 150–157. https://doi.org/10.1037/0021-9010.89.1.150

Brunner, B., Igic, I., Keller, A. C., & Wieser, S. (2019). Who gains the most from improving working conditions? Health-related absenteeism and presenteeism due to stress at work. *European Journal of Health Economics*, *20*(8), 1165–1180. https://doi.org/10.1007/s10198-019-01084-9

Credé, M., Bashshur, M., & Niehorster, S. (2010). Reference group effects in the measurement of personality and attitudes. *Journal of Personality Assessment*, *92*(5), 390–399. https://doi.org/10.1080/00223891.2010.497393

Debus, M. E., Greulich, B., König, C. J., & Kleinmann, M. (2019). Insecure about how to rate your job insecurity? A two-study investigation into time frames applied to job insecurity measures. *Occupational Health Science*, *3*(4), 421–435. https://doi.org/10.1007/s41542-019-00049-x

Ganster, D. C., & Rosen, C. C. (2013). Work stress and employee health: A multidisciplinary review. *Journal of Management*, *39*(5), 1085–1122. https://doi.org/10.1177/0149206313475815

Gerhardt, C., Semmer, N. K., Sauter, S., Walker, A., de Wijn, N., Kälin, W., Kottwitz, M. U., Kersten, B., Ulrich, B., & Elfering, A. (2021). How are social stressors at work related to well-being and health? A systematic review and meta-analysis. *BMC Public Health*, *21*(1), Article 890. https://doi.org/10.1186/s12889-021-10894-7

Gilliland, S. W. (1993). The perceived fairness of selection systems: An organizational justice perspective. *Academy of Management Review*, *18*(4), 694–734. https://doi.org/10.5465/amr.1993.9402210155

Goldberg, D. P., & Hillier, V. F. (1979). A scaled version of the General Health Questionnaire. *Psychological Medicine*, *9*(1), 139–145. https://doi.org/10.1017/s0033291700021644

Greulich, B., Debus, M. E., Kleinmann, M., & König, C. J. (2021). Response behavior in work stress surveys: A qualitative study on motivational and cognitive processes in self- and other-reports. *European Journal of Work and Organizational Psychology*, *30*(1), 40–57. https://doi.org/10.1080/1359432X.2020.1812580

Greulich, B., König, C. J., & Fischer, N. (in press). How relative is stress? The influence of social comparisons when responding to work stress surveys. *Journal of Personnel Psychology.*

Harvey, S. B., Modini, M., Joyce, S., Milligan-Saville, J. S., Tan, L., Mykletun, A., Bryant, R. A., Christensen, H., & Mitchell, P. B. (2017). Can work make you mentally ill? A systematic meta-review of work-related risk factors for common mental health problems. *Occupational and Environmental Medicine*, *74*(4), 301–310. https://doi.org/10.1136/oemed-2016-104015

Hassard, J., Teoh, K. R. H., Visockaite, G., Dewe, P., & Cox, T. (2018). The cost of work-related stress to society: A systematic review. *Journal of Occupational Health Psychology*, *23*(1), 1–17. https://doi.org/10.1037/ocp0000069

Havermans, B. M., Brouwers, E. P. M., Hoek, R. J. A., Anema, J. R., van der Beek, A. J., & Boot, C. R. L. (2018). Work stress prevention needs of employees and supervisors. *BMC Public Health*, *18*(1), 642. https://doi.org/10.1186/s12889-018-5535-1

Holtrop, D., Born, M. Ph., de Vries, A., & de Vries, R. E. (2014). A matter of context: A comparison of two types of contextualized personality measures. *Personality and Individual Differences*, *68*, 234–240. https://doi.org/10.1016/j.paid.2014.04.029

Holtz, B. C., Ployhart, R. E., & Dominguez, A. (2005). Testing the rules of justice: The effects of frame-of-reference and pre-test validity information on personality test responses and test perceptions. *International Journal of Selection & Assessment*, *13*(1), 75–86. https://doi.org/10.1111/j.0965-075X.2005.00301.x

Karasek, R. A. (1979). Job demands, job decision latitude, and mental strain: Implications for job redesign. *Administrative Science Quarterly*, *24*(2), 285–308. https://doi.org/10.2307/2392498

Kristensen, T. S., Borritz, M., Villadsen, E., & Christensen, K. B. (2005). The Copenhagen Burnout Inventory: A new tool for the assessment of burnout. *Work & Stress*, *19*(3), 192–207. https://doi.org/10.1080/02678370500297720

Lautenschlager, G. J., & Meade, A. W. (2008). AlphaTest: A Windows program for tests of hypotheses about coefficient alpha. *Applied Psychological Measurement*, *32*(6), 502–503. https://doi.org/10.1177/0146621607312307

Lazarus, R. S., & Folkman, S. (1987). Transactional theory and research on emotions and coping. *European Journal of Personality*, *1*(3), 141–169. https://doi.org/10.1002/per.2410010304

Lesener, T., Gusy, B., & Wolter, C. (2019). The job demands-resources model: A meta-analytic review of longitudinal studies. *Work & Stress*, *33*(1), 76–103. https://doi.org/10.1080/02678373.2018.1529065

Lievens, F., de Corte, W., & Schollaert, E. (2008). A closer look at the frame-of-reference effect in personality scale scores and validity. *Journal of Applied Psychology*, *93*(2), 268–279. https://doi.org/10.1037/0021-9010.93.2.268

Lievens, F., & Sackett, P. R. (2017). The effects of predictor method factors on selection outcomes: A modular approach to personnel selection procedures. *Journal of Applied Psychology*, *102*(1), 43–66. https://doi.org/10.1037/apl0000160

Linden, M., Maier, W., Achberger, M., Herr, R., Helmchen, H., & Benkert, O. (1996). Psychische Erkrankungen und ihre Behandlung in Allgemeinarztpraxen in Deutschland: Ergebnisse aus einer Studie der Weltgesundheitsorganisation (WHO). [Psychological disorders and their treatment in general practice in Germany: Results of a WHO study.]. *Der Nervenarzt*, *67*(3), 205–215.

Morgeson, F. P., & Humphrey, S. E. (2006). The Work Design Questionnaire (WDQ): Developing and validating a comprehensive measure for assessing job design and the nature of work. *Journal of Applied Psychology*, *91*(6), 1321–1339. https://doi.org/10.1037/0021-9010.91.6.1321

Mount, M. K., Barrick, M. R., & Strauss, J. P. (1994). Validity of observer ratings of the Big Five personality factors. *Journal of Applied Psychology*, *79*(2), 272–280. https://doi.org/10.1037/0021-9010.79.2.272

Nübling, M., Stößel, U., Hasselhorn, H.-M., Michaelis, M., & Hofmann, F. (2005). *Methoden zur Erfassung psychischer Belastungen: Erprobung eines Messinstrumentes (COPSOQ) [Methods for recording mental stress: Testing of a measuring instrument (COPSOQ)]*. Wirtschaftsverlag NW.

Reddock, C. M., Biderman, M. D., & Nguyen, N. T. (2011). The relationship of reliability and validity of personality tests to frame-of-reference instructions and within-person inconsistency. *International Journal of Selection and Assessment*, *19*(2), 119–131. https://doi.org/10.1111/j.1468-2389.2011.00540.x

Schmit, M. J., Ryan, A. M., Stierwalt, S. L., & Powell, A. B. (1995). Frame-of-reference effects on personality scale scores and criterion-related validity. *Journal of Applied Psychology*, *80*(5), 607–620. https://doi.org/10.1037/0021-9010.80.5.607

Semmer, N. K., Zapf, D., & Dunckel, H. (1995). Assessing stress at work: A framework and an instrument. In O. Svane & C. Johansen (Eds.), *Work and health—Scientific basis of progress in the working environment* (pp. 105–113). Office for Official Publications of the European Communities.

Semmer, N. K., Zapf, D., & Dunckel, H. (1999). Instrument zur stressbezogenen Tätigkeitsanalyse ISTA [Instrument for Stress-Oriented Task Analysis ISTA]. In H. Dunckel (Ed.), *Handbuch psychologischer Arbeitsanalyseverfahren* (pp. 179–204). vdf Hochschulverlag.

Shaffer, J. A., & Postlethwaite, B. E. (2012). A matter of context: A meta-analytic investigation of the relative validity of contextualized and noncontextualized personality measures. *Personnel Psychology*, *65*(3), 445–494. https://doi.org/10.1111/j.1744-6570.2012.01250.x

Smither, J. W., Reilly, R. R., Millsap, R. E., Pearlman, K., & Stoffey, R. W. (1993). Applicant reactions to selection procedures. *Personnel Psychology*, *46*(1), 49–76. https://doi.org/10.1111/j.1744-6570.1993.tb00867.x

Spector, P. E., & Eatough, E. (2013). Quantitative self-report methods in occupational health psychology research. In M. Wang, S. Sinclair, & L. E. Tetrick (Eds.), *Research methods in occupational health psychology* (pp. 248–267). Taylor & Francis.

Swift, V., & Peterson, J. B. (2019). Contextualization as a means to improve the predictive validity of personality models. *Personality and Individual Differences*, *144*, 153–163. https://doi.org/10.1016/j.paid.2019.03.007

Taibi, Y., Metzler, Y. A., Bellingrath, S., & Müller, A. (2021). A systematic overview on the risk effects of psychosocial work characteristics on musculoskeletal disorders, absenteeism, and workplace accidents. *Applied Ergonomics*, *95*, 103434. https://doi.org/10.1016/j.apergo.2021.103434

Tourangeau, R. (1987). Attitude measurement: A cognitive perspective. In H.-J. Hippler, N. Schwarz, & S. Sudman (Eds.), *Social information processing and survey methodology* (pp. 149–162). Springer.

Voss, N. M., Chlevin-Thiele, C., Lake, C. J., & Warren, C.-L. (2022). Using item response theory to understand the effects of scale contextualization: An illustration using decision making style scales. *International Journal of Testing, advance online publication*. https://doi.org/10.1080/15305058.2022.2047692

Voss, N. M., Chlevin-Thiele, C., Lake, C. J., & Warren, C.-L. (2023). Using item response theory to understand the effects of scale contextualization: An illustration using decision making style scales. *International Journal of Testing*, *23*(1), 34–55. https://doi.org/10.1080/15305058.2022.2047692

Wiechmann, D., & Ryan, A. M. (2003). Reactions to computerized testing in selection contexts. *International Journal of Selection and Assessment*, *11*(2–3), 215–229. https://doi.org/10.1111/1468-2389.00245

## 3.3 Study 3

# How Relative is Stress?

# The Influence of Social Comparisons When Responding to Work Stress Surveys

Berit Greulich, Cornelius J. König, and Nancy Fischer

Universität des Saarlandes, Germany

## Author -Note

Berit Greulich https://orcid.org/0000-0002-5955-7423, Cornelius J. König https://orcid.org/0000-0003-0477-8293, and Nancy Fischer, Industrial and Organizational Psychology, Saarland University, Saarland, Germany.

Data available on request due to privacy/ethical restrictions. We have no conflicts of interest to disclose.

We gratefully acknowledge funding from the German Research Foundation (to Cornelius J. König), Grant KO4244/4-1.

Correspondence concerning this article should be addressed to Berit Greulich, Saarland University, Campus A1 3, 66123 Saarbrücken, Germany. E-mail: berit.greulich@gmail.com

**Abstract**

In self-reports, employees frequently use self-selected social comparisons to assess workplace stressors and resources, but these comparisons vary within and between individuals. This study investigates how standardizing social comparison processes by adding a prescribed comparison to each item affects the reliability and validity of self-report scales measuring work-related stressors and resources, and how the standardized comparison affects scale means. A total of 208 employees were randomly assigned to one of two groups, comparing their perceptions to either their direct colleagues or without instructed comparison. Results indicate no effect on reliability, improvement in validity for one scale, and differences in means between groups for stressors and a resource scale. These findings suggest potential benefits and drawbacks of standardized social comparisons in self-report measures.

*Keywords*: work stress survey, social comparison, contextualization, reliability, validity

# How Relative is Stress?

## The Influence of Social Comparisons When Responding to Work Stress Surveys

Work stress surveys typically contain stressors (e.g., time pressure) and resources (e.g., task control) that should be rated regarding their frequency or intensity (e.g., Kristensen et al., 2005; Morgeson & Humphrey, 2006). Employees typically do this themselves in the form of self-report questionnaires, which is convenient and economical because a large number of employees can participate and are thus able to provide important information about their workplaces (Spector & Eatough, 2013). As noted by Greulich and colleagues (2021), employees use various social comparisons to evaluate their work-related stress, which might be a source of response bias. Given the diversity of social comparison objects, it might be an alternative to standardize the process of social comparison in order to investigate its impact on the reliability and validity of work stress surveys. Thus, the current study aims to investigate the potential effects of standardizing social comparison on the assessment of stressors and resources in the workplace with a focus on the psychometric implications of these findings. Additionally, the study will examine the role of social comparison in relation to work-related stress and how this impacts employees' evaluations.

## Theoretical Background

In the context of occupational health, job characteristics have been a particular focus of stress theories. The Job Demand Control Model was originally developed to explain the impact and interaction of job demands and job control as a resource on employee health and well-being (Karasek, 1979; Karasek & Theorell, 1990). Since social support from colleagues and supervisors also play an important role in promoting employee health and well-being, the Job Demand Control Model has been

expanded to include social support as a key resource (Johnson & Hall, 1988; Van der Doef & Maes, 1999). However, to fully capture the complex and multifaceted nature of work-related stress, it may still be too limited (de Jonge et al., 2010). In response to this limitation, researchers have proposed the Job Demand Resource (JDR) model, which offers a more comprehensive and nuanced understanding of work-related stress (Demerouti et al., 2001). This model posits that all job characteristics can be categorized as either demands or resources. Demands (in the survey instrument used here, demands correspond to stressors, which is why this term is used in the following, Irmer et al., 2019) refer to the physical, psychological, social, and organizational aspects of work that require sustained effort and are associated with physical and psychological costs, such as burnout, fatigue, and job strain. Resources refer to aspects of work that are functional in achieving work goals, reduce job stressors, and are associated with positive outcomes, such as work engagement, job satisfaction, and well-being. As long as employees have sufficient resources, they can compensate for some stressors; however, if this balance is disrupted, stressors often lead to dissatisfaction, turnover intentions, and negative safety outcomes in the short term, as well as health detriments and sickness absence in the long term (e.g., Bakker et al., 2023). By focusing on job characteristics as stressors and resources, organizations can develop strategies to reduce stress and promote employee well-being. This might include adjusting workloads, providing opportunities for skill-building and career advancement, and creating a supportive work environment that values social connections and work-life balance. It is also important for organizations to consider a broader range of job characteristics when developing strategies to promote occupational health and well-being and to create a positive and inclusive work culture that values employee well-being. In particular, if organizations are

interested in taking steps to reduce stressors and strengthen resources, a good starting point will be the assessment of work-related stressors and resources.

A large number of questionnaires have already been developed to assess stressors and resources at work (e.g., Kristensen et al., 2005; Morgeson & Humphrey, 2006). In many cases, these are self-reports of the frequency or intensity of stressors and resources evaluated on a five- to seven-point Likert scale. Self-reports are popular in this context because they are convenient and economical for organizations. Furthermore, employees likely know their workplace better than any other potential rater (Semmer et al., 1995). Typical stressors measured with such questionnaires are time pressure, concentration demands, and one-sidedness of physical stressors; typical resources that are surveyed are task control, support from colleagues or supervisors, and development opportunities. Many of these questionnaire items are phrased in general terms and can therefore be applied to many different workplaces. In addition to the causes of psychological strain, some questionnaires also include the potential consequences of stressors such as exhaustion, turnover intentions, and psychosomatic complaints (e.g., Kristensen et al., 2005).

Self-report questionnaires are widely used in psychology, but debates about measurement issues are prevalent in the literature (e.g., Einola & Alvesson, 2021). One challenge identified within this debate is the failure to consider the continuum that defines theoretical constructs, which can lead to incomplete measurement and compromised validity evidence. For example, Tay and Jebb (2018) propose a continuum specification approach to create an appropriate measure. Another challenge is to determine which kind of scaling is appropriate. For example, the General Health Questionnaire's (GHQ) traditional binary scoring system fails to distinguish between individuals with chronic symptoms and those without psychological distress (Whaley

et al., 2005). Scaling issues matter even more if researchers create ratio variables as, for instance, in the case of effort-reward imbalance (Lang et al., 2020). Furthermore, there is an ongoing debate about whether stress researchers should control for personality. This debate tends to center around the trait of negative affectivity (e.g., Debus et al., 2015): Whereas some researchers recommend controlling for NA bias, Spector and colleagues (2000) argue that NA should not be routinely partialled out, as it may play a substantial role in the job stress process. In addition, stress surveys might be affected by response styles such as acquiescence (e.g., Lee et al., 2020). This study aims to contribute to these debates on measurement issues in the well-being literature by providing a better understanding of work stress measurement with particular reference to social comparison processes.

**Social Comparisons in Self-Reports**

Accurately completing self-reports can be a challenge, and it requires several cognitive processes for the respondents (Greulich et al., 2021; Tourangeau, 1987). Respondents have to understand the question asked of them, to activate mental representations of their job, and to compare these representations to the scale anchors. Therefore, they can experience some uncertainty about whether “sometimes” is the right answer to a question like “Can you decide for yourself how you want to do your job?”.

In such a situation of uncertainty, Festinger's (1954) theory of social comparisons predicts that answers are based on a comparison with others. Generally speaking, uncertainty about aspects of the self (e.g., abilities, traits, and feelings) increases the probability that social comparison processes will be used to make an evaluation (Festinger, 1954; Gerber et al., 2018). In most cases, the use of social comparison processes for assessment happens unintentionally and is not a conscious

decision. Therefore, self-perceptions can be influenced by unaware cognitive processes like social comparisons (Heine et al., 2002). The use of social comparison processes by employees when filling out work stress surveys may be unproblematic when measuring subjective stress levels. However, a potential problem could arise when many very different comparison objects are used to assess one's stress level. As Greulich et al. (2021) found, the comparison objects vary significantly both within a person and between individuals. As a result, interindividual and intraindividual differences may exist when using individual social comparisons to answer survey items. The inconsistent use of social comparison processes when answering questionnaires could thus affect the measurement accuracy of questionnaires. Consequently, the addition of a specified comparison object to standardize the social comparison process may positively impact psychometric properties, such as the reliability and validity, of work stress surveys.

Research on personality questionnaires has already shown that standardization of a frame-of-reference has positive effects on measurement accuracy (e.g., Lievens et al., 2008), and researchers in cross-cultural psychology who have been interested in understanding answers to self-reports in different cultures (e.g., Heine et al., 2008; Peng et al., 1997; Song et al., 2019) found that respondents self-rating on a Likert scale is influenced by social comparison, which they refer to as the reference-group effect. For instance, when Chinese people rate their level of individualism on a Likert scale they relate their response to the level of individualism of a salient reference group of other Chinese people and not on their absolute level of individualism (Heine et al., 2002). There is also preliminary evidence that this reference-group effect can be found in the assessment of personality: Credé and colleagues (2010) added four versions of an explicit reference group (e.g., their immediate family and people of the

same age and gender) on items of a personality inventory. These explicit comparisons substantially influenced personality scores.

### Social Comparisons in Work-Related Stressor and Resource Scales

Festinger's (1954) social comparison theory can also be applied to self-reports of work-related stressors and resources. Although this argument has not been empirically tested, it has been mentioned by several authors. For example, Semmer and colleagues (1995) noted that the perception of stressors and resources at work may be influenced by "social comparisons like 'I am much better (worse) off than many of my colleagues'" (p. 105), and Buunk and Gibbons (2007) stated that they expect people to compare themselves more often in situations of high stressors. In addition, the results of a recent qualitative study indicate that employees use social comparison processes when filling out work stress surveys. Participants reported that they used different types of social comparisons (e.g., comparisons with direct co-workers, with people of their organization but with different tasks, and with different professions), that partly vary depending on the item (Greulich et al., 2021). Furthermore, colleagues tend to compete against each other when there are few opportunities for promotion, status, and recognition (Baumann et al., 2019; Prendergast, 1999), and situations that foster competition are likely to promote interest in social comparison for many people (Garcia et al., 2013).

### Effects of Item Standardization on Reliability and Validity of Work-Related Stressor and Resource Scales

Individuals tend to employ diverse social comparisons when completing work stress questionnaires (Goodman & Haisley, 2007; Greulich et al., 2021). On the one hand, this phenomenon pertains to the individual respondent, whose choice of reference group may vary depending on the item being evaluated, thereby resulting in

a greater degree of intraindividual variability. On the other hand, it applies to all respondents, as each individual engages in individualized comparisons, ultimately resulting in a greater degree of interindividual variability. To address this issue and standardize the interpretation of items, it has been proposed in the literature on personality questionnaires that items should be standardized (Schmit et al., 1995). While items lacking a context are more open to interpretation (Robie et al., 2000), by standardizing an item through the inclusion of a specific context (e.g., "at school" added to every personality item), respondents are expected to make consistent contextual references for each item, thus reducing intraindividual variance. Consequently, this reduction in variance is assumed to increase the reliability of the inventory, a proposition that has received empirical support in prior personality studies (Lievens et al., 2008; Reddock et al., 2011; Swift & Peterson, 2019).

Applying this concept to work stress surveys and considering respondents' inclination towards individual social comparisons, it is anticipated that item standardization through the inclusion of a social comparison object will reduce intraindividual variance in scales assessing work-related stressors and resources resulting in higher reliability. To examine this hypothesis, we have selected direct colleagues as the comparison group, as employees are presumed to be familiar with their colleagues to a relatively high degree, thus making such comparisons relatively effortless for respondents. It can be formally stated as follows:

> *Hypothesis 1: The reliabilities of work-related stressor and resource scales with an explicit comparison to a direct colleague are higher than the reliabilities of work-related stressor and resource scales without an explicit comparison to a direct colleague.*

According to both classical and modern test theory, an increase in reliability leads to a corresponding increase in validity (e.g., Crocker & Algina, 1986). As a result of standardizing the social comparison process, it can be expected that all respondents will similarly approach the items, leading to a reduction in interindividual variance. This, in turn, could be accompanied by an increase in validity. Empirical studies investigating contextualized personality items have supported this argument (e.g., Lievens et al., 2008; Swift & Peterson, 2019; Voss et al., 2023). The current approach aims to improve criterion validity by standardizing the items in work stress surveys through the incorporation of a specific social comparison (in this case, a direct colleague). Although the use of the term "close colleague" may still lead to individual variation in the selection of comparison objects, it is anticipated that this specification will nevertheless result in standardized item interpretation, as respondents are instructed to compare themselves to an individual with the same job responsibilities in the same work environment. Consequently, while the specific colleague may vary across respondents, the comparison object meets the same criteria for all respondents.

However, there is also evidence suggesting that the use of reference groups in stress questionnaires may not necessarily improve criterion validity. Specifically, Credé et al. (2010) found that the more specific the reference group, the lower the criterion validity. This suggests that individuals may have different perceptions of the specified reference group, leading to distinct scaling standards and potential bias in results. The greater the degree of specificity with which a comparison object is defined, the greater the likelihood that idiosyncratic individual characteristics such as behavior patterns or salient personality traits will shape the social comparison process. Such distinctive attributes of the comparison object can give rise to unique

perceptions of the object that may vary significantly across individuals. Therefore, comparisons with a direct colleague could result in responses that are too specific referring to the content of the construct, which might reduce validity. Given that there are arguments for both sides, we formulate a research question.

*Research Question 1: What is the impact of adding an explicit social comparison in stressor and resource items on the validity of work stress surveys?*

**The Direction of Social Comparisons in Work-Related Stressor and Resource Scales**

In addition to the underlying motivations that prompt individuals to engage in social comparisons, the direction of the comparison represents a significant distinction within the existing body of research on social comparisons. Comparisons can be classified as either upward (where the target of comparison is perceived as superior) or downward (where the target of comparison is perceived as inferior). Downward comparisons have been found to evoke more positive emotions by enhancing self-esteem, whereas upward comparisons tend to elicit more negative emotions (Brickman & Bulman, 1977; Buunk & Gibbons, 2007; Taylor et al., 1996; Wood, 1989). Depending on the specific context, an upward comparison can also serve as a source of motivation for individuals to improve their situation (Collins, 1996; Taylor & Lobel, 1989). While Festinger's (1954) theory does not explicitly address the direction of social comparisons (i.e., whether individuals engage in downward or upward comparisons), it is pertinent to consider this aspect within the context of work-related stress.

The Stress as Offense to Self theory (SOS, Semmer et al., 2007, 2019) offers a framework that can provide insights into upward and downward social comparisons in

the context of work-related stress. According to the SOS theory, stress arises when individuals perceive a threat to their sense of self. This occurs when demands at work (or in other life domains) exceed the individual's ability to cope, and when the individual perceives these demands as a challenge to their personal values or competencies. The SOS theory posits that stress is not solely the result of external demands or objective conditions, but is also influenced by subjective perceptions and evaluations. When individuals experience stress, they are motivated to restore their sense of self by reducing the perceived threat to their self-image.

According to the SOS theory (Semmer et al., 2007, 2019), employees strongly identify with their professional roles, which become integral to their identity and self-image. The maintenance of a positive personal and social self-image is considered a fundamental psychological need, and job performance and success assessments are particularly relevant to personal self-esteem. High levels of stressors at work may indicate that the employee is overwhelmed with tasks and lacks resources or support, leading to feelings of diminished worth and threatening self-esteem. As such, experiencing high stressors and low resources could undermine professional identity, and one potential coping strategy is to present oneself as being less stressed and possessing greater resources in comparison to a colleague. Consequently, employees may engage in downward social comparisons when evaluating their stressors and resources, comparing themselves to individuals with worse work conditions to maintain a positive self-image. At the same time, employees experiencing high stressors and low resources may fear that their performance will decline in response to stress, which could be observed by others. To protect their self-esteem and justify any performance drops, they may resort to upward comparisons, rating their own stressors as much higher and their resources as much less pronounced than is the case with

their colleagues, suggesting that their heightened stress is a result of having more responsibilities or tasks compared to their colleagues. Based on the argument presented by the SOS theory (Semmer et al., 2007, 2019), it is possible that employees use social comparisons to protect their self-esteem when facing high work stress. If the social comparison is predominantly downward directed, employees may report lower average stressors and higher average resources compared to a group without an explicit comparison. Conversely, if the social comparison is predominantly upward directed, employees may report higher average stressors and lower average resources. It remains unclear whether an explicit social comparison, specifically with colleagues in our case, is more likely to trigger a downward or an upward comparison. Thus, the following research question is formulated:

> *Research Question 2: Does the incorporation of an explicit comparison in work stress survey items elicit a predominantly downward comparison, resulting in lower mean stressor values and higher mean resource values relative to the control group, or an upward comparison, leading to higher mean stressor values and lower mean resource values relative to the control group?*

## Method

### Participants

We recruited employees from nine rehabilitation hospitals in Germany by sending out 600 paper-and-pencil questionnaires to employees. The final sample consisted of 208 employees (i.e., a response rate of 34.7%) and was mostly female (78.8%). The majority of the participants (49.5%) were over 50 years old, 24.5% were between 41 and 50 years old, 16.7% were 30 to 40 years old, and 9.3% were under 30 years old. The participants' professional fields primarily consisted of facility

management (30.0%) and healthcare (21.0%). Other fields were psychotherapy (20.7%), public administration (13.5%), and childcare (14.8%). The control ($n$ = 105) and experimental ($n$ = 103) groups did not differ in terms of gender, $F(1, 183) = .006$, $p = .940$, age, $F(1, 183) = .740$, $p = .391$, language, $F(1, 183) = .06$, $p = .806$, professional fields, $F(1, 183) = 2.02$, $p = .157$, and job satisfaction, $F(1, 183) = .510$, $p = .476$. Participation was voluntary.

**Procedure and Experimental Manipulation**

In a between-subjects design, participants were randomly assigned to one of two groups: The control group ($n$ = 105) was given the original questions without changes in the wording (e.g., “How often are you under time pressure?”) and instructed to evaluate the stressors and resources in their workplace; The items of the experimental group ($n$ = 103) were manipulated by adding the wording “in comparison to your close colleague” at the end of the same items (e.g., “How often are you under time pressure compared to your close colleague?”). Participants placed their questionnaires in a ballot box that was emptied after approximately three weeks, and the questionnaires were then sent back to us.

**Measures**

***Work Stress Survey***

Four scales were used to assess typical stressors: time pressure, work environment, one-sided physical stressors, and work-life balance. Two scales were used to assess typical resources at work: task control and development opportunities/meaning of work. The scales were taken from two common German work stress surveys: the “Instrument for Stress-Oriented Task Analysis” (ISTA; Irmer et al., 2019; Semmer et al., 1999) and the “Copenhagen Psychosocial Questionnaire” (COPSOQ; Kristensen et al., 2005; Nübling et al., 2005). The criteria for selection

were that the stressors and resources should be relevant for all occupational groups at a hospital and that adding the comparison should be feasible without changing items. The items measure stressors and resources either in terms of frequency or intensity on a five-point Likert scale ranging from 1 (disagree strongly/never) to 5 (agree strongly/always; for details see Table 1). However, the scales "one-sided physical stressors" and "work-life balance" yielded low reliability (Cronbach's $\alpha = .23$ and .43 respectively) and were therefore excluded from further analysis.

### *Job Satisfaction*

To test the effects of an explicit comparison with colleagues on validity, we used job satisfaction as a criterion, assessed with the Neuberger and Allerbeck (1978) job satisfaction scale, which is a widely used job satisfaction scale in Germany (see e.g., Maier & Brunstein, 2001). Participants rate satisfaction with their job, working conditions, the relationship with their supervisor, the relationship with their colleagues, promotion opportunities, organization and management, and benefits and pay. The participants indicate their level of agreement with the seven items on a five-point Likert scale ranging from 1 (disagree strongly) to 5 (agree strongly).

### *Control Variables*

Gender and age have been shown to influence ratings of stressors and resources at work (e.g., Matud, 2004; Shultz et al., 2010; Watson et al., 2011), and we thus controlled for both variables.

### *Stress Mindset*

To additionally check whether participants perceive stress as something rather negative, their general attitude toward stress was assessed with the following four items from the Stress Mindset Measure (Crum et al., 2013, in its German translation, Schollmeyer, 2004): (1) "The effects of stress are negative and should be avoided.",

(2) “Experiencing stress increases my performance and productivity.”, (3) ”Experiencing stress affects my health and vitality.”, (4) “The effects of stress are positive and should be used.” Respondents indicate their level of agreement with a five-point Likert scale ranging from 1 (disagree strongly) to 5 (agree strongly), and items (2) and (4) are reverse coded. Therefore, a high score indicates a negative attitude towards stress.

**Table 1**

*Work-Related Stressor and Resource Items, Chosen From the Instrument of Stress-Oriented Task Analysis (Semmer et al., 1999) and Copenhagen Psychosocial Questionnaire (Nübling et al., 2005)*

| Scale | Item | Inventory (Number) |
|---|---|---|
| Stressors | | |
| Time pressure | How often are you pressed for time? | ISTA (ZD1) |
| | How often must you finish work later because you have too much to do? | ISTA (ZD4) |
| | How often is a fast pace of work required of you? | ISTA (ZD6) |
| Work environment | How often do you have to do physically strenuous work? | COPSOQ (B8b-1) |
| | How often are you exposed to noise or loud background noise at your workplace? | COPSOQ (B8b-2) |
| | How often do you come in contact with chemicals or hazardous substances at your work? | COPSOQ (B8b-3) |
| | How often are you exposed to extreme temperatures or a draft at your workplace? | COPSOQ (B8b-4) |
| One-sided physical stressors | How often does your work activity require you to bend from the waist? | ISTA (EBA2) |
| | How often does your work activity require you to take a twisted or unusual posture? | ISTA (EBA3) |
| | How typical is sitting for long periods of time in your work activity? | ISTA (EBA4) |
| | How typical is standing for long periods of time in your work activity? | ISTA (EBA5) |

| | | |
|---|---|---|
| Work-life balance | I take care of work-related tasks outside of my working time as well. | COPSOQ (B2-5) |
| | I'm available in my free time for people I deal with professionally. | COPSOQ (B2-6) |
| Resources | | |
| Task control | Considering your workplace in general, how much can you change the sequence of your different tasks yourself? | ISTA (HS1) |
| | How much influence do you have on the work that is assigned to you? | ISTA (HS2) |
| | Considering your work activity in general, how much opportunity is there for you to make your own decisions? | ISTA (HS3) |
| | Can you yourself decide on which way to carry out your work? | ISTA (HS4) |
| Development opportunities | Do you have the possibility of learning new things through your work? | COPSOQ (B5-1) |
| | Can you use your skills or expertise in your work? | COPSOQ (B5-2) |
| | Is your work meaningful? | COPSOQ (B5-3) |

## Results

### Preliminary Analyses

Participants had on average a fairly negative attitude towards stress, as indicated by high Stress Mindset Measurement scale values in both groups (the group with a colleague comparison: $M = 4.11$, $SD = 0.82$; the group without a colleague

comparison: $M = 4.17$, $SD = 0.75$). The two means were not significantly different from each other, $t(206) = .52$, $p = .60$).

**Test of the Hypothesis and the Research Questions**

To test for reliability differences, we calculated Cronbach's α for each scale and each condition and compared them using the computer program Alpha Test (Lautenschlager & Meade, 2008). Contrary to Hypothesis 1, Cronbach's αs did not significantly differ between groups..

**Table 2**

*Reliabilities for Each Group*

| Scale | Cronbach's α | | | |
|---|---|---|---|---|
| | Without comparison ($n$ = 105) | With comparison ($n$ = 103) | $\chi^2$ | $p$ |
| Stressors | | | | |
| Time pressure | .80 | .74 | .90 | .34 |
| Work environment | .82 | .79 | .27 | .61 |
| Resources | | | | |
| Task control | .88 | .87 | .20 | .66 |
| Development opportunities | .77 | .81 | .44 | .51 |

*Note*. The $\chi^2$ and $p$ values describe the significance of the reliability differences.

Table 3 shows how well the stressors and resources with and without the comparison predicted job satisfaction. Correlation coefficients were compared for statistical significance (Lautenschlager & Meade, 2008). As shown in Table 5, only

the ‘work environment’ stressor scale predicted job satisfaction better with social comparison vs. without comparison.

**Table 3**

*Criterion Validities for Each Group*

| Scale | Correlations with job satisfaction | | |
|---|---|---|---|
| | Without comparison ($n$ = 105) | With comparison ($n$ = 103) | $p$ |
| Stressors | | | |
| Time pressure | -.38 | -.30 | .24 |
| Work environment | -.25 | -.47 | .03* |
| Resources | | | |
| Task control | .30 | .37 | .32 |
| Development opportunities | .42 | .36 | .32 |

*Note*. The $p$ values describe the significance of the differences between correlations.

*$p < .05$.

Because previous literature suggests that age and gender influence the assessment of stressors and resources, these variables were considered as covariates. The assumptions of a MANCOVA were met: An ANOVA showed the independence of the covariates and the treatment effect with $F(1, 202) = 0.96$, $p = .33$, for age and $F(1, 206) = 0.56$, $p = .46$, for gender. Furthermore, the assumption of homogeneity of regression slopes could be accepted for the stressor and resource scales. The means, adjusted means, standard deviations, and standard errors for the work-related stressor and resource scales are shown in Table 4.

**Table 4**

*Means, Adjusted Means, Standard Deviations, and Standard Errors of the Work-Related Stressor and Resource Scales for the two Groups*

| Scale | | Without comparison ($n$ = 105) | With comparison ($n$ = 103) |
|---|---|---|---|
| Stressors | | | |
| Time pressure | *M* | 3.29 | 3.05 |
| | *(SD)* | (0.80) | (0.82) |
| | $M_{adj}$ | 3.29 | 3.05 |
| | *(SE)* | (0.08) | (0.08) |
| Work environment | *M* | 2.56 | 2.25 |
| | *(SD)* | (1.17) | (0.94) |
| | $M_{adj}$ | 2.56 | 2.25 |
| | *(SE)* | (0.10) | (0.11) |
| Resources | | | |
| Task control | *M* | 2.50 | 2.26 |
| | *(SD)* | (1.00) | (0.87) |
| | $M_{adj}$ | 2.50 | 2.27 |
| | *(SE)* | (0.09) | (0.09) |
| Development opportunities | *M* | 2.43 | 2.78 |
| | *(SD)* | (0.99) | (1.01) |
| | $M_{adj}$ | 2.43 | 2.79 |
| | *(SE)* | (0.10) | (0.10) |

*Note*. Adjustments for age and gender as control variables.

**Table 5**

*Results of the MANCOVAs of the Stressor and Resource Scales and Their Respective Covariates*

| Scale | *F* | *p* | Partial $\eta^2$ |
|---|---|---|---|
| Stressors | | | |
| Time pressure | 4.60 | < .05* | .02 |
| Work environment | 4.14 | < .05* | .02 |
| Age | 0.73 | .48 | .01 |
| Gender | 0.41 | .67 | .00 |
| Resources | | | |
| Task control | 3.08 | .08 | .02 |
| Development opportunities | 6.65 | < .05* | .03 |
| Age | 0.11 | .90 | .00 |
| Gender | 2.46 | .09 | .02 |

*Note*. *$p < .05$.

The MANCOVA for the stressor scales (see Table 5) showed a significant overall difference between the groups with and without explicit social comparison, $F(2, 199) = 3.44$, $p < .05$. The tests of between-subjects effects were significant for both time pressure, $F(1, 200) = 4.60$, $p < .05$, and work environment, $F(1, 200) = 4.14$, $p < .05$. Therefore, the group with colleague comparisons had lower average stressor scores than the group without comparisons, supporting a downward comparison.

The MANCOVA for resources (see Table 5) showed a significant overall difference between the groups with and without explicit social comparison, $F(2, 199) = 11.67$, $p < .001$. The test of between-subjects-effects was significant for development opportunities, $F(1, 200) = 6.65$, $p < .05$, with a higher mean for the group with comparisons compared to the group without comparisons. The effect on task control did not reach conventional significance levels (and was descriptively even opposite to the hypothesized direction). Accordingly, these data do not allow a clear to be drawn.

## Discussion

The purpose of this study was to investigate the impact of standardized social comparisons on the evaluation of work-related stressors and resources. Specifically, we tested the hypothesis that contextualizing items through the inclusion of a defined comparison object would affect the reliability of the measurement. Furthermore, as part of a research question, we were interested in investigating the impact of integrating a standardized social comparison process on the validity of assessing work-related stress. Although contextualization did not have a significant impact on reliability, it did improve the validity of one stressor scale by reducing interindividual error variance. Regarding the research question of the directionality of social comparison processes, the results indicated that for stressors, the group that included colleague comparisons had lower average stressor scores in comparison to the group without comparisons. Conversely, for resources, higher mean values were observed only for the development opportunities scale. Overall, our findings indicate that the standardization of social comparison in work-related stressors and resources had a discernible influence on the evaluation. However, the effects on the psychometric variables were not as substantial as anticipated. Nevertheless, this

approach presents a potential alternative to previous measurement techniques that requires further elucidation.

The standardized social comparison influenced the means of the scales *time pressure, work environment,* and *development opportunities*. This finding aligns with the results of Credé et al. (2010), who found that mean values of a scale for personality vary depending on the reference-group. Going beyond previous research in the personality realm, we derived from SOS theory (Semmer et al., 2007, 2019) that stress, if seen as something negative, threatens self-esteem and that employees tend to protect it. Our data show that employees have negative attitudes toward stress and tend to take downward comparisons to protect their self-esteem: Lower stressor values and higher resource values were observed in the group with peer comparison.

Based on studies in the personality domain (e.g., Reddock et al., 2011), we expected that the addition of a context in the form of a social comparison should increase scale reliability because the unification of the context should reduce *intra*individual variance (Lievens et al., 2008). However, the data did not support this hypothesis, which implies a difference between personality test items and stressor/resources items: Whereas people might consider different comparison targets when filling out (e.g., a friend for extraversion item no. 1 vs. the sister for extraversion item no. 2), people filling out stressor and resources items do not seem to vary strongly enough in their comparison targets between items to make this measurable in the context of this study.

The incorporation of a targeted comparison was anticipated to improve the validity of the evaluation. However, it was found to only have a significant impact on one scale, namely the work environment scale. This finding may support the theoretical argument that standardizing the context can reduce interindividual error

variance (Lievens et al., 2008). In the absence of an explicit comparison, employees may compare themselves to other occupational groups, family members, or colleagues (see Greulich et al., 2021), which increases interindividual error variance.

**Limitations and Future Research**

As with any study, this work is subject to limitations. The choice of items from common work stress questionnaires was restricted by the manipulation (e.g., because items already had a different kind of comparison in their formulation, or their content did not allow us to add a comparison). Thus, the pool of potential scales available in the end was quite limited, and some scales consisted of only two to four items, which was likely the reason why their reliabilities were too low to use. Ideally, future research is able to work with other scales. Although the sample size in this study was rather small and maybe not be sufficient to detect true small differences in two independent correlations (Vanhove, 2015), we considered the investigation of the influence of social comparison processes on reliability and validity appropriate to get first insights into possible effects. A replication with a larger sample would also allow the use of other statistical techniques to investigate possible biases (e.g., confirmatory factor analyses or differential item analyses based on item response theory, Somaraju et al., 2022; Thissen et al., 1993, but see also Robitzsch & Lüdtke, 2022). Furthermore, all participants were hospital staff, and the majority of the participants were female. Although a female majority is typical for German hospital workers (see Brehm et al., 2021), future research should try to replicate these findings with other samples of employees. Care should then also be taken to ensure that the gender variable is more equally distributed than in our sample. Even though this study did not show any differences depending on gender, there are indications that men and women differ in their use of downward comparisons to protect themselves (Kemmelmeier &

Oyserman, 2001). In addition, the effects of adding a social comparison in this study were not as strong as initially expected, maybe because participants worked in hospitals where working conditions may not be as competitive as they would be in other sectors. Further research should also investigate whether there are other comparisons that are also relevant in the context of work-related stressors and resources. For example, Greulich et al. (2021) mentioned that employees might also compare themselves to other professions or to family members – or to themselves in a previous job.

### Practical Implications

This study implies that organizations should pay more attention to the details of administration when collecting survey data about workplace stressors and resources. For example, the text introducing a survey might trigger certain comparisons (or only certain comparisons for some employees but not others). Organizations could also attempt to standardize such comparison processes by adding an explicit comparison to the introductory text (or to each item). Furthermore, organizations should become more aware that data from standard stress surveys are also influenced by comparison processes. Thus, before taking such data too literally and before using it to initiate actions (e.g., workplace changes for a certain unit), it is likely beneficial to reach out to this unit and its employees and to try to understand the situation and the perception of the situation in this unit, maybe by using qualitative interviews.

### Conclusion

In this study, we have contributed to the ongoing discourse surrounding measurement issues in the field of work stress research, which has a long-standing tradition of discussion and debate (e.g., Lang et al., 2020; Whaley et al., 2005).

Through our examination of work stress measurement with a particular focus on social comparison processes, we have gained insights into the complexities of measuring subjective experiences in the workplace. A notable finding of this study is that the inclusion of a comparison variable can influence the means of work stress survey scales, potentially reducing interindividual error variance. The reason why the reduction was not observed for all scales requires further investigation. Therefore, this study encourages researchers to continue to engage in these measurement debates and to develop innovative approaches to measurement that are better aligned with theoretical constructs, and thereby advance our understanding of the complex nature of human experiences in the workplace.

**References**

Bakker, A. B., Demerouti, E., & Sanz-Vergel, A. (2023). Job demands-resources theory: Ten years later. *Annual Review of Organizational Psychology and Organizational Behavior*, *10*(1), 25–53. https://doi.org/10.1146/annurev-orgpsych-120920-053933

Baumann, O., Eggers, J. P., & Stieglitz, N. (2019). Colleagues and competitors: How internal social comparisons shape organizational search and adaptation. *Administrative Science Quarterly*, *64*(2), 275–309. https://doi.org/10.1177/0001839218766310

Brehm, T. T., Schwinge, D., Lampalzer, S., Schlicker, V., Küchen, J., Thompson, M., Ullrich, F., Huber, S., Schmiedel, S., Addo, M. M., Lütgehetmann, M., Knobloch, J. K., Schulze zur Wiesch, J., & Lohse, A. W. (2021). Seroprevalence of SARS-CoV-2 antibodies among hospital workers in a German tertiary care center: A sequential follow-up study. *International Journal of Hygiene and Environmental Health*, *232*, 113671. https://doi.org/10.1016/j.ijheh.2020.113671

Brickman, P., & Bulman, R. J. (1977). Pleasure and pain in social comparison. In L. M. Sulsky & R. L. Miller (Eds.), *Social comparison processes: Theoretical and empirical perspectives.* (pp. 149–186). Hemisphere.

Buunk, A. P., & Gibbons, F. X. (2007). Social comparison: The end of a theory and the emergence of a field. *Organizational Behavior and Human Decision Processes*, *102*(1), 3–21. https://doi.org/10.1016/j.obhdp.2006.09.007

Collins, R. L. (1996). For better or worse: The impact of upward social comparison on self-evaluations. *Psychological Bulletin*, *119*(1), 51–69. https://doi.org/10.1037/0033-2909.119.1.51

Credé, M., Bashshur, M., & Niehorster, S. (2010). Reference group effects in the measurement of personality and attitudes. *Journal of Personality Assessment*, *92*(5), 390–399. https://doi.org/10.1080/00223891.2010.497393

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Holt Rinehart & Winston.

Crum, A. J., Salovey, P., & Achor, S. (2013). Rethinking stress: The role of mindsets in determining the stress response. *Journal of Personality and Social Psychology*, *104*(4), 716–733. https://doi.org/10.1037/a0031201

de Jonge, J., van Vegchel, N., Shimazu, A., Schaufeli, W., & Dormann, C. (2010). A longitudinal test of the demand-control model using specific job demands and specific job control. *International Journal of Behavioral Medicine*, *17*(2), 125–133. https://doi.org/10.1007/s12529-010-9081-1

Debus, M. E., König, C. J., Kleinmann, M., & Werner, C. S. (2015). Examining the effects of negative affectivity on self- and supervisor ratings of job stressors: The role of stressor observability. *Work & Stress*, *29*(4), 341–361. https://doi.org/10.1080/02678373.2015.1075233

Demerouti, E., Bakker, A. B., Nachreiner, F., & Schaufeli, W. B. (2001). The job demands-resources model of burnout. *Journal of Applied Psychology*, *86*(3), 499–512. https://doi.org/10.1037/0021-9010.86.3.499

Einola, K., & Alvesson, M. (2021). Behind the numbers: Questioning questionnaires. *Journal of Management Inquiry*, *30*(1), 102–114. https://doi.org/10.1177/1056492620938139

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*(2), 117–140. https://doi.org/10.1177/001872675400700202

Garcia, S. M., Tor, A., & Schiff, T. M. (2013). The psychology of competition: A social comparison perspective. *Perspectives on Psychological Science*, *8*(6), 634–650. https://doi.org/10.1177/1745691613504114

Gerber, J. P., Wheeler, L., & Suls, J. (2018). A social comparison theory meta-analysis 60+ years on. *Psychological Bulletin*, *144*(2), 177–197. https://doi.org/10.1037/bul0000127

Goodman, P. S., & Haisley, E. (2007). Social comparison processes in an organizational context: New directions. *Organizational Behavior and Human Decision Processes*, *102*(1), 109–125. https://doi.org/10.1016/j.obhdp.2006.10.005

Greulich, B., Debus, M. E., Kleinmann, M., & König, C. J. (2021). Response behavior in work stress surveys: A qualitative study on motivational and cognitive processes in self- and other-reports. *European Journal of Work and Organizational Psychology*, *30*(1), 40–57. https://doi.org/10.1080/1359432X.2020.1812580

Heine, S. J., Buchtel, E. E., & Norenzayan, A. (2008). What do cross-national comparisons of personality traits tell us? The case of conscientiousness. *Psychological Science*, *19*(4), 309–313. https://doi.org/10.1111/j.1467-9280.2008.02085.x

Heine, S. J., Lehman, D. R., Peng, K., & Greenholtz, J. (2002). What's wrong with cross-cultural comparisons of subjective Likert scales? The reference-group effect. *Journal of Personality and Social Psychology*, *82*(6), 903–918. https://doi.org/10.1037//0022-3514.82.6.903

Irmer, J. P., Kern, M., Schermelleh-Engel, K., Semmer, N. K., & Zapf, D. (2019). The Instrument for Stress-Oriented Task Analysis (ISTA): A meta-analysis. *Zeitschrift für Arbeits- und Organisationspsychologie*, *63*(4), 217–237. https://doi.org/10.1026/0932-4089/a000312

Johnson, J. V., & Hall, E. M. (1988). Job strain, work place social support, and cardiovascular disease: A cross-sectional study of a random sample of the Swedish working population. *American Journal of Public Health*, *78*(10), 1336–1342. https://doi.org/10.2105/ajph.78.10.1336

Karasek, R. A. (1979). Job demands, job decision latitude, and mental strain: Implications for job redesign. *Administrative Science Quarterly*, *24*(2), 285–308. https://doi.org/10.2307/2392498

Karasek, R. A., & Theorell, T. (1990). *Healthy work: Stress productivity and reconstruction of working life*. Basic Books.

Kemmelmeier, M., & Oyserman, D. (2001). Gendered influence of downward social comparisons on current and possible selves. *Journal of Social Issues*, *57*(1), 129–148. https://doi.org/10.1111/0022-4537.00205

Kristensen, T. S., Borritz, M., Villadsen, E., & Christensen, K. B. (2005). The Copenhagen Burnout Inventory: A new tool for the assessment of burnout. *Work & Stress*, *19*(3), 192–207. https://doi.org/10.1080/02678370500297720

Lang, J. W. B., Van Hoeck, S., & Runge, J. M. (2020). Methodological and conceptual issues in studying effort-reward fit. *Journal of Managerial Psychology*, *37*(5), 498–512. https://doi.org/10.1108/JMP-11-2019-0659

Lautenschlager, G. J., & Meade, A. W. (2008). AlphaTest: A Windows program for tests of hypotheses about coefficient alpha. *Applied Psychological Measurement*, *32*(6), 502–503. https://doi.org/10.1177/0146621607312307

Lee, S., Vasquez, E., Ryan, L., & Smith, J. (2020). Measurement equivalence of subjective well-being scales under the presence of acquiescent response style for the racially and ethnically diverse older population in the United States. *Survey Research Methods*, *14*(4), Article 4. https://doi.org/10.18148/srm/2020.v14i4.7413

Lievens, F., de Corte, W., & Schollaert, E. (2008). A closer look at the frame-of-reference effect in personality scale scores and validity. *Journal of Applied Psychology*, *93*(2), 268–279. https://doi.org/10.1037/0021-9010.93.2.268

Maier, G. W., & Brunstein, J. C. (2001). The role of personal work goals in newcomers' job satisfaction and organizational commitment: A longitudinal analysis. *Journal of Applied Psychology*, *86*(5), 1034–1042. https://doi.org/10.1037/0021-9010.86.5.1034

Matud, M. P. (2004). Gender differences in stress and coping styles. *Personality and Individual Differences*, *37*(7), 1401–1415. https://doi.org/10.1016/j.paid.2004.01.010

Morgeson, F. P., & Humphrey, S. E. (2006). The Work Design Questionnaire (WDQ): Developing and validating a comprehensive measure for assessing job design and the nature of work. *Journal of Applied Psychology*, *91*(6), 1321–1339. https://doi.org/10.1037/0021-9010.91.6.1321

Neuberger, O., & Allerbeck, M. (1978). *Messung und Analyse von Arbeitszufriedenheit: Erfahrungen mit dem "Arbeitsbeschreibungsbogen (ABB)" [Measurement and analysis of job satisfaction: experiences with the "Job Description Questionnaire (ABB)"]*. Hans Huber.

Nübling, M., Stößel, U., Hasselhorn, H. M., Michaelis, M., & Hofmann, F. (2005). COPSOQ: Ein Instrument zur Messung psychischer Belastungen am Arbeitsplatz [COPSOQ: An instrument for measuring mental stress in the workplace]. *Gesundheitswesen*, *67*, V60. https://doi.org/10.1055/s-2005-920554

Peng, K., Nisbett, R. E., & Wong, N. Y. C. (1997). Validity problems comparing values across cultures and possible solutions. *Psychological Methods*, *2*(4), 329–344. https://doi.org/10.1037/1082-989X.2.4.329

Prendergast, C. (1999). The provision of incentives in firms. *Journal of Economic Literature*, *37*(1), 7–63. https://doi.org/10.1257/jel.37.1.7

Reddock, C. M., Biderman, M. D., & Nguyen, N. T. (2011). The relationship of reliability and validity of personality tests to frame-of-reference instructions and within-person inconsistency. *International Journal of Selection and Assessment*, *19*(2), 119–131. https://doi.org/10.1111/j.1468-2389.2011.00540.x

Robie, C., Schmit, M. J., Ryan, A. M., & Zickar, M. J. (2000). Effects of item context specificity on the measurement equivalence of a personality inventory. *Organizational Research Methods*, *3*(4), 348–365. https://doi.org/10.1177/109442810034003

Robitzsch, A., & Lüdtke, O. (2022). *Why measurement invariance is not necessary for valid group comparisons* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/cjyqp

Schmit, M. J., Ryan, A. M., Stierwalt, S. L., & Powell, A. B. (1995). Frame-of-reference effects on personality scale scores and criterion-related validity. *Journal of Applied Psychology*, *80*(5), 607–620. https://doi.org/10.1037/0021-9010.80.5.607

Schollmeyer, M. (2004). *Erstellung und Validierung der deutschen Version des Stress Mental Measure Testinventars [Develepment and validation of the German version of the Stress Mental Measure Inventory]* [Bachelorarbeit]. Saarland University.

Semmer, N. K., Jacobshagen, N., Meier, L. L., & Elfering, A. (2007). Occupational stress research: The "stress-as-offence-to-self" perspective. In J. Houdmont & S. McIntyre (Eds.), *Occupational health psychology: European perspectives on research, and practices* (Vol. 2, pp. 42–66). ISMAI.

Semmer, N. K., Tschan, F., Jacobshagen, N., Beehr, T. A., Elfering, A., Kälin, W., & Meier, L. L. (2019). Stress as Offense to Self: A promising approach comes of age. *Occupational Health Science*, *3*(3), 205–238. https://doi.org/10.1007/s41542-019-00041-5

Semmer, N. K., Zapf, D., & Dunckel, H. (1995). Assessing stress at work: A framework and an instrument. In O. Svane & C. Johansen (Eds.), *Work and health—Scientific basis of progress in the working environment* (pp. 105–113). Office for Official Publications of the European Communities.

Semmer, N. K., Zapf, D., & Dunckel, H. (1999). Instrument zur stressbezogenen Tätigkeitsanalyse ISTA [Instrument for Stress-Oriented Task Analysis ISTA]. In H. Dunckel (Ed.), *Handbuch psychologischer Arbeitsanalyseverfahren* (pp. 179–204). vdf Hochschulverlag.

Shultz, K. S., Wang, M., Crimmins, E. M., & Fisher, G. G. (2010). Age differences in the demand-control model of work stress: An examination of data from 15 European countries. *Journal of Applied Gerontology*, *29*(1), 21–47. https://doi.org/10.1177/0733464809334286

Somaraju, A. V., Nye, C. D., & Olenick, J. (2022). A review of measurement equivalence in organizational research: What's old, what's new, what's next? *Organizational Research Methods*, *25*(4), 741–785. https://doi.org/10.1177/10944281211056524

Song, H., Cramer, E. M., & Park, N. (2019). Cultural differences in social comparison on Facebook. *Behaviour & Information Technology*, *38*(2), 172–183. https://doi.org/10.1080/0144929X.2018.1519037

Spector, P. E., & Eatough, E. (2013). Quantitative self-report methods in occupational health psychology research. In M. Wang, S. Sinclair, & L. E. Tetrick (Eds.), *Research methods in occupational health psychology* (pp. 248–267). Taylor & Francis.

Spector, P. E., Zapf, D., Chen, P. Y., & Frese, M. (2000). Why negative affectivity should not be controlled in job stress research: Don't throw out the baby with the bath water. *Journal of Organizational Behavior*, *21*(1), 79–95. https://doi.org/10.1002/(SICI)1099-1379(200002)21:1<79::AID-JOB964>3.0.CO;2-G

Swift, V., & Peterson, J. B. (2019). Contextualization as a means to improve the predictive validity of personality models. *Personality and Individual Differences*, *144*, 153–163. https://doi.org/10.1016/j.paid.2019.03.007

Tay, L., & Jebb, A. T. (2018). Establishing construct continua in construct validation: The process of continuum specification. *Advances in Methods and Practices in Psychological Science*, *1*(3), 375–388. https://doi.org/10.1177/2515245918775707

Taylor, S. E., & Lobel, M. (1989). Social comparison activity under threat: Downward evaluation and upward contacts. *Psychological Review*, *96*(4), 569–575. https://doi.org/10.1037/0033-295X.96.4.569

Taylor, S. E., Wayment, H. A., & Carrillo, M. (1996). Social comparison, self-regulation, and motivation. In R. M. S. E. T. Higgins (Ed.), *Handbook of motivation and cognition, Vol. 3: The interpersonal context* (Vol. 3, pp. 3–27). Guilford Press.

Thissen, D., Steinberg, L., & Wainer, H. (1993). Detection of differential item functioning using the parameters of item response models. In *Differential item functioning* (pp. 67–113). Erlbaum. https://doi.org/10.1075/z.62.13kok

Tourangeau, R. (1987). Attitude measurement: A cognitive perspective. In H.-J. Hippler, N. Schwarz, & S. Sudman (Eds.), *Social information processing and survey methodology* (pp. 149–162). Springer.

Van der Doef, M., & Maes, S. (1999). The Job Demand-Control (-Support) Model and psychological well-being: A review of 20 years of empirical research. *Work & Stress*, *13*(2), 87–114. https://doi.org/10.1080/026783799296084

Vanhove, J. (2015, April 14). *Power simulations for comparing independent correlations*. https://janhove.github.io/design/2015/04/14/power-simulations-for-comparing-independent-correlations

Voss, N. M., Chlevin-Thiele, C., Lake, C. J., & Warren, C.-L. (2023). Using item response theory to understand the effects of scale contextualization: An illustration using decision making style scales. *International Journal of Testing*, *23*(1), 34–55. https://doi.org/10.1080/15305058.2022.2047692

Watson, S. B., Goh, Y. W., & Sawang, S. (2011). Gender influences on the work-related stress-coping process. *Journal of Individual Differences*, *32*(1), 39–46. https://doi.org/10.1027/1614-0001/a000033

Whaley, C. J., Morrison, D. L., Payne, R. L., Fritschi, L., & Wall, T. D. (2005). Chronicity of psychological strain in occupational settings and the accuracy of the General Health Questionnaire. *Journal of Occupational Health Psychology*, *10*(4), 310–319. https://doi.org/10.1037/1076-8998.10.4.310

Wood, J. V. (1989). Theory and research concerning social comparisons of personal attributes. *Psychological Bulletin*, *106*(2), 231–248. https://doi.org/10.1037/0033-2909.106.2.231

## 3.4 Study 4

# When Employees Understate Their Stress: Defensive Biasing in Work Stress Surveys

Berit Greulich, Cornelius J. König, and Ramona Mohr

Universität des Saarlandes, Germany

### Author Note

Berit Greulich https://orcid.org/0000-0002-5955-7423, Cornelius J. König https://orcid.org/0000-0003-0477-8293, and Ramona Mohr, Industrial and Organizational Psychology, Saarland University, Saarland, Germany.

We have no conflicts of interest to disclose.

We gratefully acknowledge funding from the German Research Foundation (to Cornelius J. König), Grant KO4244/4-1.

Correspondence concerning this article should be addressed to Berit Greulich, Saarland University, Campus A1 3, 66123 Saarbrücken, Germany. E-mail: berit.greulich@gmail.com

**Abstract**

Purpose: The purpose of this study is to investigate the phenomenon of defensive biasing in work stress surveys, which occurs when employees trivialize potential stressors and strains due to fear of negative consequences from their supervisors or management. This study aims to better understand the factors that influence this behavior and to develop a scale to measure it.

Design/methodology/approach: The study used an online survey of 200 employees to investigate the factors influencing defensive biasing behavior. The researchers developed a scale for defensive biasing with the help of subject matter experts and derived possible factors from the literature. Participants were presented with a scenario in which they imagined a work stress survey in their organization and were asked to answer related items. The data were analyzed by using regression analysis.

Findings: The study found that defensive biasing behavior was significantly predicted by perceived anonymity and neuroticism. Participants who felt more anonymous and had higher levels of neuroticism were more likely to engage in defensive biasing. Job insecurity and trust in supervisor were not found to be significant predictors of defensive biasing.

Originality: This study contributes to the literature on work stress surveys by developing a scale for defensive biasing and investigating the factors that influence this behavior. The study highlights the importance of making the survey process more transparent to reduce defensive biasing and obtain trustworthy results.

*Keywords*: work-related stress, employee survey, defensive biasing, perceived anonymity, neuroticism

## When Employees Understate Their Stress: Defensive Biasing in Work Stress Surveys

Response bias, a tendency where individuals respond in particular ways that may not reflect their genuine attitudes, beliefs, or behaviors, can significantly affect the validity of survey results, potentially leading to inaccurate conclusions (Navarro-González et al., 2016; Paulhus, 1991; Rammstedt et al., 2010). This bias is particularly critical in work stress assessment, typically measured by self-report. According to a study on response behavior in work stress surveys (Greulich et al., 2021), some employees may downplay their stressful work conditions to avoid potential negative repercussions from superiors. This self-protection strategy, termed 'defensive biasing,' may lead to the under-reporting of work stress. To gain a deeper understanding of defensive biasing, which has so far only been qualitatively examined in one study (Greulich et al., 2021), we have developed a scale for quantitative examination. Moreover, we propose theoretical arguments regarding factors influencing defensive biasing, incorporating the social identity model of deindividuation effects (Reicher et al., 1995). This study thereby aims to enrich the understanding of motivational processes underlying response behavior in work stress surveys.

## Theoretical Background

Changes in working life regarding the use of complex technology, globalization, and high competitive pressure lead to a high workload for many employees, and these stressful working conditions result in psychological strain at the workplace that affects both psychological wellbeing and physical health (Niedhammer et al., 2021). Outcomes of a high workload include job dissatisfaction, turnover intentions, and counterproductive work behavior (e.g., Fida et al., 2015). In

the long run, psychological strain at work might result in physical and mental illnesses such as migraine, cardiovascular diseases, sleeping disorders, depression, and burnout (e.g., Taouk et al., 2020).

Sickness caused by work stressors is associated with considerable suffering for the person concerned, but also with economic problems for the organization due to sick leave and costs for society (Han et al., 2019). To mitigate these negative consequences, organizations need to know the (potential) sources of work stress. Suitable interventions could then be taken to minimize sources of stress and thus maintain the well-being and health of the employee. This makes it important for organizations to assess working conditions that result in health problems for their employees.

**Measuring Work Stress**

Work stress encompasses three pivotal elements: stressors, resources, and strain (e.g., Bakker et al., 2023). Stressors are stimuli requiring sustained physical, cognitive, or emotional effort in the work environment, such as high demands on concentration or role ambiguity. Resources, conceptualized as an energetic reservoir an individual draws upon to cope with stressors (Lazarus & Folkman, 1987), include support from colleagues or action latitude. Hence, stress surveys typically assess these elements and their repercussions.

The primary method for evaluating work stressors, resources, and strain is self-report surveys, where participants rate the intensity of specific work stressors, resources, and strain (e.g., Kristensen et al., 2005; Schulte et al., 2021; Semmer et al., 1999).This method is efficient, allowing for substantial data collection with minimal time and resource investment. Furthermore, employees, being intimately familiar with their job demands, serve as expert informants. Moreover, subjective perceptions of

work situations are often more suited for strain reactions compared to objective indicators (Perrewé & Zellars, 1999; Spector & Jex, 1998).

Nevertheless, self-report surveys suffer from some disadvantages because they might be biased due to factors such as negative affectivity, acquiescence, and situational contexts. Evidence for these biases can be found in various research fields that use self-reports (e.g., Debus et al., 2015; Morgeson & Campion, 1997; Schmit et al., 1995). Results of a qualitative study by Greulich et al. (2021) about the response behavior in work stress surveys show a new form of biasing. Participants reported that they deliberately understate or trivialize their ratings of stressors or strain when they fear negative consequences from their supervisors to protect themselves against material and immaterial loss (e.g., ostracism or job loss; Greulich et al., 2021).

Numerous studies have demonstrated that fear-based motives can result in employee silence, where individuals intentionally withhold work-related information or criticism (e.g., Pinder & Harlos, 2001; Van Dyne et al., 2003). The literature has established the concept of "silence" (e.g., Hao et al., 2022; Morrison & Milliken, 2000), defined as the withholding of work-related information due to underlying motives (Brinsfield et al., 2009). Brinsfield et al. (2009) investigated the specific fears that may cause employee silence and found that fear of negative consequences was a common motivation. This form of silence, referred to as "defensive silence," is for self-protection and may arise from concerns of job loss, appearing incompetent, being labeled a complainer, or causing conflict in the workplace (Jahanzeb et al., 2018; Milliken et al., 2003). De Clercq et al. (2020) demonstrated that belief in verbally abusive leaders can prompt employees to engage in defensive silence, which can ultimately reduce the risk of negative performance evaluations. This highlights the

significance of defensive silence as a mechanism used by employees to avoid negative performance evaluations by abusive leaders.

Although the fear of potential negative consequences has been found to make employees remain silent (Kish-Gephart et al., 2009; Morrison & Milliken, 2000) and could be considered a special case of impression management, this fear has largely been unexplored in the context of work stress surveys. When people are forced to make a statement by participating in a questionnaire, they have to resort to another kind of self-protection: trivializing stressors and strains and presenting resources as more pronounced. The concealment of stressful working conditions to protect oneself from negative consequences can thus bias stress surveys. Based on the construct of "defensive silence" (see e.g., Jahanzeb et al., 2018; Milliken et al., 2003; Van Dyne et al., 2003), this intentional and proactive response behavior in work stress surveys can be called "defensive biasing". This defensive biasing likely distorts the quality of the obtained data and leads organizations to make ineffective management and human resources decisions. To improve the validity of work stress surveys, it is crucial to better understand the phenomenon of defensive biasing.

The qualitative results of the study by Greulich et al. (2021) and the literature on survey response behavior and defensive silence offer arguments which factors likely influence defensive biasing. Factors that lead to voice and silence are often divided into motivators and inhibitors, respectively (see Morrison, 2014). These factors can in turn be divided into situational circumstances, such as characteristics of the relationship with the supervisor, and also into individual dispositions, such as certain personality traits. Furthermore, beliefs about personal consequences and sensitivity to anonymity seem to have a significant influence on the response behavior (Mueller et al., 2014).

**Anonymity as a Predictor for Defensive Biasing**

The employees interviewed for the qualitative study of Greulich et al. (2021) emphasized that the anonymity of the survey plays a major role in defensive biasing. They mentioned a fear of being recognized despite assured anonymity, for example, based on demographic data. They kept this possibility in mind when answering questions for employee work stress surveys. This suggests that the perceived anonymity of the survey influences the fear-related response behavior of survey participants, which is also consistent with the literature on organizational surveys in general that also assigns great importance to the perceptions of anonymity (Rogelberg et al., 2006).

There are two forms of anonymity: literal anonymity and perceived anonymity (Dunnette & Heneman, 1956). It is quite possible that a survey objectively preserves anonymity (i.e., has literal anonymity), but respondents still do not perceive it as such. Perceived anonymity has already been established as a relevant factor for self-reports (e.g., Saari & Scherbaum, 2011). A meta-analysis summarized that different objective implementation strategies of anonymity yielded only small effects on the response behavior, which speaks for a rather subjective impression of anonymity (Singer et al., 1995).

The assumption of the influence of perceived anonymity on specific fear-based response behavior is plausible against the background of the so-called social identity model of deindividuation effects (Reicher et al., 1995), which originally stems from the research about computer-mediated communication, a field of social psychology. Social identity theory (Tajfel & Turner, 1986) and self-categorization theory (Turner et al., 1987) form the basis of this model, postulating that a low level of identifiability leads individuals to orient themselves to their values and norms since

they do not have to fear sanctions. In contrast, as identifiability increases, individuals become more oriented toward the norms and values of the outside group. The social identity model of deindividuation can also be applied to anonymous employee surveys (Chudziak & Maus, 2008) and may better explain the relationship between subjectively low perceived anonymity and defensive biasing in an organizational context than the classic social desirability models. In the context of work stress surveys, high perceived identifiability toward the outside group (i.e., the supervisors or management) leads to increased adoption of the norms and values of this group. To avoid sanctions by more powerful supervisors or management, employees refrain from response behavior that can lead to punishment. Therefore, answers may be intentionally biased due to fear as doubts about anonymity grow. These theoretical considerations and findings lead to the assumption that the subjectively low perceived anonymity of the survey reinforces the defensive biasing of answers. In the context of work stress surveys, being sure that employees are not identifiable seems crucial for honest evaluations (so employees are not engaging in defensive biasing). In the case of stressful working conditions, employees may perceive that a true rating could cause negative actions from their supervisors (e.g., ostracism). By trivializing possible stressors or strains and presenting resources more positively, employees can try to protect themselves from negative consequences. This leads to the following hypothesis:

*Hypothesis 1: Perceived anonymity is negatively related to defensive biasing.*

**Neuroticism**

In the realm of silence research, personality traits are frequently examined as potential motivators or inhibitors of employee voice and silence (e.g., Hao et al., 2022). Among the Big Five personality traits (Costa & McCrae, 1992), neuroticism is

one of the most widely studied and frequently examined factors in personality psychology (Tackett & Lahey, 2017). Furthermore, this trait is also commonly examined in the context of stress research (e.g., Roloff et al., 2022). Neurotic individuals tend to experience negative emotions such as anxiety or anger more frequently and are more sensitive to stressors than their emotionally stable counterparts (Costa & McCrae, 1992). They are also more prone to focusing on negative stimuli in the environment, which contributes to their perception of situations as more threatening (Nettle, 2006). Consequently, their elevated fear and insecurity can lead to reduced expression of opinions, concerns, and problems in the workplace (LePine & Van Dyne, 2001). This finding is supported by Li and Xu (2020), who demonstrated that employees with higher levels of neuroticism exhibit lower levels of expressive behavior at work, which is mediated by emotional exhaustion. Brinsfield's (2013) research also indicated that high levels of employee neuroticism are associated with increased defensive silence. Additionally, De Clercq et al. (2020) suggest that the effectiveness of defensive silence may be particularly pronounced among individuals with high levels of neuroticism. This argument also extends to defensive biasing in work stress surveys, as those who are more neurotic may fear negative consequences more than their emotionally stable counterparts when asked to report work stressors and strain. Therefore, we expect that there is a relationship between neuroticism and defensive biasing:

*Hypothesis 2: Neuroticism is positively related to defensive biasing.*

**Job Insecurity**

Job insecurity reflects employees' worry about losing their present job and has strong psychological impact on employees (Lee et al., 2018). Furthermore, job insecurity decreases employee voice and increases employee silence (Breevaart et al.,

2020). There is also evidence that job insecurity also affects respondents' answering behavior. In interviews, employees expressed job loss as a feared negative consequence of answering work stress surveys truthfully (Greulich et al., 2021). These interviewees also mentioned that staff reductions led to a tense atmosphere and fear in the company which was reflected in a previous employee survey in the form of trivialization of answers (Greulich et al., 2021). One way of coping with job insecurity is to ensure that others see the value they bring to their organization by engaging in impression management. Empirical support for this argument comes from a study by Huang et al. (2013), who found job insecurity and impression management to be positively related. Furthermore, job-insecure people might also increase their work efforts and their performance to protect their jobs (Staufenbiel & König, 2010). Defensive biasing in surveys could be seen as a further protection strategy against potential job loss because it helps make the impression that the work situation is still good and that the employee is still able to maintain high performance and not be impaired by stress or emotional exhaustion. In the context of work stress surveys, job loss fear could increase the motivation for defensive biasing. For example, if the company's economic situation is poor and the probability of losing one's job is thus greater, it is conceivable that employees in work stress surveys are more likely to present stressors and strains in a mitigated form because they want to avoid behavior that could jeopardize their job. The mentioned empirical findings and theoretical considerations lead to the arguments that employees will increasingly engage in defensive biasing when having doubts regarding the stability of their jobs. More formally, we can state the following hypothesis:

*Hypothesis 3: Job insecurity is positively related to defensive biasing.*

**Trust in Supervisor**

A fourth main predictor of defensive biasing should be trust in the supervisor. Trust can be defined as "the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party" (Mayer et al., 1995, p. 712). The amount of trust employees have in their supervisors appears to be particularly important for how they respond to fear (Detert & Burris, 2007). The influence of trust on employees' voice and silence behavior has already been supported by empirical findings: A significant negative correlation was found between trust and employees' fear-based silence (Brinsfield et al., 2009), and the less trust there was in the supervisor, the more likely employees were to remain silent for fear of negative consequences (Dedahanov et al., 2015). Thus, a good relationship of trust between supervisor and employee can lead to a reduction in the fear of negative consequences when expressing opinions or criticism because the supervisor takes the employees' problems seriously and is prepared to listen to their suggestions. The goodwill expressed by the supervisor toward their employees and the resulting trust can have an encouraging effect not only on voicing criticism, for example of work processes, but also on communicating personal burdens, as is also the case in a work stress survey. Therefore, if employees trust that their supervisors will take appropriate actions when getting negative feedback, this might also generalize to employees' strategies while filling out work stress surveys. We can thus propose the following hypothesis:

*Hypothesis 4: Trust in supervisor is negatively related to defensive basing.*

## Method

### Sample

The sample comprised 200 German employees (61% female) across various professions (30.5% service providers, 25.5% industry, 16.5% health sector, 7.5% education sector, 5.5% social sector, 14.5% other). Participants' ages were categorized: 13.5% under 24, 31.5% between 25-34, 13.5% between 35-44, 22.5% between 45-54, and 19% over 55. All were non-self-employed, with 23.5% holding management positions and 82.5% having permanent contracts. The average professional experience was 15.8 years (*SD* = 12.8), with 10.7 years in their current job (*SD* = 10.5). About 11.5% had prior experience with work stress surveys. The snowball principle was employed for recruitment, with voluntary participation and no remuneration. Anonymity was assured, and participants were informed that the study served scientific purposes, not requiring their employer's name.

### Questionnaire

#### *Scenario*

The questionnaire, presented on an online platform, began with a scenario where participants imagined a work stress survey being distributed in their organization. They were provided with excerpts from the German version of the Copenhagen Psychosocial Questionnaire (COPSOQ, see below, Kristensen et al., 2005), a reliable and valid instrument for assessing work stressors in German-speaking countries (Nübling et al., 2005). Participants first filled in demographic data before answering selected items from the COPSOQ about stressors, resources, and strain. The COPSOQ items were chosen in consultation with one of its authors (M. Nübling, personal communication, 07.05.2018), focusing on those most likely to cause fear of consequences if answered honestly with low perceived anonymity. The

COPSOQ comprises five domains, and items from different scales were selected for the online questionnaire. Table 1 displays the ten selected items. After completing the work stress survey section, respondents qualitatively described their response behavior and reported their experience with work stress questionnaires.

**Table 1**

*The Items Used in This Study (From the Copenhagen Psychosocial Questionnaire, Kristensen et al., 2005) and the Scales They Belong To*

| Item | Scale |
|---|---|
| Do you feel that the work you do is important? | Meaning of work |
| Are you proud of being part of this company? | Commitment to workplace |
| To what extent would you say that your immediate superior gives high priority to job satisfaction? | Quality of leadership |
| To what extent would you say that your immediate superior is good at work planning? | Quality of leadership |
| How often is your immediate superior willing to listen to your problems at work, if needed? | Support at work |
| How often do you feel unjustly criticised, bullied or shown up in front of others by your colleagues and your superior? | Unfair treatment |
| Can the employees trust the information that comes from the management? | Trust and justice |
| Is your work recognized and appreciated by the management? | Recognition |
| In the past 12 months, how often have you thought about changing your job? | Intention to leave Profession/job |
| How often do you feel emotionally exhausted? | Burnout symptoms |

### *Defensive Biasing*

No existing instrument was found in the literature search that specifically measures defensive biasing, necessitating the development of a new instrument. To capture defensive biasing, a direct measure was employed, wherein participants were asked whether they would respond to work stress surveys in a deliberately understated manner given the described situation. This approach utilized behavioral intention as a proxy for actual or future behavior. To enable this inference, it is crucial to provide a detailed description of the survey context, including its implementation (Ajzen & Fishbein, 1980).

Drawing upon qualitative interviews conducted by Greulich et al. (2021), interviews with eight subject matter experts, and relevant empirical studies on voice and silence research, four items were developed to describe potential reasons for defensive biasing (see Table 2). Items 1 and 4 address the fear of material losses. By expressing job-related grievances in a mitigated manner and providing more positive responses regarding their supervisor, individuals may seek to avoid negative consequences such as job loss, conflicts with their supervisor, and a deterioration of the work climate. On the other hand, items 2 and 3 pertain to the fear of intangible losses if participants were to answer the questions truthfully. To avoid appearing weak or incompetent, individuals may downplay personal feelings, worries, and stressors. Thus, the scale encompasses behaviors aimed at mitigating both material and immaterial losses. Participants rated the items on a five-point scale ranging from 1 (strongly disagree) to 5 (strongly agree). The wording of the items was designed such that agreement indicates a stronger tendency towards defensive biasing.

**Table 2**

*Defensive Biasing Scale*

| | Item | Factor loadings |
|---|---|---|
| | If a risk assessment of psychological stress was conducted in my company in the manner and with the given questions mentioned above, I would... | |
| 1 | …present work-related grievances in an attenuated way. | .89 |
| 2 | …slightly downplay circumstances that burden me. | .89 |
| 3 | …trivialize statements regarding my personal condition or worries. | .75 |
| 4 | …answer questions regarding my superiors or colleagues a bit more positive than reality might reflect them. | .64 |

*Note*. The original German items are available from the authors.

***Perceived Anonymity***

To assess perceived anonymity, the Perceived Anonymity Scale (PANON, Whelan & Thompson, 2009) was translated into German by a professional translator and then back-translated into English by another translator. During the translation process, both translators deemed the item "My responses will blend in with the responses of other people" as untranslatable into German, leading to its exclusion. Consequently, the scale consisted of five remaining items. These items included statements such as "I feel my responses are indistinguishable from the responses of others that have taken this survey" and "It would be impossible to trace my responses

to this survey back to me." Participants were instructed to consider the preceding section of the work stress survey (i.e., the scenario part) when responding to these items. A five-point scale ranging from 1 (strongly disagree) to 5 (strongly agree) was used for participants to rate their agreement, with higher scores indicating greater perceived anonymity.

### *Neuroticism*

Neuroticism was measured using the 8-item scale from the German version of the Big Five Inventory (BFI, John et al., 1991; Rammstedt & John, 2005). A sample item is “I see myself as someone who worries a lot.” The items were answered on a five-point scale from 1 (strongly disagree) to 5 (strongly agree).

### *Job Insecurity*

Job insecurity was measured using the 4-item scale of Staufenbiel and König (2010), coded so that high values indicated high levels of job insecurity. A sample item is "In my opinion, I will be employed long-term at my current job."

### *Trust in Supervisor*

Trust in supervisor was assessed with the 9-item scale from the German version of the Workplace Trust Survey (Ferres & Travaglione, 2003; Lehmann-Willenbrock & Kauffeld, 2010), with agreement indicating high trust. A sample item is "My supervisor treats personal conversations confidentially."

## Results

### Preliminary Analyses

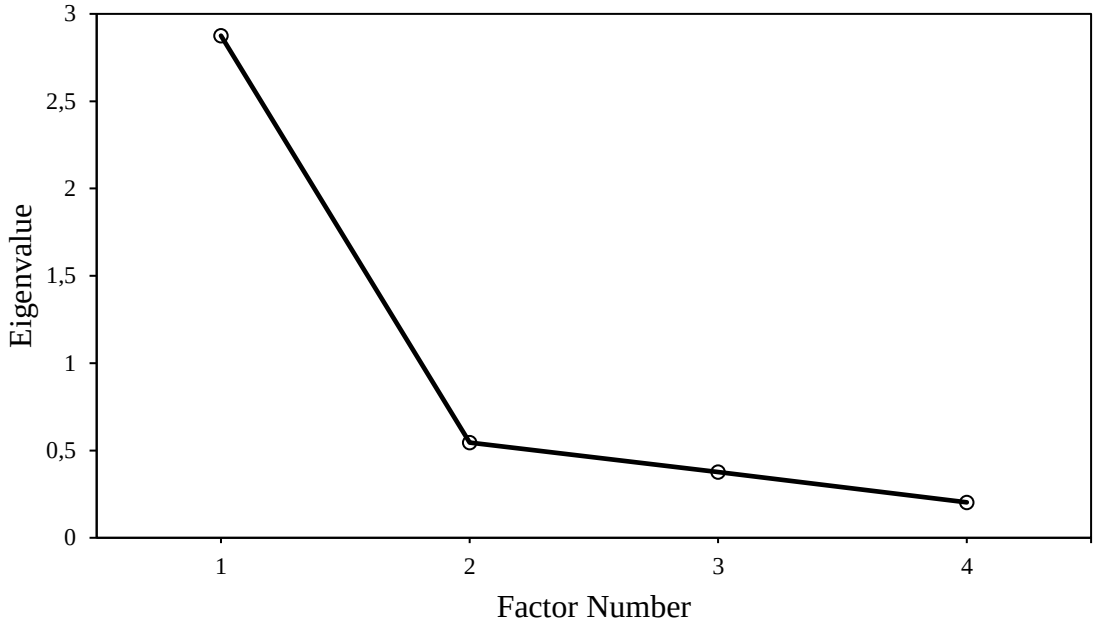
### *Factor Analysis*

A principal axis factor analysis was performed on the four defensive bias items, with oblique rotation. The Kaiser-Meyer-Olkin (KMO) measure confirmed sampling adequacy, KMO = .8 (Hutcheson & Sofroniou, 1999), and all individual

item KMO values exceeded .74, surpassing the acceptable limit of .5 (Field, 2013). Initial analysis for the defensive bias scale produced eigenvalues for each factor. One factor had an eigenvalue above Kaiser's criterion of 1, explaining 71.88% of the variance. The scree plot (Figure 1) displayed inflections supporting the retention of one factor. Table 3 presents factor loadings, with items clustering on one factor, indicating defensive biasing. The defensive biasing scale demonstrated high reliability, with Cronbach's Alpha at .84 and McDonald's Omega at .87.

**Figure 3**

*Screeplot of the Exploratory Factor Analysis*




***Experience with Work Stress Questionnaires***

A minority of participants (11.5%) had prior experience with work stress surveys. Among them, 26% reported positive experiences and changes in their company, while 30% reported negative experiences and no changes, particularly at the management level.

### *Qualitative Data*

After the scenario, only 57.5% of respondents provided qualitative descriptions of their response behavior. Among them, 20.5% could realistically imagine themselves in the situation and 12% did not trust the survey's anonymity and feared negative consequences, with one participant stating s/he answered "cautiously" due to mistrust in the word "anonymously." Seven percent followed their feelings when answering, while four participants mentioned feeling stressed and worrying about it. A fifth (20.5%) said they answered the questions truthfully, with one person writing that they felt the questionnaire was anonymous. Another person added that although they felt inferences could be made about the individual, this had not mattered to them because of their good relationship with their supervisor.

## Test of Hypotheses

A multiple linear regression analysis was conducted to examine the prediction of defensive biasing from perceived anonymity, neuroticism, job insecurity, and trust in supervisor (see Table 3 for mean values, standard deviations, and correlations). The results (Table 4) showed a significant association between perceived anonymity and neuroticism. Perceived anonymity negatively predicted defensive biasing ($\beta = -.18$, $p < .01$), confirming Hypothesis 1. Neuroticism positively predicted defensive biasing ($\beta = .17$, $p < .05$), supporting Hypothesis 2. However, job insecurity ($\beta = .03$, $p = .73$) and trust in supervisor ($\beta = .05$, $p = .49$) were not significant predictors, failing to confirm Hypotheses 3

**Table 3**

*Means, Standard Deviations, Cronbach's Alpha, and Correlations*

| Variables | *M* | *SD* | Cronbach's α | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 Defensive biasing | 2.42 | 0.92 | .87 | | | | | | | | | | | |
| 2 Perceived anonymity | 3.35 | 0.80 | .79 | -.19** | | | | | | | | | | |
| 3 Neuroticism | 2.64 | 0.69 | .90 | .18** | -.09 | | | | | | | | | |
| 4 Job insecurity | 2.19 | 1.10 | .88 | .08 | -.03 | .36** | | | | | | | | |
| 5 Trust in supervisor | 3.97 | 0.80 | .94 | -.02 | .10 | -.28** | -.24** | | | | | | | |
| 6 Gender | 1.61 | 0.49 | - | .02 | -.07 | .23** | .01 | -.14 | | | | | | |
| 7 Age | 3.02 | 1.36 | - | -.18* | .01 | -.12 | -.15* | -.14 | .01 | | | | | |
| 8 Work experience | 15.80 | 12.83 | - | -.20** | .04 | -.03 | -.14* | -.18** | -.07 | .82** | | | | |
| 9 Employment in current job | 10.70 | 10.52 | - | -.19** | -.04 | .01 | -.17* | -.15* | -.07 | .66** | .76** | | | |
| 10 Employment contract | .18 | .38 | - | .10 | .05 | .04 | .35** | .15* | .04 | -.32** | -.39** | -.38** | | |
| 11 Management responsibility | .24 | .43 | - | -.18* | .05 | -.08 | -.06 | .02 | -.23** | .15* | .21** | .23** | -.13 | |

*Notes*. $N = 200$; gender: 1 = male, 2 = female; age 1 ≥ 24 years, 2 = 25-24 years, 3 = 35-44 years, 4 = 45-54 years, 5 ≥ 55; employment contract: 1 = permanent, 0 = temporary; management responsibility: 1= yes, 0 = no; work experience and employment at current job in years.

**Table 4**

*Regressing Defensive Biasing on Perceived Anonymity, Neuroticism, Job Insecurity and Trust in Supervisor*

| Variables | *B* | *SE* | *t* | *p* | 95% CI |
|---|---|---|---|---|---|
| Constant | 2.25 | 0.55 | 4.10 | .000 | [1.17, 3.30] |
| Perceived anonymity | -0.21 | 0.08 | -2.63 | .009 | [-0.37, -0.05] |
| Neuroticism | 0.23 | 0.10 | 2.25 | .026 | [0.03, 0.43] |
| Job insecurity | 0.02 | 0.06 | 0.35 | .729 | [-0.10, 0.15] |
| Trust in supervisor | 0.06 | 0.08 | 0.70 | .490 | [-0.11, 0.22] |

*Note*. CI = Confidence interval. $N = 200$.

## Discussion

This study aimed to quantitatively investigate a new form of biasing specifically in work stress self-reports: defensive biasing. This phenomenon occurs when respondents bias their answers in employee surveys for self-protection. To investigate this phenomenon, we first developed a scale to measure the construct of defensive biasing. Second, we made a first attempt to link other concepts to defensive biasing. We empirically demonstrated that perceived anonymity and neuroticism significantly predicted defensive biasing, whereas job insecurity and trust in superiors could not be confirmed as predictors.

Initially, the process of developing a psychometric instrument for measuring defensive biasing resulted in a four-item scale that can be used to assess deliberate understatement or trivialization of employee ratings of stressors or strain. The use of this scale will help to understand employees' motives to bias their responses in work stress surveys. A factor analysis confirmed that the new scale works well as a

univariate measure of defensive biasing. Furthermore, the scale demonstrated good reliability.

The results of this study show that low perceived anonymity leads to an increased response behavior of defensive biasing. This result also confirms the arguments of Mueller et al. (2014) that the perception of anonymity plays a decisive role in answering items in a survey. Furthermore, the social identity model of deindividuation (Reicher et al., 1995) provides a plausible explanation for this connection: In cases of low perceived anonymity of the work stress survey (i.e., high identifiability of the employees by supervisors or management), employees' response behavior is adjusted to the norms and values of supervisors and management for fear of sanctions. In the context of a work stress survey, the behavior that is avoided is truthfully rating of bad working conditions, because this would mean negative feedback for the foreign group of supervisors or management. The qualitative comments of the employees collected in this study also support the relevance of perceived anonymity. As an answer to the question about their response behavior, participants already noted concerns about assured anonymity (e.g., "I don't trust the phrase anonymously very much, so I responded carefully."). Furthermore, it was stated that "[…] anonymity cannot be guaranteed at all in smaller organizations."

The significant main effect of neuroticism on defensive biasing indicates the important role of personality traits in response behavior. Neurotic individuals appear to do more defensive biasing in surveys. It could be that these employees are in principle less likely to express opinions, concerns, or criticism, regardless of whether an inference can be personally drawn about them or not. A conceivable explanation for this lies in the characteristics of neurotic individuals: A high level of neuroticism is often associated with insecurity and less self-confidence (e.g., Abdellaoui et al.,

2019), which may make these individuals more cautious with their statements, triggering them to formulate statements in a weakened manner. Moreover, they tend to perceive situations more threateningly than they actually are (Schneider, 2004). Possible negative consequences could be interpreted by them as much more likely. To protect themselves from this, they are left with only defensive biasing.

Although the interviewees in the qualitative study of Greulich et al. (2021) had explicitly spoken of the relevance of job insecurity to the biasing response behavior, this relationship could not be shown quantitatively in this study. In fact, the type of employment contract also had no significant influence. This may be due to the strong protection for employees in Germany (where the data was collected) due to the German employment protection legislation, which makes firing employees rather complicated. This also fits with the rather low mean value of job insecurity (i.e., 2.19). Furthermore, unemployment rates in Germany are not very high (Federal Employment Agency, 2021). In countries with different legal regulations and higher unemployment rates, it might be possible to find a moderating effect of job insecurity.

The assumption that trust in supervisor influences the likelihood of defensive biasing could not be confirmed by our data. We had argued that a trustful relationship between employees and supervisors should increase the likelihood of employees admitting to being overtaxed and exhausted, but this does not seem to be the case. A possible explanation is that employees often do not know who exactly is meant by the term supervisor in surveys. Especially in organizations with many hierarchical levels, several persons could be considered (cf., Hackman & Wageman, 2004; Therkelsen & Fiebich, 2004). On the one hand, the participants could think of their team leader, who would allocate tasks to them. On the other hand, they could also think of a higher level of superiors who are familiar with personnel decisions. Depending on who they

refer to (team leader, management, CEO) they have more or less confidence that could affect defensive biasing. A methodological justification could be the conspicuously right-skewed distribution of the variable trust in the superior.

Scores of defensive biasing were also significantly correlated with age and work experience, employment in the current job, and management responsibility. These results suggest that defensive biasing might be a phenomenon that tends to affect younger employees with less work experience and less management responsibility.

**Limitations**

As with all studies, this study has its limitations. First, it relies on participants' capacity to imagine a hypothetical scenario of a work stress survey in their organization, which may impact its effectiveness. However, qualitative feedback suggested that a substantial proportion of participants, specifically 40.9%, were able to effectively simulate this scenario, with comments such as "I treated the question as if my employer were asking it." This lends some validity to the employed method. Second, all data were gathered exclusively through self-report in a single questionnaire, which could introduce potential common method variance, thus limiting the interpretability of the results. Nonetheless, the nature of the study made self-reporting indispensable as it was the only way for participants to express their perceptions of anonymity and their response behavior. Moreover, it should be noted that the common method effect is counterbalanced by the unreliability of the measuring instruments (Lance et al., 2010), indicating it should not strongly affect results. Nonetheless, future research could consider supplementing self-reports with data collected from other sources.

**Future Research and Implications**

This study highlights the use of defensive biasing by employees in work stress surveys as a means of self-protection. Future research should investigate the extent to which defensive biasing skews relationships between constructs assessed in organizational surveys. Additionally, further exploration is needed to understand the subjective perception of anonymity and identify potential influencing factors. The newly developed scale for measuring defensive biasing offers significant advantages for future research in this area. Its practical length and high reliability make it a valuable tool. Unlike indirect deductions of defensive biasing from questionnaire responses, which may involve interpretation and yield less reliable findings, this scale directly measures behavior. Thus, future research should focus on validating the scale and employing it in further investigations of defensive biasing.

This study also provides practical implications for conducting employee surveys. Transparent communication regarding the purpose and process of surveys can alleviate employees' concerns and fears (see also Schweiger & DeNisi, 1991). Explaining to employees why specific questions are being asked can reduce the fear of negative consequences associated with describing work stress. Additionally, when surveys are conducted by external companies, employees may perceive them as fair and unbiased, benefiting both supervisors and management. Another practical implication emerges from qualitative responses regarding past experiences with work stress surveys, where participants often criticized the lack of resulting changes. This underscores the importance of meeting employees' expectations and implementing changes based on survey findings. Failure to satisfy these expectations may negatively impact future surveys of a similar nature. Therefore, special attention should be given to implementing meaningful changes following survey administration.

## Conclusion

The purpose of the current study was to take the first step towards quantitatively investigating the occurrence of defensive biasing and its cause-effect relationships. Perceived anonymity and high neuroticism were associated with an increased occurrence of defensive biasing. Hopefully, this study stimulates more research on this important phenomenon: employees who protect themselves by not honestly describing how much work stress they experience.

**References**

Abdellaoui, A., Chen, H.-Y., Willemsen, G., Ehli, E. A., Davies, G. E., Verweij, K. J. H., Nivard, M. G., de Geus, E. J. C., Boomsma, D. I., & Cacioppo, J. T. (2019). Associations between loneliness and personality are mostly driven by a genetic association with neuroticism. *Journal of Personality*, *87*(2), 386–397. https://doi.org/10.1111/jopy.12397

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Prentice-Hall.

Bakker, A. B., Demerouti, E., & Sanz-Vergel, A. (2023). Job demands-resources theory: Ten years later. *Annual Review of Organizational Psychology and Organizational Behavior*, *10*(1), 25–53. https://doi.org/10.1146/annurev-orgpsych-120920-053933

Breevaart, K., Lopez Bohle, S., Pletzer, J. L., & Muñoz Medina, F. (2020). Voice and silence as immediate consequences of job insecurity. *Career Development International*, *25*(2), 204–220. https://doi.org/10.1108/CDI-09-2018-0226

Brinsfield, C. T. (2013). Employee silence motives: Investigation of dimensionality and development of measures. *Journal of Organizational Behavior*, *34*(5), 671–697. https://doi.org/10.1002/job.1829

Brinsfield, C. T., Edwards, M. S., & Greenberg, J. (2009). Voice and silence in organizations: Historical review and current conceptualizations. In J. Greenberg & M. S. Edwards (Eds.), *Voice and silence in organizations* (pp. 1–33). Emerald.

Bundesagentur für Arbeit. (2021, October 3). *Arbeitslosenquote & Arbeitslosenzahlen 2021 [Unemployment rate & unemployment figures 2021]*. https://www.arbeitsagentur.de/news/arbeitsmarkt-2021#:~:text=Die%20Arbeitslosenquote%20sank%20um%200%2C1%20Prozentpunkte%20auf%205%2C6,die%20Arbeitslosenzahl%20um%20320.000%20gesunken.

Chudziak, N., & Maus, D. (2008). Einfluss der Anonymitätswahrnehmung auf das Antwortverhalten in Online-Mitarbeiterbefragungen im interkulturellen Vergleich [Influence of anonymity perception on response behavior in online employee surveys in cross-cultural comparison]. *Mannheimer Beiträge zur Wirtschafts- und Organisationspsychologie*, *1*, 3–10.

Costa, P. T., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: The NEO Personality Inventory. *Psychological Assessment*, *4*(1), 5–13. https://doi.org/10.1037/1040-3590.4.1.5

De Clercq, D., Jahanzeb, S., & Fatima, T. (2020). How abusive supervision ultimately might enhance performance ratings among silent, neurotic employees. *Personnel Review*, *50*(5), 1297–1315. https://doi.org/10.1108/PR-01-2020-0007

Debus, M. E., König, C. J., Kleinmann, M., & Werner, C. S. (2015). Examining the effects of negative affectivity on self- and supervisor ratings of job stressors: The role of stressor observability. *Work & Stress*, *29*(4), 341–361. https://doi.org/10.1080/02678373.2015.1075233

Dedahanov, A. T., Kim, C., & Rhee, J. (2015). Centralization and communication opportunities as predictors of acquiescent or prosocial Silence. *Social Behavior and Personality: An International Journal*, *43*(3), 481–492. https://doi.org/10.2224/sbp.2015.43.3.481

Detert, J. R., & Burris, E. R. (2007). Leadership behavior and employee voice: Is the door really open? *The Academy of Management Journal*, *50*(4), 869–884. https://doi.org/10.5465/amj.2007.26279183

Dunnette, M. D., & Heneman, H. G., Jr. (1956). Influence of scale administrator on employee attitude responses. *Journal of Applied Psychology*, *40*(2), 73–77. https://doi.org/10.1037/h0040428

Ferres, N., & Travaglione, A. (2003). *The development and validation of the workplace trust survey (WTS): Combining qualitative and quantitative methodologies*. APROS Emotions, Attitudes, and Culture Stream, Oaxaca, Mexico. https://www.semanticscholar.org/paper/Development-and-Validation-of-the-WTS-1-The-and-of-Ferres/00dd98535dceef1b7c22380b5ed5dc154a358ee2

Fida, R., Paciello, M., Tramontano, C., Fontaine, R. G., Barbaranelli, C., & Farnese, M. L. (2015). An integrative approach to understanding counterproductive work behavior: The roles of stressors, negative emotions, and moral disengagement. *Journal of Business Ethics*, *130*(1), 131–144. https://doi.org/10.1007/s10551-014-2209-5

Field, A. (2013). *Discovering statistics using IBM SPSS Statistics* (4th ed.). Sage.

Greulich, B., Debus, M. E., Kleinmann, M., & König, C. J. (2021). Response behavior in work stress surveys: A qualitative study on motivational and cognitive processes in self- and other-reports. *European Journal of Work and Organizational Psychology*, *30*(1), 40–57. https://doi.org/10.1080/1359432X.2020.1812580

Hackman, J. R., & Wageman, R. (2004). When and how team leaders matter. *Research in Organizational Behavior*, *26*, 37–74. https://doi.org/10.1016/S0191-3085(04)26002-6

Han, S., Shanafelt, T. D., Sinsky, C. A., Awad, K. M., Dyrbye, L. N., Fiscus, L. C., Trockel, M., & Goh, J. (2019). Estimating the attributable cost of physician burnout in the United States. *Annals of Internal Medicine*, *170*(11), 784–790. https://doi.org/10.7326/M18-1422

Hao, L., Zhu, H., He, Y., Duan, J., Zhao, T., & Meng, H. (2022). When is silence golden? A meta-analysis on antecedents and outcomes of employee silence. *Journal of Business and Psychology*, *37*(5), 1039–1063. https://doi.org/10.1007/s10869-021-09788-7

Huang, G., Zhao, H. H., Niu, X., Ashford, S. J., & Lee, C. (2013). Reducing job insecurity and increasing performance ratings: Does impression management matter? *Journal of Applied Psychology*, *98*(5), 852–862. https://doi.org/10.1037/a0033151

Hutcheson, G., & Sofroniou, N. (1999). *The multivariate social scientist*. Sage.

Jahanzeb, S., Fatima, T., & Malik, M. A. R. (2018). Supervisor ostracism and defensive silence: A differential needs approach. *European Journal of Work and Organizational Psychology*, *27*(4), 430–440. https://doi.org/10.1080/1359432X.2018.1465411

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *Big Five Inventory—Versions 4a and 5*. University of California, Berkeley, Institute of Personality and Social Research.

Kish-Gephart, J. J., Detert, J. R., Treviño, L. K., & Edmondson, A. C. (2009). Silenced by fear: The nature, sources, and consequences of fear at work. *Research in Organizational Behavior*, *29*, 163–193. https://doi.org/10.1016/j.riob.2009.07.002

Kristensen, T. S., Borritz, M., Villadsen, E., & Christensen, K. B. (2005). The Copenhagen Burnout Inventory: A new tool for the assessment of burnout. *Work & Stress*, *19*(3), 192–207. https://doi.org/10.1080/02678370500297720

Lance, C. E., Dawson, B., Birkelbach, D., & Hoffman, B. J. (2010). Method effects, measurement error, and substantive conclusions. *Organizational Research Methods*, *13*(3), 435–455. https://doi.org/10.1177/1094428109352528

Lazarus, R. S., & Folkman, S. (1987). Transactional theory and research on emotions and coping. *European Journal of Personality*, *1*(3), 141–169. https://doi.org/10.1002/per.2410010304

Lee, C., Huang, G.-H., & Ashford, S. J. (2018). Job insecurity and the changing workplace: Recent developments and the future trends in job insecurity research. *Annual Review of Organizational Psychology and Organizational Behavior*, *5*(1), 335–359. https://doi.org/10.1146/annurev-orgpsych-032117-104651

Lehmann-Willenbrock, N., & Kauffeld, S. (2010). Development and construct validation of the German workplace trust survey (G-WTS). *European Journal of Psychological Assessment*, *26*(1), 3–10. https://doi.org/10.1027/1015-5759/a000002

LePine, J. A., & Van Dyne, L. (2001). Voice and cooperative behavior as contrasting forms of contextual performance: Evidence of differential relationships with big five personality characteristics and cognitive ability. *Journal of Applied Psychology*, *86*(2), 326–336. https://doi.org/10.1037/0021-9010.86.2.326

Li, J., & Xu, S. (2020). Extraversion, neuroticism, and employee voice: A conservation of resources perspective. *Frontiers in Psychology*, *11*(1281). https://doi.org/doi.org/10.3389/fpsyg.2020.01281

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *The Academy of Management Review*, *20*(3), 709–734. https://doi.org/10.2307/258792

Milliken, F. J., Morrison, E. W., & Hewlin, P. F. (2003). An exploratory study of employee silence: Issues that employees don't communicate upward and why. *Journal of Management Studies*, *40*(6), 1453–1476. https://doi.org/10.1111/1467-6486.00387

Morgeson, F. P., & Campion, M. A. (1997). Social and cognitive sources of potential inaccuracy in job analysis. *Journal of Applied Psychology*, *82*(5), 627–655. https://doi.org/10.1037//0021-9010.82.5.627

Morrison, E. W. (2014). Employee voice and silence. *Annual Review of Organizational Psychology and Organizational Behavior*, *1*(1), 173–197. https://doi.org/10.1146/annurev-orgpsych-031413-091328

Morrison, E. W., & Milliken, F. J. (2000). Organizational silence: A barrier to change and development in a pluralistic world. *Academy of Management Review*, *25*(4), 706–725. https://doi.org/10.5465/amr.2000.3707697

Mueller, K., Straatmann, T., Hattrup, K., & Jochum, M. (2014). Effects of personalized versus generic implementation of an intra-organizational online survey on psychological anonymity and response behavior: A field experiment. *Journal of Business and Psychology*, *29*(2), 169–181. https://doi.org/10.1007/s10869-012-9262-9

Navarro-González, D., Lorenzo-Seva, U., & Vigil-Colet, A. (2016). How response bias affects the factorial structure of personality self-reports. *Psicothema*, *28*(4), 465–470. https://doi.org/10.7334/psicothema2016.113

Nettle, D. (2006). The evolution of personality variation in humans and other animals. *American Psychologist*, *61*(6), 622–631. https://doi.org/10.1037/0003-066X.61.6.622

Niedhammer, I., Bertrais, S., & Witt, K. (2021). Psychosocial work exposures and health outcomes: A meta-review of 72 literature reviews with meta-analysis. *Scandinavian Journal of Work, Environment & Health*, *47*(7), 489–508. https://doi.org/10.5271/sjweh.3968

Nübling, M., Stößel, U., Hasselhorn, H.-M., Michaelis, M., & Hofmann, F. (2005). *Methoden zur Erfassung psychischer Belastungen: Erprobung eines Messinstrumentes (COPSOQ) [Methods for recording mental stress: Testing of a measuring instrument (COPSOQ)]*. Wirtschaftsverlag NW.

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of Personality and Social Psychological Attitudes* (pp. 17–59). Academic Press. https://doi.org/10.1016/B978-0-12-590241-0.50006-X

Perrewé, P. L., & Zellars, K. L. (1999). An examination of attributions and emotions in the transactional approach to the organizational stress process. *Journal of Organizational Behavior*, *20*(5), 739–752. https://doi.org/10.1002/(SICI)1099-1379(199909)20:5<739::AID-JOB1949>3.0.CO;2-C

Pinder, C. C., & Harlos, K. P. (2001). Employee silence: Quiescence and acquiescence as responses to perceived injustice. In K. M. Rowland & G. R. Ferris (Eds.), *Research in personnel and human resources management* (Vol. 20, pp. 331–369). Emerald. https://doi.org/10.1016/S0742-7301(01)20007-3

Rammstedt, B., Goldberg, L. R., & Borg, I. (2010). The measurement equivalence of Big-Five factor markers for persons with different levels of education. *Journal of Research in Personality*, *44*(1), 53–61. https://doi.org/10.1016/j.jrp.2009.10.005

Rammstedt, B., & John, O. P. (2005). Kurzversion des Big Five Inventory (BFI-K) [Short version of the Big Five Inventory (BFI-K)]. *Diagnostica*, *51*(4), 195–206. https://doi.org/10.1026/0012-1924.51.4.195

Reicher, S. D., Spears, R., & Postmes, T. (1995). A social identity model of deindividuation phenomena. *European Review of Social Psychology*, *6*(1), 161–198. https://doi.org/10.1080/14792779443000049

Rogelberg, S. C., Spitzmüeller, C., Little, I., & Reeve, C. L. (2006). Understanding response behavior to an online special topics organizational satisfaction survey. *Personnel Psychology*, *59*(4), 903–923. https://doi.org/10.1111/j.1744-6570.2006.00058.x

Roloff, J., Kirstges, J., Grund, S., & Klusmann, U. (2022). How strongly is personality associated with burnout among teachers? A meta-analysis. *Educational Psychology Review*, *34*(3), 1613–1650. https://doi.org/10.1007/s10648-022-09672-7

Saari, L. M., & Scherbaum, C. A. (2011). Identified employee surveys: Potential promise, perils, and professional practice guidelines. *Industrial and Organizational Psychology*, *4*(4), 435–448. https://doi.org/10.1111/j.1754-9434.2011.01369.x

Schmit, M. J., Ryan, A. M., Stierwalt, S. L., & Powell, A. B. (1995). Frame-of-reference effects on personality scale scores and criterion-related validity. *Journal of Applied Psychology*, *80*(5), 607–620. https://doi.org/10.1037/0021-9010.80.5.607

Schneider, T. R. (2004). The role of neuroticism on psychological and physiological stress responses. *Journal of Experimental Social Psychology*, *40*(6), 795–804. https://doi.org/10.1016/j.jesp.2004.04.005

Schulte, E.-M., Wittner, B., & Kauffeld, S. (2021). Ressourcen und Anforderungen (ReA) in der Arbeitswelt: Entwicklung und erste Validierung eines Fragebogens [Resources and demands in the world of work: Development and first validation of a questionnaire]. *Gruppe. Interaktion. Organisation. Zeitschrift für angewandte Organisationspsychologie (GIO)*, *52*(2), 405–415. https://doi.org/10.1007/s11612-021-00565-x

Schweiger, D. M., & DeNisi, A. S. (1991). Communication with employees following a merger: A longitudinal field experiment. *Academy of Management Journal*, *34*(1), 110–135. https://doi.org/10.5465/256304

Semmer, N. K., Zapf, D., & Dunckel, H. (1999). Instrument zur stressbezogenen Tätigkeitsanalyse ISTA [Instrument for Stress-Oriented Task Analysis ISTA]. In H. Dunckel (Ed.), *Handbuch psychologischer Arbeitsanalyseverfahren* (pp. 179–204). vdf Hochschulverlag.

Singer, E., von Thurn, D. R., & Miller, E. R. (1995). Confidentiality assurances and response: A quantitative review of the experimental literature. *Public Opinion Quarterly*, *59*(1), 66–77. https://doi.org/10.1086/269458

Spector, P. E., & Jex, S. M. (1998). Development of four self-report measures of job stressors and strain: Interpersonal Conflict at Work Scale, Organizational Constraints Scale, Quantitative Workload Inventory, and Physical Symptoms Inventory. *Journal of Occupational Health Psychology*, *3*(4), 356–367. https://doi.org/10.1037/1076-8998.3.4.356

Staufenbiel, T., & König, C. J. (2010). A model for the effects of job insecurity on performance, turnover intention, and absenteeism. *Journal of Occupational and Organizational Psychology*, *83*(1), 101–117. https://doi.org/10.1348/096317908X401912

Tackett, J. L., & Lahey, B. B. (2017). Neuroticism. In T. A. Widiger (Ed.), *The Oxford handbook of the Five Factor Model* (pp. 39–56). Oxford University Press.

Tajfel, H., & Turner, J. (1986). The social identitiy theory of intergroup behavior. In S. Worchel & W. Austin (Eds.), *Psychology of intergroup relations* (pp. 7–24). Nelson-Hall.

Taouk, Y., Spittal, M. J., LaMontagne, A. D., & Milner, A. J. (2020). Psychosocial work stressors and risk of all-cause and coronary heart disease mortality: A systematic review and meta-analysis. *Scandinavian Journal of Work, Environment & Health*, *46*(1), 19–31. https://doi.org/10.5271/sjweh.3854

Therkelsen, D. J., & Fiebich, C. L. (2004). The supervisor: The linchpin of employee relations. *Journal of Communication Management*, *8*(2), 120–129. https://doi.org/10.1108/13632540410807592

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Blackwell.

Van Dyne, L., Ang, S., & Botero, I. C. (2003). Conceptualizing employee silence and employee voice as multidimensional constructs. *Journal of Management Studies*, *40*(6), 1359–1392. https://doi.org/10.1111/1467-6486.00384

Whelan, T. J., & Thompson, L. F. (2009). *Development/validation of the PANON Scale assessing survey anonymity perceptions*. Paper presented at the 24th annual conference of the Society for Industrial and Organizational Psychology, New Orleans, LA.

# 4 General Discussion

The aim of my dissertation was to improve the quality of work stress surveys in order to obtain reliable and meaningful data on work-related stressors and resources. To address the three research questions, I used a mixed-methods approach that combined qualitative and quantitative study designs. The procedure followed a sequential process, where subsequent quantitative studies were informed by the findings of a qualitative interview study.

In summary, my dissertation achieved the following results: First, it presents a comprehensive conceptual model of response behavior in work stress surveys that encompasses cognitive and motivational processes influenced by various contextual factors such as time frames, situational frames, and comparisons. Second, it explores the similarities and differences in response patterns among three perspectives: self-report, supervisor-report, and coworker-report. The investigation includes cognitive and motivational processes, the utilization of comparisons and time frames, and their evaluation. Third, it is shown that the inclusion of highly job-related contextualized items in work stress surveys has a limited impact on criterion validity. In addition, participants attributed higher predictive validity to the contextualized items. Fourth, it suggests that the inclusion of explicit comparisons through tagging influences responses to widely used work stress surveys and has a limited effect on validity. Finally, it quantitatively demonstrates the existence of defensive biasing as a motivational source of bias employed by employees in responding to work stress surveys as a means of protecting themselves from negative consequences, such as ostracism by supervisors. This response behavior is found to be associated with perceived anonymity and neuroticism.

The first research question of this dissertation aimed to gain a comprehensive understanding of the response behavior in work stress surveys, including the differences between self-report and other-report. To address this question, a qualitative research approach (i.e., grounded theory, Glaser & Strauss, 1967) was employed in Study 1 allowing for a thorough exploration of existing knowledge from the work stress and survey literature, as well as the identification and understanding of novel and specific aspects related to responding to work stress surveys. Study 1 yielded valuable and extensive insights, culminating in the development of a conceptual model that enhances our understanding of response behavior in work stress surveys. The model encompasses several key findings, shedding light on various aspects of the survey response process. First, cognitive processes emerged as foundational components of the model, uncovering new theoretical connections and arguments. These processes not only replicated previous research (e.g., Menon, 1994; Tourangeau, 1984) but also unveiled novel insights, emphasizing the significant influence of motivational processes on cognition. Participants reported diverse motivations, such as protecting themselves from negative consequences or alleviating frustration, which shaped their intentional and behavioral responses. Consequently, theories related to impression management (e.g., Bolino et al., 2016) and faking (e.g., Melchers et al., 2020), typically applied in personnel selection settings, may be applicable to comprehending response behavior in stress surveys.

Second, the study underscored the importance of contextual factors in work stress survey item responses, particularly in relation to time frames and social comparisons. The utilization of different time frames was found to impact participants' interpretation and information retrieval processes, potentially leading to variations in responses. This finding is consistent with previous research in this field

(e.g., Debus et al., 2019). Moreover, social comparisons played a substantial role, as individuals compared themselves to colleagues or supervisors when assessing stressors and resources. These findings emphasize the need to incorporate contextual elements into stress theories and consider the interplay between comparisons and fairness theories in understanding stress perceptions. Last, the study delved into the processes involved in other-reports of work-related stress, specifically from supervisors and coworkers. The results support that there are significant discrepancies between these perspectives and self-reports, with supervisors generally underestimating stress levels compared to employees themselves (Debus et al., 2015). This discrepancy may be attributed to supervisors' limited understanding of employee tasks and processes. In contrast, coworker reports exhibited greater convergence with self-reports, likely due to their shared job or similar roles, deeper knowledge of the workplace, and absence of personal consequences. These findings align with role theory, which posits that individuals occupying similar roles are more likely to hold congruent views (Sluss et al., 2011). Overall, the qualitative interviews yielded a substantial amount of findings that partly supported existing research but also provided entirely novel insights, thereby enabling further investigation of these potential effects.

The second research question aimed to explore the impact of contextualization, specifically the level of contextualization and social comparison processes, as cognitive factors that may influence the reliability, validity, and predictive power of work stress surveys in self-reports. Previous literature on personality tests has consistently shown positive effects of contextualization on scale quality, enhancing reliability and validity (e.g., Lievens et al., 2008; Schmit et al., 1995; Swift & Peterson, 2019). Contextualization is believed to standardize

responses, reducing intra- and interindividual variance and ultimately improving the quality of scales. Study 1 provided evidence that contextualization also plays a role in work stress survey items, with respondents individually employing contextual cues for generally formulated items. Building upon these findings, Study 2 and 3 quantitatively demonstrated that different forms of contextualization influence response behavior in work stress surveys and have the potential to enhance scale quality.

The results of Study 2 raise the question of whether contextualization can become too specific. The detailed nature of the items may make it harder rather than easier for respondents to see themselves and their activities reflected in them. Therefore, it may be sufficient to make a reference to the job, department, or position within the company through a tag, rather than adapting each item to a specific context. Furthermore, the focus should be on identifying profession-specific stressors and resources (Bakker & Demerouti, 2017; Menghini & Balducci, 2021). For example, the Instrument for Stress-Related Job Analysis for Hospital Physicians (ISAK, Keller et al., 2013) incorporates specific stressor scales such as "social stressors with patients" and "treatment of patients with difficult courses of disease" into the ISTA scales (Semmer et al., 1999). This instrument has demonstrated reliable and valid results in assessing stressors and resources in physician working conditions in hospitals. Nonetheless, the authors acknowledge that departments within hospitals exhibit heterogeneity in terms of physicians' positions, professional experience, and workplaces (e.g., ward vs. operating room). Consequently, they suggest aggregating the results at the department level to identify problem areas. Workplace-specific contextualization is expected to yield positive effects for work-related stress items, but it is crucial to determine the appropriate level for each specific work group.

Study 3 examined the influence of contextualization through social comparison processes, a method frequently mentioned by the interviewees in Study 1. When confronted with uncertainty, responders tend to engage in social comparisons to make assessments (Festinger, 1954). Contextualization again influenced response behavior in this case by eliciting social comparisons. However, the effects observed were relatively small and did not significantly influence all scales. This may be attributed to the fact that while the tagged items consistently asked participants to engage in a social comparison process, they did not refer to the same person, as each participant had an individual colleague to compare. And this direct colleague, in turn, could have very specific characteristics. Research on the reference-group effect in personality inventories has also indicated that comparison groups can be too specific, resulting in decreased validity (Credé et al., 2010). The most influential reference group was found to be individuals of the same age as the participants (Lenhausen et al., 2022). Using a more general comparison object or group would likely be more appropriate and yield stronger effects. Therefore, it is necessary to identify a suitable reference group for assessing work-related stressors and resources. Despite the interviewees in Study 1 reporting comparisons with colleagues engaged in similar activities, specifying a somewhat broader reference group (e.g., workers in the same industry) reduces intra- and interindividual variance, thus enhancing the reliability and validity of the scales. Additionally, to ensure more objectivity in the comparison processes, one possibility is to create a hypothetical person to whom all participants should compare themselves, as exemplified in certain items of the ISTA (Semmer et al., 1999).

The third research question addressed in this dissertation focuses on a novel motivational process that emerges during the completion of work stress surveys:

defensive biasing. Interview participants reported deliberately presenting their professional situation in a more favorable light than it actually is, motivated by a desire to protect themselves from negative consequences. They expressed fears of being perceived as incapable of handling pressure, being overlooked for promotions, losing visibility with superiors, or facing dismissal or contract non-renewal. While defensive biasing shares similarities with constructs commonly studied in questionnaire and interview literature, such as impression management (e.g., Bolino et al., 2016), faking (e.g., Melchers et al., 2020), or social desirability (e.g., van de Mortel, 2008), it is regarded as a distinct phenomenon due to its proactive nature driven by fear within the specific context of work stress surveys. Study 4 represents an initial exploration of this phenomenon and its associated variables. The findings revealed a correlation between defensive biasing and perceived anonymity and neuroticism, contrary to expectations, trust in the supervisor and leadership quality did not demonstrate a significant influence. Although further investigation is needed to delve into additional influencing factors and effects, it is important to acknowledge that motivational processes, alongside cognitive processes, can impact the measurement of work-related stressors and resources. This recognition highlights the need for improvements not only in survey instruments but also in the overall approach to administering such questionnaires within organizations. Particularly concerning the sensitive nature of stress and mental health, the way in which a work stress measurement is communicated to employees can significantly affect motivational processes, with perceived anonymity playing a pivotal role. For employees and their response behavior, it is crucial to understand the purpose and potential outcomes of a work stress measurement. The development of action plans can serve as a catalyst for enhancing subsequent survey results. Therefore, it is essential to identify the factors

within the general administration that can influence motivational processes and, consequently, the measurement of work stress (Huebner & Zacher, 2022). By gaining insight into these factors, organizations can improve their understanding of the nuances involved in measuring work stress and contribute to the overall well-being of employees.

This dissertation follows a mixed methods research approach, which has become increasingly popular in broader psychological research (Mertens, 2019), and aims to leverage the advantages and mitigate the limitations of both qualitative and quantitative methods (e.g., Venkatesh et al., 2016). The use of a grounded theory approach facilitated an in-depth exploration of the response behavior in work stress surveys, providing valuable insights into the intricate processes involved. Through qualitative interviews, the experiences and perspectives of participants were thoroughly explored, resulting in a comprehensive understanding of the complexities associated with responding to work stress surveys (Misoch, 2019). The integration of quantitative methods in Studies 2, 3, and 4 offers several advantages, including the ability to conduct a comprehensive analysis, generalize findings, test hypotheses, and provide an objective evaluation of the findings (Spector & Pindek, 2016). By leveraging the strengths of both qualitative and quantitative approaches, these investigations build on the findings gained from Study 1 and contribute to a more robust and nuanced understanding of various aspects related to work stress surveys. By supplementing qualitative findings with numerical data, quantitative methods allows for a more comprehensive analysis and a deeper understanding of the phenomenon under investigation. In addition, quantitative approaches facilitate the generalization of the findings to a larger population, thereby increasing the external validity of the results. These methods also provide the means to test hypotheses and

draw meaningful inferences, further strengthening the empirical foundation of the research (Creswell, 2015; Teddlie & Tashakkori, 2011).

While the combination of qualitative and quantitative research methods offers numerous advantages, applying quantitative methods to test and refine qualitative study results can present challenges (Baur, 2019). As evident from the individual results sections of the studies, the quantitative mapping of qualitative findings yielded smaller effects than anticipated. This discrepancy may stem from the inherent complexity of the phenomena being studied, as a quantitative research approach tends to oversimplify intricate concepts by reducing them to numerical values (Morgan, 2013). Furthermore, it is crucial to consider that Studies 2 and 3, despite being built upon qualitative results regarding contextualization, relied on study designs derived from the field of personality questionnaires. While adopting a methodological framework with a proven track record can provide a sense of certainty, it is possible that the actual construct reported by the interviewees was not captured with the same accuracy in these studies. These challenges highlight the need for caution when translating qualitative findings into quantitative measures. The richness and nuance of qualitative data may be difficult to fully capture using numerical approaches alone. It is important for researchers to critically evaluate the fit between the qualitative insights and the quantitative operationalization, ensuring that the essence of the construct under investigation is accurately represented (Heyvaert et al., 2013).

An alternative explanation for the limited effects observed in the quantitative studies could be the potential overestimation of the frequency and intensity of qualitatively reported content. This could be attributed to two principles of the grounded theory methodology: theoretical sampling and theoretical saturation. Theoretical sampling involves actively selecting specific cases or instances to study in

order to develop and refine the theory (Byrne, 2001). It is an iterative process where the researcher collects and analyzes data, using it to guide further sampling. Consequently, qualitative studies typically have smaller sample sizes compared to quantitative studies. Theoretical saturation occurs when further data collection and analysis no longer yield new concepts or categories, indicating that the theory is well-developed and complete (Fusch & Ness, 2015). At this point, the researcher has gathered sufficient data to fully comprehend the phenomenon under investigation, and additional data is unlikely to reveal new insights. Due to the pursuit of gaining new insights, the focus of interview content tends to become more specific over the course of data collection. This means that potential indications of a new phenomenon may emerge during interviews, leading to the selection of specific interviewees who can confirm and describe the phenomenon in greater detail. However, it is important to note that the presence of a phenomenon in the data does not necessarily imply its prevalence in society; it simply suggests that it might occur. Moreover, the qualitative interview setting offers a larger space for reflection compared to the brief moment of completing a quantitative questionnaire.

Additionally, individuals who are inclined towards self-reflection may be more willing to participate in qualitative interviews. This allows for a more nuanced description of constructs that may not be captured in as much detail among mainstream employees or represent their complexity quantitatively (Collins et al., 2005). Social desirability may also play a role in qualitative research, as the interview situation may create an expectation for participants to provide more elaborate accounts compared to the limited resources required for responding to quantitative questionnaires, potentially leading to an overestimation or biasing of reality (Bergen & Labonté, 2020).

In summary, while combining qualitative and quantitative methods can be advantageous, employing quantitative methods to refine qualitative study results requires careful consideration. Researchers must be mindful of the potential oversimplification of complex phenomena and the potential mismatch between qualitative insights and the operationalization in quantitative studies. By acknowledging these challenges, future research can strive for a more comprehensive and accurate integration of qualitative and quantitative approaches (Schonfeld & Mazzola, 2012).

### 4.1 Limitations

In addition to the specific limitations identified within each individual study, there are three broader limitations that pertain to this dissertation. First, the selection of work-related stressor and resource items was constrained to two survey instruments: ISTA (Semmer et al., 1999) and COPSOQ (Kristensen et al., 2005; Nübling et al., 2005). Furthermore, the inclusion of items in the studies was carefully guided by the research questions, ensuring their relevance and applicability to the work context of medical assistants or their suitability for incorporating a social comparison. However, due to the necessity of mitigating participant burden, only a limited number of items were included, resulting in the potential omission of certain stressors and resources. Although both survey instruments align with the overarching J-DR model of work stress (Bakker et al., 2007), they were not explicitly developed within this theoretical framework, which may limit the comprehensive coverage of resources (Schulte et al., 2021). While the selection of validated instruments tailored to the German-speaking context was crucial, future research could benefit from the inclusion of a broader range of items from instruments such as the Energy Compass

(EC, Schaufeli, 2017) and the Resources and Demands Questionnaire (Schulte et al., 2021), which were specifically designed for the German-speaking population.

Second, no items were specifically examined that targeted the psychological and physical manifestations of stress, such as typical burnout symptoms or their precursors, including difficulties in switching off after work, problems with falling asleep or staying asleep, and concentration difficulties. Obtaining information about existing health issues resulting from stress can be crucial for companies, as it allows for the assessment of employees' well-being and the identification of potential early signs of conditions like burnout. Understanding the response patterns related to these aspects would provide valuable insights into the relationship between the response behavior of stressors and resources. This deeper understanding would enable organizations to implement timely interventions and effectively address these issues.

Third, in the second and third studies, the sample composition was focused on employees in the healthcare sector, specifically medical assistants and employees of rehabilitation centers. These occupational groups are known to be exposed to high levels of work stress and encounter specific stress-related situations. Additionally, the healthcare sector is characterized by a significant proportion of female employees (Krystal, 2020). It is important to acknowledge that this sample composition may limit the generalizability of the results, as working conditions and stress levels can vary across different sectors. Therefore, a more diverse sample representing various industries would provide a stronger foundation for drawing valid conclusions about workplace stress analysis.

### 4.2 Future Research and Directions

Future studies should continue to aim at improving and enhancing the quality of work stress surveys in order to ensure that the data obtained closely aligns with

reality, enabling the design and implementation of targeted and effective intervention measures. The first study of this dissertation provided valuable insights into the response behavior of employees, supervisors, and coworkers in work stress surveys, and the qualitative findings present further potential for exploration through quantitative methodologies. Consequently, additional investigations are warranted, particularly regarding the inclusion of other-reports from supervisors and coworkers in the measurement of work stress. While utilizing other-reports can be beneficial in various domains (e.g., counterproductive work behavior, Fox et al., 2007), it has been observed in work stress research that the convergence of stressor ratings between supervisors and employees is influenced by stressor observability, with less observable stressors leading to lower convergence (Debus et al., 2015). Future research should elucidate ways to increase this convergence and explore how other-reports can add value to the measurement of work stress. For example, coworker reports could potentially compensate for the limitations posed by stressor observability. Additionally, as revealed by the results of Study 1, employees, coworkers, and supervisors have different time frames in mind. Standardizing the time frames may help mitigate the discrepancies arising from these varying frames of reference. The influence of induced self-comparison processes and their potential improvement of scale quality compared to social comparisons should also be investigated, considering that the perception of time stress, for instance, has been shown to be influenced by different temporal frames (Debus et al., 2019). Instead of excluding other-reports entirely, given the subjective nature of stress perceptions, further research is needed to determine how other-reports can meaningfully complement self-reports to obtain reliable results. Study 4 suggests that motivational processes are significantly stronger and more pronounced in self-reports compared to

other-reports. Mitigating the bias caused by these motivational processes could be achieved through the utilization of other-reports from supervisors and coworkers.

Another avenue to enhance the quality of work stress measurement is the utilization of digital diaries, a method already employed in occupational health research (e.g., de Bloom et al., 2010). For example, workers could be prompted digitally to assess stressors and resources multiple times a day over the course of a workweek, and these assessments could be aggregated to provide a comprehensive picture. This approach allows for the inclusion of short-term influencing factors by capturing data at multiple time points. Digital tools could also be employed to capture the frequency of specific stressors, such as time stress, by having employees press a buzzer whenever they experience such stressors, and the same procedure could be applied to assess the occurrence of resources, such as support from colleagues. This would enable the examination of the temporal relationship between stressors and resources, shedding light on whether support is provided following the perception of time stress, for example. In addition, the integration of physiological measurement methods, such as heart rate, blood pressure, or cortisol levels, into the survey period alongside digital diaries could further enhance the understanding of psychophysiological responses to work-related stressors (e.g., Buckert et al., 2017; Langer et al., 2022; Schaafsma et al., 2021). The combination of digital diaries and physiological measurements has been utilized to explore correlations in common stress models (e.g., Johnston et al., 2016). It is imperative to investigate whether this approach adds value and improves the quality of work stress measurement, leading to results that align more closely with reality and serve as a foundation for further interventions.

Furthermore, the field of contextualization in work stress surveys still holds numerous unanswered questions. The findings of this dissertation indicate that contextualization has an impact on response behavior and scale quality. The objective of contextualization is to ensure that respondents utilize the same context across all items and that a consistent context is employed across all survey participants (Lievens et al., 2008). However, it remains necessary to determine the appropriate level of contextualization, considering potential variations among different occupational groups. As previously mentioned in the general discussion, the influence of tagged contextualization (medium level, following Holtrop et al., 2014) should be further examined. It is conceivable that adding items referring to the participants' respective company or department may already establish a shared context among respondents.

Contextualization through additional comparisons has also demonstrated an impact on the measurement of work-related stress. Again, the most suitable comparison for participants to reference needs to be identified. Moreover, there is evidence highlighting the importance of self-comparison, such as comparing current working conditions to previous ones, in assessing stress (e.g., Weiß, 2020). It is crucial to investigate the extent to which induced self-comparison processes can contribute to enhancing scale quality compared to social comparisons. The ISTA (Semmer et al., 1999) incorporates a unique question format known as the A-B format, which also triggers a comparison process. In this format, respondents are presented with two hypothetical individuals, A and B, and are asked to indicate which one they identify with more. The benefits and limitations of this method are not fully understood. Some experts suggest that this forced comparison may facilitate the response process, while others argue that it may pose challenges for respondents who cannot easily align themselves with either option. Further research is needed to

deepen our understanding of the relationship between these question types and the decision-making processes they elicit. To address the drawback of respondents being unable to assign themselves to either option, an alternative approach could involve utilizing a single hypothetical person as a reference. This individual would be presented within the context of a case description, ensuring that all participants refer to the same object of comparison across all items.

A similar approach can be found in the use of anchoring vignettes, which aim to uncover different interpretations of items and response scales among respondents and adjust responses accordingly (King & Wand, 2007). Respondents are presented with hypothetical case studies featuring people and their behaviors or experiences, which are then rated on a multilevel scale. Interpretive tendencies can be inferred from response patterns and mathematically extracted from respondents' self-reported ratings, thereby reducing bias. Anchoring vignettes are designed to standardize responses and minimize response bias by providing concrete examples that help respondents understand the meaning of the question and the expected type of response. This can increase the reliability and validity of survey data by reducing response variability and improving the accuracy of results.

Motivational processes in surveys have become an increasingly important area of research due to their potential to influence the data collected (e.g., impression management, Bolino et al., 2016; social desirability, van de Mortel, 2008). The concept of defensive biasing as a motivational process in the context of work stress surveys is a relatively new area of investigation, but it holds great promise for advancing our understanding of the factors that influence response behavior in these surveys. To further advance our understanding of defensive biasing, several key areas of research should be pursued. First, the development and validation of measures of

defensive biasing are critical to understanding the nature and extent of this phenomenon. While an initial four-item scale was developed in this dissertation, further validation with larger and more diverse samples is needed. In addition, the identification of predictors of defensive biasing is an important avenue of research. Perceived anonymity has been identified as a predictor of defensive biasing, and future studies should continue to explore the role of perceived anonymity and other factors that contribute to defensive biasing. Also, the phenomenon of respondents deliberately exacerbating their stressors and depleting their resources out of frustration, in an attempt to effect change warrants further investigation into motivational factors.

Another line of research could be to examine external factors in the administration of work stress surveys, rather than focusing solely on the respondents themselves. For example, the impact of different survey administration practices (Morrel-Samuels, 2002) on response behavior in work stress surveys should be explored. This includes examining the role of communication with employees (Macey & Fink, 2020) and the process of obtaining informed consent and providing education about work stress surveys, which may influence employee response behavior, particularly in terms of perceived anonymity (Mueller et al., 2014). In addition, examining whether the use of external specialists to administer work stress surveys within an organization differs from the use of internal staff may shed light on the impact of external services on perceived anonymity and potential defensive bias. Furthermore, the feasibility of conducting work stress surveys outside the workplace and having them administered by external specialists is also worth investigating.

In conclusion, future studies should aim to refine and improve the quality of work stress surveys by exploring various methodological approaches. This includes

further investigation of other-reports, the use of digital diaries, the integration of physiological measures, and the investigation of contextualization strategies. In addition, the study of defensive bias as a motivational process and the exploration of external factors in survey administration should be pursued. By addressing these areas of research, we can advance our understanding of work-related stress and contribute to the development of effective interventions to promote psychological well-being in the workplace.

### 4.3 General Conclusion

The dissertation significantly contributes to the advancement of knowledge regarding response behavior in work stress surveys and the enhancement of work-related stressor and resource scales. The findings underscore the necessity for more reliable and valid measures that effectively capture employees' experiences of work stress. The assessment of work-related stressors and resources is profoundly influenced by cognitive and motivational processes, as well as various contextualization effects, which present challenges for obtaining objective and accurate data on work stress. Additionally, the studies emphasize the importance of considering response biases in work stress survey research and highlight the potential advantages of utilizing improved measurement scales in workplace health and well-being evaluations. Overall, this research makes a significant contribution to ongoing efforts to understand and manage work stress, particularly with respect to its assessment.

**References**

Alarcon, G. M. (2011). A meta-analysis of burnout with job demands, resources, and attitudes. *Journal of Vocational Behavior*, *79*(2), 549–562. https://doi.org/10.1016/j.jvb.2011.03.007

Arbeitsschutzgesetz [Occupational Safety and Health Act]. (2013). https://www.gesetze-im-internet.de/arbschg/__5.html

Bakker, A. B., & Demerouti, E. (2017). Job demands-resources theory: Taking stock and looking forward. *Journal of Occupational Health Psychology*, *22*(3), 273–285. https://doi.org/10.1037/ocp0000056

Bakker, A. B., Demerouti, E., & Sanz-Vergel, A. (2023). Job demands-resources theory: Ten years later. *Annual Review of Organizational Psychology and Organizational Behavior*, *10*(1), 25–53. https://doi.org/10.1146/annurev-orgpsych-120920-053933

Bakker, A. B., Hakanen, J. J., Demerouti, E., & Xanthopoulou, D. (2007). Job resources boost work engagement, particularly when job demands are high. *Journal of Educational Psychology*, *99*(2), 274–284. https://doi.org/10.1037/0022-0663.99.2.274

Baur, N. (2019). Linearity vs. circularity? On some common misconceptions on the differences in the research process in qualitative and quantitative research. *Frontiers in Education*, *4*, Article 53. https://doi.org/doi: 10.3389/feduc.2019.00053

Bergen, N., & Labonté, R. (2020). "Everything is perfect, and we have no problems": Detecting and limiting social desirability bias in qualitative research. *Qualitative Health Research*, *30*(5), 783–792. https://doi.org/10.1177/1049732319889354

Blair, E., & Burton, S. (1987). Cognitive processes used by survey respondents to answer behavioral frequency questions. *Journal of Consumer Research*, *14*(2), 280–288. https://doi.org/10.1086/209112

Blustein, D. (2006). *The psychology of working: A new perspective for career development, counseling, and public policy*. Routledge. https://doi.org/10.4324/9780203935477

Bolino, M., Long, D., & Turnley, W. (2016). Impression management in organizations: Critical questions, answers, and areas for future research. *Annual Review of Organizational Psychology and Organizational Behavior*, *3*(1), 377–406. https://doi.org/10.1146/annurev-orgpsych-041015-062337

Bradburn, N., Rips, L., & Shevell, S. (1987). Answering autobiographical questions: The impact of memory and inference on surveys. *Science*, *236*(4798), 157–161. https://doi.org/10.1126/science.3563494

Brinsfield, C. T., Edwards, M. S., & Greenberg, J. (2009). Voice and silence in organizations: Historical review and current conceptualizations. In J. Greenberg & M. S. Edwards (Eds.), *Voice and silence in organizations* (pp. 1–33). Emerald.

Buckert, M., Schwieren, C., Kudielka, B. M., & Fiebach, C. J. (2017). How stressful are economic competitions in the lab? An investigation with physiological measures. *Journal of Economic Psychology*, *62*, 231–245. https://doi.org/10.1016/j.joep.2017.07.004

Collins, M., Shattell, M., & Thomas, S. P. (2005). Problematic interviewee behaviors in qualitative research. *Western Journal of Nursing Research*, *27*(2), 188–199. https://doi.org/10.1177/0193945904268068

Crawford, E. R., LePine, J. A., & Rich, B. L. (2010). Linking job demands and resources to employee engagement and burnout: A theoretical extension and meta-analytic test. *Journal of Applied Psychology*, *95*(5), 834–848. https://doi.org/10.1037/a0019364

Credé, M., Bashshur, M., & Niehorster, S. (2010). Reference group effects in the measurement of personality and attitudes. *Journal of Personality Assessment*, *92*(5), 390–399. https://doi.org/10.1080/00223891.2010.497393

Creswell, J. W. (2015). *A concise introduction to mixed methods research*. SAGE Publications.

DAK Gesundheit. (2022). *Psychreport 2022: Entwicklungen der psychischen Erkrankungen im Job: 2011-2021 [Psychreport 2022: Developments in mental illness on the job: 2011-2021]*. https://www.dak.de/dak/bundesthemen/psychreport-2022-2533048.html#/

Davison, H. K., & Bing, M. N. (2009). Content validity does matter for the criterion-related validity of personality tests. *Industrial and Organizational Psychology*, *2*(4), 501–503. https://doi.org/10.1111/j.1754-9434.2009.01182.x

de Bloom, J., Geurts, S. A. E., Taris, T. W., Sonnentag, S., de Weerth, C., & Kompier, M. A. J. (2010). Effects of vacation from work on health and well-being: Lots of fun, quickly gone. *Work & Stress*, *24*(2), 196–216. https://doi.org/10.1080/02678373.2010.493385

de Jonge, J., Bosma, H., Peter, R., & Siegrist, J. (2000). Job strain, effort-reward imbalance and employee well-being: A large-scale cross-sectional study. *Social Science & Medicine*, *50*(9), 1317–1327. https://doi.org/10.1016/S0277-9536(99)00388-3

de Jonge, J., van Vegchel, N., Shimazu, A., Schaufeli, W., & Dormann, C. (2010). A longitudinal test of the demand-control model using specific job demands and specific job control. *International Journal of Behavioral Medicine*, *17*(2), 125–133. https://doi.org/10.1007/s12529-010-9081-1

de Leeuw, E. D., Hox, J., & Dillman, D. (Eds.). (2008). *International handbook of survey methodology*. Routledge. https://doi.org/10.4324/9780203843123

Debus, M. E., Greulich, B., König, C. J., & Kleinmann, M. (2019). Insecure about how to rate your job insecurity? A two-study investigation into time frames applied to job insecurity measures. *Occupational Health Science*, *3*(4), 421–435. https://doi.org/10.1007/s41542-019-00049-x

Debus, M. E., König, C. J., Kleinmann, M., & Werner, C. S. (2015). Examining the effects of negative affectivity on self- and supervisor ratings of job stressors: The role of stressor observability. *Work & Stress*, *29*(4), 341–361. https://doi.org/10.1080/02678373.2015.1075233

Demerouti, E., Bakker, A. B., Nachreiner, F., & Schaufeli, W. B. (2001). The job demands-resources model of burnout. *Journal of Applied Psychology*, *86*(3), 499–512. https://doi.org/10.1037/0021-9010.86.3.499

Dewe, P. J., O'Driscoll, M. P., & Cooper, C. L. (2012). Theories of psychological stress at work. In R. J. Gatchel & I. Z. Schultz (Eds.), *Handbook of Occupational Health and Wellness* (pp. 23–38). Springer. https://doi.org/10.1007/978-1-4614-4839-6_2

Donaldson, S. I., & Grant-Vallone, E. J. (2002). Understanding self-report bias in organizational behavior research. *Journal of Business and Psychology*, *17*(2), 245–260. https://doi.org/10.1023/A:1019637632584

Erazo-Chamorro, V. C., Arciniega-Rocha, R. P., Rudolf, N., Tibor, B., & Gyula, S. (2022). Safety workplace: The prevention of industrial security risk factors. *Applied Sciences*, *12*(21), 10726. https://doi.org/10.3390/app122110726

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*(2), 117–140. https://doi.org/10.1177/001872675400700202

Ford, M. T., Cerasoli, C. P., Higgins, J. A., & Decesare, A. L. (2011). Relationships between psychological, physical, and behavioural health and work performance: A review and meta-analysis. *Work & Stress*, *25*(3), 185–204. https://doi.org/10.1080/02678373.2011.609035

Fox, S., Spector, P. E., Goh, A., & Bruursema, K. (2007). Does your coworker know what you're doing? Convergence of self- and peer-reports of counterproductive work behavior. *International Journal of Stress Management*, *14*(1), 41–60. https://doi.org/10.1037/1072-5245.14.1.41

Fox, S., Spector, P. E., & Miles, D. (2001). Counterproductive work behavior (CWB) in response to job stressors and organizational justice: Some mediator and moderator tests for autonomy and emotions. *Journal of Vocational Behavior*, *59*(3), 291–309. https://doi.org/10.1006/jvbe.2001.1803

Fusch, P., & Ness, L. (2015). Are we there yet? Data saturation in qualitative research. *The Qualitative Report*, *20*(9), 1408–1416. https://doi.org/10.46743/2160-3715/2015.2281

Gerhardt, C., Semmer, N. K., Sauter, S., Walker, A., de Wijn, N., Kälin, W., Kottwitz, M. U., Kersten, B., Ulrich, B., & Elfering, A. (2021). How are social stressors at work related to well-being and health? A systematic review and meta-analysis. *BMC Public Health*, *21*(1), Article 890. https://doi.org/10.1186/s12889-021-10894-7

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Aldine.

Hacker, W. (1973). *Allgemeine Arbeits- und Ingenieurpsychologie [General work and engineering psychology]*. Deutscher Verlag der Wissenschaft.

Hacker, W. (1986). *Arbeitspsychologie [Work psychology]*. Huber.

Hassard, J., Teoh, K. R. H., Visockaite, G., Dewe, P., & Cox, T. (2018). The cost of work-related stress to society: A systematic review. *Journal of Occupational Health Psychology*, *23*(1), 1–17. https://doi.org/10.1037/ocp0000069

Heine, S. J., Buchtel, E. E., & Norenzayan, A. (2008). What do cross-national comparisons of personality traits tell us? The case of conscientiousness. *Psychological Science*, *19*(4), 309–313. https://doi.org/10.1111/j.1467-9280.2008.02085.x

Heyvaert, M., Hannes, K., Maes, B., & Onghena, P. (2013). Critical appraisal of mixed methods studies. *Journal of Mixed Methods Research*, *7*(4), 302–327. https://doi.org/10.1177/1558689813479449

Hill, N. L., Mogle, J., Whitaker, E. B., Gilmore-Bykovskyi, A., Bhargava, S., Bhang, I. Y., Sweeder, L., Tiwari, P. A., & Van Haitsma, K. (2019). Sources of response bias in cognitive self-report items: “Which memory are you talking about?” *The Gerontologist*, *59*(5), 912–924. https://doi.org/10.1093/geront/gny087

Holtrop, D., Born, M. Ph., de Vries, A., & de Vries, R. E. (2014). A matter of context: A comparison of two types of contextualized personality measures. *Personality and Individual Differences*, *68*, 234–240. https://doi.org/10.1016/j.paid.2014.04.029

Holtz, B. C., Ployhart, R. E., & Dominguez, A. (2005). Testing the rules of justice: The effects of frame-of-reference and pre-test validity information on personality test responses and test perceptions. *International Journal of Selection and Assessment*, *13*, 75–86. https://doi.org/10.1111/j.0965-075X.2005.00301.x

Hu, Q., Schaufeli, W. B., & Taris, T. W. (2017). How are changes in exposure to job demands and job resources related to burnout and engagement? A longitudinal study among Chinese nurses and police officers. *Stress and Health*, *33*(5), 631–644. https://doi.org/10.1002/smi.2750

Huebner, L.-A., & Zacher, H. (2022). Effects of action planning after employee surveys. *Journal of Personnel Psychology*, *21*(1), 23–36. https://doi.org/10.1027/1866-5888/a000285

Ingold, P. V., Kleinmann, M., König, C. J., & Melchers, K. G. (2015). Shall we continue or stop disapproving of self-presentation? Evidence on impression management and faking in a selection context and their relation to job performance. *European Journal of Work and Organizational Psychology*, *24*(3), 420–432. https://doi.org/10.1080/1359432X.2014.915215

Irmer, J. P., Kern, M., Schermelleh-Engel, K., Semmer, N. K., & Zapf, D. (2019). The Instrument for Stress-Oriented Task Analysis (ISTA): A meta-analysis. *Zeitschrift für Arbeits- und Organisationspsychologie*, *63*(4), 217–237. https://doi.org/10.1026/0932-4089/a000312

Johnson, J. V., & Hall, E. M. (1988). Job strain, work place social support, and cardiovascular disease: A cross-sectional study of a random sample of the Swedish working population. *American Journal of Public Health*, *78*(10), 1336–1342. https://doi.org/10.2105/ajph.78.10.1336

Johnston, D., Bell, C., Jones, M., Farquharson, B., Allan, J., Schofield, P., Ricketts, I., & Johnston, M. (2016). Stressors, appraisal of stressors, experienced stress and cardiac response: A real-time, real-life investigation of work stress in nurses. *Annals of Behavioral Medicine*, *50*(2), 187–197. https://doi.org/10.1007/s12160-015-9746-8

Karasek, R. A. (1979). Job demands, job decision latitude, and mental strain: Implications for job redesign. *Administrative Science Quarterly*, *24*(2), 285–308. https://doi.org/10.2307/2392498

Karasek, R. A., & Theorell, T. (1990). *Healthy work: Stress productivity and reconstruction of working life*. Basic Books.

Keller, M., Bamberg, E., Kersten, M., & Nienhaus, A. (2013). Validierung des Instruments zur stressbezogenen Arbeitsanalyse für Klinikärztinnen und -ärzte (ISAK) [Validation of the Instrument for Stress-Related Job Analysis for Hospital Physicians]. *Zeitschrift für Arbeits- und Organisationspsychologie A&O*, *57*(1), 3–21. https://doi.org/10.1026/0932-4089/a000096

King, G., & Wand, J. (2007). Comparing incomparable survey responses: Evaluating and selecting anchoring vignettes. *Political Analysis*, *15*(1), 46–66. https://doi.org/10.1093/pan/mpl011

Kocakulah, M. C., Kelley, A. G., Mitchell, K. M., & Ruggieri, M. P. (2016). Absenteeism problems and costs: Causes, effects and cures. *International Business & Economics Research Journal*, *15*(3), 89–96. https://doi.org/10.19030/iber.v15i3.9673

Koolhaas, J. M., Bartolomucci, A., Buwalda, B., de Boer, S. F., Flügge, G., Korte, S. M., Meerlo, P., Murison, R., Olivier, B., Palanza, P., Richter-Levin, G., Sgoifo, A., Steimer, T., Stiedl, O., van Dijk, G., Wöhr, M., & Fuchs, E. (2011). Stress revisited: A critical evaluation of the stress concept. *Neuroscience & Biobehavioral Reviews*, *35*(5), 1291–1301. https://doi.org/10.1016/j.neubiorev.2011.02.003

Kristensen, T. S., Borritz, M., Villadsen, E., & Christensen, K. B. (2005). The Copenhagen Burnout Inventory: A new tool for the assessment of burnout. *Work & Stress*, *19*(3), 192–207. https://doi.org/10.1080/02678370500297720

Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality & Quantity*, *47*(4), 2025–2047. https://doi.org/10.1007/s11135-011-9640-9

Krystal, J. H. (2020). Responding to the hidden pandemic for healthcare workers: Stress. *Nature Medicine*, *26*(5), 639–639. https://doi.org/10.1038/s41591-020-0878-4

Langer, M., König, C. J., Siegel, R., Fredenhagen, T., Schunck, A. G., Hähne, V., & Baur, T. (2022). Vocal-stress diary: A longitudinal investigation of the association of everyday work stressors and human voice features. *Psychological Science*, *33*(7), 1027–1039. https://doi.org/10.1177/09567976211068110

Lazarus, R. S., & Folkman, S. (1987). Transactional theory and research on emotions and coping. *European Journal of Personality*, *1*(3), 141–169. https://doi.org/10.1002/per.2410010304

Leka, S., Jain, A., & Lerouge, L. (2017). Work-related psychosocial risks: Key definitions and an overview of the policy context in Europe. In L. Lerouge (Ed.), *Psychosocial risks in labour and social security law: A comparative legal overview from Europe, North America, Australia and Japan* (pp. 1–12). Springer International Publishing. https://doi.org/10.1007/978-3-319-63065-6_1

Lenhausen, M. R., Hopwood, C. J., & Bleidorn, W. (2022). Nature and impact of reference group effects in personality assessment data. *Journal of Personality Assessment*, Advance online publication. https://doi.org/10.1080/00223891.2022.2132504

Lesener, T., Gusy, B., & Wolter, C. (2019). The job demands-resources model: A meta-analytic review of longitudinal studies. *Work & Stress*, *33*(1), 76–103. https://doi.org/10.1080/02678373.2018.1529065

Lievens, F., de Corte, W., & Schollaert, E. (2008). A closer look at the frame-of-reference effect in personality scale scores and validity. *Journal of Applied Psychology*, *93*(2), 268–279. https://doi.org/10.1037/0021-9010.93.2.268

Macey, W. H., & Fink, A. A. (Eds.). (2020). *Employee surveys and sensing: Challenges and opportunities*. Oxford University Press. https://doi.org/10.1093/oso/9780190939717.001.0001

Meijman, T. F., & Mulder, G. (1998). Psychological aspects of workload. In P. J. D. Drenth, H. Thierry, & C. J. de Wolff (Eds.), *A handbook of work and organizational psychology* (pp. 5–33). Psychology Press.

Melchers, K. G., Roulin, N., & Buehl, A.-K. (2020). A review of applicant faking in selection interviews. *International Journal of Selection and Assessment*, *28*(2), 123–142. https://doi.org/10.1111/ijsa.12280

Menghini, L., & Balducci, C. (2021). The importance of contextualized psychosocial risk indicators in workplace stress assessment: Evidence from the healthcare sector. *International Journal of Environmental Research and Public Health*, *18*(6), 3263. https://doi.org/10.3390/ijerph18063263

Menon, G. (1994). Judgments of behavioral frequencies: Memory search and retrieval strategies. In N. Schwarz & S. Sudman (Eds.), *Autobiographical memory and the validity of retrospective reports* (pp. 161–172). Springer.

Mertens, D. M. (2019). *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods*. SAGE Publications. https://study.sagepub.com/mertens4e

Misoch, S. (2019). *Qualitative Interviews [Qualitative interviews]*. De Gruyter Oldenburg. https://doi.org/10.1515/9783110545982

Mitchell, R. J., & Bates, P. (2011). Measuring health-related productivity loss. *Population Health Management*, *14*(2), 93–98. https://doi.org/10.1089/pop.2010.0014

Morgan, D. L. (2013). *Integrating qualitative and quantitative methods: A pragmatic approach*. SAGE Publications. https://doi.org/10.4135/9781544304533

Morgeson, F. P., Brannick, M. T., & Levine, E. L. (2019). *Job and work analysis: Methods, research, and applications for human resource management* (3rd ed.). SAGE Publications. https://us.sagepub.com/en-us/nam/job-and-work-analysis/book258522

Morgeson, F. P., & Campion, M. A. (1997). Social and cognitive sources of potential inaccuracy in job analysis. *Journal of Applied Psychology*, *82*(5), 627–655. https://doi.org/10.1037//0021-9010.82.5.627

Morgeson, F. P., & Campion, M. A. (2012). A framework of potential sources of inaccuracy in job analysis. In M. A. Wilson (Ed.), *The handbook of work analysis: The methods, systems, applications and science of work measurement in organizations* (pp. 593–601). Routledge.

Morgeson, F. P., & Dierdorff, E. C. (2011). Work analysis: From technique to theory. In S. Zedeck (Ed.), *APA handbook of industrial and organizational psychology, Vol 2: Selecting and developing members for the organization* (pp. 3–41). American Psychological Association. https://doi.org/10.1037/12170-001

Morgeson, F. P., & Humphrey, S. E. (2006). The Work Design Questionnaire (WDQ): Developing and validating a comprehensive measure for assessing job design and the nature of work. *Journal of Applied Psychology*, *91*(6), 1321–1339. https://doi.org/10.1037/0021-9010.91.6.1321

Morrel-Samuels, P. (2002, February). Getting the truth into workplace surveys. *Harvard Business Review*, *80*(2), 111–118.

Mueller, K., Straatmann, T., Hattrup, K., & Jochum, M. (2014). Effects of personalized versus generic implementation of an intra-organizational online survey on psychological anonymity and response behavior: A field experiment. *Journal of Business and Psychology*, *29*(2), 169–181. https://doi.org/10.1007/s10869-012-9262-9

Nahrgang, J. D., Morgeson, F. P., & Hofmann, D. A. (2011). Safety at work: A meta-analytic investigation of the link between job demands, job resources, burnout, engagement, and safety outcomes. *Journal of Applied Psychology*, *96*(1), 71–94. https://doi.org/10.1037/a0021484

Niedhammer, I., Bertrais, S., & Witt, K. (2021). Psychosocial work exposures and health outcomes: A meta-review of 72 literature reviews with meta-analysis. *Scandinavian Journal of Work, Environment & Health*, *47*(7), 489–508. https://doi.org/10.5271/sjweh.3968

Nübling, M., Stößel, U., Hasselhorn, H. M., Michaelis, M., & Hofmann, F. (2005). COPSOQ: Ein Instrument zur Messung psychischer Belastungen am Arbeitsplatz [COPSOQ: An instrument for measuring mental stress in the workplace]. *Gesundheitswesen*, *67*, V60. https://doi.org/10.1055/s-2005-920554

O'Connor, D. B., Thayer, J. F., & Vedhara, K. (2021). Stress and health: A review of psychobiological processes. *Annual Review of Psychology*, *72*(1), 663–688. https://doi.org/10.1146/annurev-psych-062520-122331

OECD. (2015). *Fit mind, fit job: From evidence to practice in mental health and work*. OECD Publishing. https://doi.org/10.1787/9789264228283-en.

OECD. (2018). *Health at a glance: Europe 2018: State of health in the EU cycle*. OECD Publishing. https://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-europe-2018_health_glance_eur-2018-en

Parker, S. K. (2014). Beyond motivation: Job and work design for development, health, ambidexterity, and more. *Annual Review of Psychology*, *65*(1), 661–691. https://doi.org/10.1146/annurev-psych-010213-115208

Paulhus, D. L., & Vazire, S. (2007). The self-report method. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personallity psychology* (pp. 224–239). Guilford Press.

Perrewé, P. L., & Zellars, K. L. (1999). An examination of attributions and emotions in the transactional approach to the organizational stress process. *Journal of Organizational Behavior*, *20*(5), 739–752. https://doi.org/10.1002/(SICI)1099-1379(199909)20:5<739::AID-JOB1949>3.0.CO;2-C

Quick, J. C., & Tetrick, L. E. (Eds.). (2011). *Handbook of occupational health psychology* (2nd ed.). American Psychological Association.

Rattrie, L. T. B., & Kittler, M. G. (2014). The job demands-resources model and the international work context: A systematic review. *Journal of Global Mobility*, *2*(3), 260–279. https://doi.org/10.1108/JGM-06-2014-0018

Rigó, M., Dragano, N., Wahrendorf, M., Siegrist, J., & Lunau, T. (2021). Work stress on rise? Comparative analysis of trends in work stressors using the European working conditions survey. *International Archives of Occupational and Environmental Health*, *94*(3), 459–474. https://doi.org/10.1007/s00420-020-01593-8

Rom, O., & Reznick, A. Z. (2016). The stress reaction: A historical perspective. In M. Pokorski (Ed.), *Respiratory contagion* (pp. 1–4). Springer International Publishing. https://doi.org/10.1007/5584_2015_195

Schaafsma, F. G., Hulsegge, G., Jong, M. A., Overvliet, J., Rossum, E. F. C., & Nieuwenhuijsen, K. (2021). The potential of using hair cortisol to measure chronic stress in occupational healthcare: A scoping review. *Journal of Occupational Health*, *63*(1), Article e12189. https://doi.org/10.1002/1348-9585.12189

Schaufeli, W. B. (2017). Applying the job demands-resources model: A ‘how to’ guide to measuring and tackling work engagement and burnout. *Organizational Dynamics*, *46*(2), 120–132. https://doi.org/10.1016/j.orgdyn.2017.04.008

Schaufeli, W. B., Bakker, A. B., & Rhenen, W. V. (2009). How changes in job demands and resources predict burnout, work engagement, and sickness absenteeism. *Journal of Organizational Behavior*, *30*(7), 893–917. https://doi.org/10.1002/job.595

Schmit, M. J., Ryan, A. M., Stierwalt, S. L., & Powell, A. B. (1995). Frame-of-reference effects on personality scale scores and criterion-related validity. *Journal of Applied Psychology*, *80*(5), 607–620. https://doi.org/10.1037/0021-9010.80.5.607

Schonfeld, I. S., & Mazzola, J. J. (2012). Strengths and limitations of qualitative approaches to research in occupational health psychology. In R. R. Sinclair, M. Wang, & L. E. Tetrick (Eds.), *Research methods in occupational health psychology* (pp. 292–313). Routledge. https://doi.org/10.4324/9780203095249-27

Schuler, R. S. (1982). An integrative transactional process model of stress in organizations. *Journal of Occupational Behaviour*, *3*(1), 5–19. https://doi.org/10.1002/job.4030030103

Schulte, E.-M., Wittner, B., & Kauffeld, S. (2021). Ressourcen und Anforderungen (ReA) in der Arbeitswelt: Entwicklung und erste Validierung eines Fragebogens [Resources and demands in the world of work: Development and first validation of a questionnaire]. *Gruppe. Interaktion. Organisation. Zeitschrift für angewandte Organisationspsychologie (GIO)*, *52*(2), 405–415. https://doi.org/10.1007/s11612-021-00565-x

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, *54*, 93–105. https://doi.org/10.1037/0003-066X.54.2.93

Schwarz, N. (2007). Cognitive aspects of survey methodology. *Applied Cognitive Psychology*, *21*(2), 277–287. https://doi.org/10.1002/acp.1340

Selye, H. (1976). The stress concept. *Canadian Medical Association Journal*, *115*(8), 718.

Semmer, N. K., Zapf, D., & Dunckel, H. (1999). Instrument zur stressbezogenen Tätigkeitsanalyse ISTA [Instrument for Stress-Oriented Task Analysis ISTA]. In H. Dunckel (Ed.), *Handbuch psychologischer Arbeitsanalyseverfahren* (pp. 179–204). vdf Hochschulverlag.

Shaffer, J. A., & Postlethwaite, B. E. (2012). A matter of context: A meta-analytic investigation of the relative validity of contextualized and noncontextualized personality measures. *Personnel Psychology*, *65*(3), 445–494. https://doi.org/10.1111/j.1744-6570.2012.01250.x

Siegrist, J. (1996). Adverse health effects of high-effort/low-reward conditions. *Journal of Occupational Health Psychology*, *1*(1), 27–41. https://doi.org/10.1037/1076-8998.1.1.27

Sluss, D. M., van Dick, R., & Thompson, B. S. (2011). Role theory in organizations: A relational perspective. In S. Zedeck (Ed.), *APA handbook of industrial and organizational psychology, Vol 1: Building and developing the organization* (pp. 505–534). American Psychological Association.

Spector, P. E. (2021). *Industrial and organizational psychology: Research and practice* (8th ed.). John Wiley & Sons.

Spector, P. E., & Eatough, E. (2013). Quantitative self-report methods in occupational health psychology research. In M. Wang, S. Sinclair, & L. E. Tetrick (Eds.), *Research methods in occupational health psychology* (pp. 248–267). Taylor & Francis.

Spector, P. E., & Pindek, S. (2016). The future of research methods in work and occupational health psychology. *Applied Psychology*, *65*(2), 412–431. https://doi.org/10.1111/apps.12056

Swift, V., & Peterson, J. B. (2019). Contextualization as a means to improve the predictive validity of personality models. *Personality and Individual Differences*, *144*, 153–163. https://doi.org/10.1016/j.paid.2019.03.007

Taibi, Y., Metzler, Y. A., Bellingrath, S., & Müller, A. (2021). A systematic overview on the risk effects of psychosocial work characteristics on musculoskeletal disorders, absenteeism, and workplace accidents. *Applied Ergonomics*, *95*, 103434. https://doi.org/10.1016/j.apergo.2021.103434

Teddlie, C., & Tashakkori, A. (2011). Mixed method research: Contemporary Issues in an Emerging Field. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (pp. 285–299). SAGE.

Tourangeau, R. (1984). Cognitive science survey methods: A cognitive perspective. In T. Jabine, M. Straf, J. Tanur, & R. Tourangeau (Eds.), *Cognitive aspects of survey methodology: Building a bridge between disciplines* (pp. 73–100). National Academy Press.

van de Mortel, T. F. (2008). Faking it: Social desirability response bias in self-report research. *The Australian Journal of Advanced Nursing*, *25*(4), 40–48. https://doi.org/10.3316/informit.210155003844269

van Vegchel, N., de Jonge, J., Bosma, H., & Schaufeli, W. (2005). Reviewing the effort-reward imbalance model: Drawing up the balance of 45 empirical studies. *Social Science & Medicine*, *60*(5), 1117–1131. https://doi.org/10.1016/j.socscimed.2004.06.043

van Wingerden, J., Bakker, A. B., & Derks, D. (2017). Fostering employee well-being via a job crafting intervention. *Journal of Vocational Behavior*, *100*, 164–174. https://doi.org/10.1016/j.jvb.2017.03.008

Venkatesh, V., Brown, S., & Sullivan, Y. (2016). Guidelines for conducting mixed-methods research: An extension and illustration. *Journal of the Association for Information Systems*, *17*(7), 435–494. https://doi.org/10.17705/1jais.00433

Viswesvaran, C., Sanchez, J. I., & Fisher, J. (1999). The role of social support in the process of work stress: A meta-analysis. *Journal of Vocational Behavior*, *54*(2), 314–334. https://doi.org/10.1006/jvbe.1998.1661

Volpert, W. (1982). The model of the hierarchical-sequential organization of action. In W. Hacker, W. Volpert, & M. V. Cranach (Eds.), *Cognitive and motivational aspects of action* (pp. 35–51). Deutscher Verlag der Wissenschaft.

Weiß, E.-E. (2020). Worktime control and work stress: The moderating effect of self-comparisons and social comparisons. *The International Journal of Human Resource Management*, *31*(5), 682–704. https://doi.org/10.1080/09585192.2017.1365747

Wheaton, B. (1999). The nature of stressors. In A. V. Horwitz & Scheid (Eds.), *A handbook for the study of mental health: Social contexts, theories, and systems* (pp. 176–197). Cambridge University Press.

Xanthopoulou, D., Bakker, A. B., Dollard, M. F., Demerouti, E., Schaufeli, W. B., Taris, T. W., & Schreurs, P. J. G. (2007). When do job demands particularly predict burnout? The moderating role of job resources. *Journal of Managerial Psychology*, *22*(8), 766–786. https://doi.org/10.1108/02683940710837714