Received: 30 September 2016 | Revised: 17 January 2018 | Accepted: 24 January 2018

DOI: 10.1002/job.2271

SPECIAL ISSUE ARTICLE

WILEY Journal of Organizational Behavior

A time to trust? The buffering effect of trust and its temporal variations in the context of high-reliability teams

Michael J. Burtscher[1] | Bertolt Meyer[2] | Klaus Jonas[1] | Sebastian Feese[3] | Gerhard Tröster[3]

[1] Department of Psychology, University of Zurich, Zürich, Switzerland

[2] Department of Psychology, Chemnitz University of Technology, Chemnitz, Germany

[3] Department of Information Technology and Electrical Engineering, ETH Zurich, Zürich, Switzerland

Correspondence
Michael J. Burtscher, Social and Business Psychology, Department of Psychology, University of Zurich, Binzmuehlestrasse 14/13, 8050 Zurich, Switzerland.
Email: m.burtscher@psychologie.uzh.ch

Funding information
Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: CR12I1_137741

Summary

This study aims to further clarify the functionality of job resources in the context of high-reliability teams. Combining extant stress models with theoretical considerations from team research, we address temporal variations in the buffering effect of trust in teammates. We hypothesize that trust buffers the negative effect of objective physical activity on perceived strain and that this buffering effect is more pronounced during later performance episodes (i.e., when employees complete a series of temporally distinguishable tasks). We tested the hypotheses with a sample of professional firefighters who completed a sequence of 3 performance episodes in a high-fidelity simulation environment. Each participant was equipped with a smartphone capturing individual motion activity, which we used as an indicator of physical activity. In line with our hypotheses, multilevel modeling revealed a buffering effect of trust on the relationship between physical activity and perceived strain. Importantly, this buffering effect was more pronounced in the second performance episode as compared with the first performance episode. Our findings add a temporal perspective to the understanding of the effectiveness of job resources. In addition, the current study illustrates the usefulness of smartphones for obtaining behavioral data in a naturalistic setting.

KEYWORDS

high-reliability occupations, social sensing, teams, time, trust

1 | INTRODUCTION

In some occupations, employees are regularly pushed to their limits and beyond. Emergency physicians experience time pressure and a turbulent work environment, police officers face life-threatening situations, and firefighters have to cope with extreme heat and smoke. These occupations have been jointly referred to as high-reliability occupations—complex work environments rich with the potential for error, where errors have serious consequences (Weick, Sutcliffe, & Obstfeld, 1999). Central characteristics of high-reliability occupations include (a) that many tasks are performed in teams (e.g., Flin, O'Connor, & Mearns, 2002; Wilson, Burke, Priest, & Salas, 2005) and (b) that these teams perform their work in recurring performance episodes (Ishak & Ballard, 2012; Marks, Mathieu, & Zaccaro, 2001). The current study aims to improve our understanding of how these characteristics interact to affect work outcomes. In particular, we investigate temporal variations in the buffering effect of job resources across multiple performance episodes.

Job resources are assumed to buffer the negative impact of job demands by affecting the perceptions evoked by said demands (Bakker & Demerouti, 2007; Kahn & Byosiere, 1992). In the team context, trust constitutes a key variable that affects a variety of processes and outcomes (e.g., De Jong, Dirks, & Gillespie, 2016; Mathieu, Maynard, Rapp, & Gilson, 2008; Salas, Sims, & Burke, 2005). Importantly, when people work together, trust influences one person's perception of their colleagues' actions or events related to those actions (Dirks & Ferrin, 2001). Following this notion, we suggest that trust in teammates represents a resource capable of affecting work perceptions in high-reliability teams; in particular, we propose that trust buffers the negative effects of objective physical activity on perceived strain.

Although research has provided useful insights into the work of high-reliability teams (e.g., Salas & Rosen, 2013), little is known about the impact of recurring performance episodes. This is a problem because recurring performance episodes constitute an integral part of the work routine in high-reliability occupations and they

have wide-ranging consequences for team members (Ilgen, Hollenbeck, Johnson, & Jundt, 2005; Ishak & Ballard, 2012; Mathieu et al., 2008). We address this issue by investigating temporal variations in the buffering effect of trust across multiple performance episodes. Combining the job demands–resources model (JD-R; Bakker & Demerouti, 2007; Demerouti, Bakker, Nachreiner, & Schaufeli, 2001) with both Hobfoll's (1989, 2002) conservation of resources (COR) theory and theoretical considerations from team research (e.g., Ilgen et al., 2005; Ishak & Ballard, 2012), we propose that trust in teammates will be more important during later performance episodes. This is because employees need to spend a significant amount of physical resources to complete a first performance episode, which increases their motivation to look for alternative resources and hence will make trust more salient later on.

We test our propositions in the context of firefighting. Given their similarities, we expect that our argumentation can be applied to other high-reliability occupations (Baumann, Gohm, & Bonner, 2011). We envision three main contributions of the current research. First, we extend the theoretical understanding regarding how psychological job resources shape employees' work outcomes across multiple performance episodes. By combining stress models with theoretical considerations from team research, the current study addresses temporal variations in the buffering effect of job resources. We consider this as a compelling extension of extant models of occupational stress that could both stimulate empirical research and further theorize regarding the effectiveness of job resources. Second, the current study clarifies the role of trust in teammates as a remedy for attenuating some of the negative effects of working in high-reliability occupations. Given the high stakes in these occupations and the inevitability of high demands, this represents an important contribution that could help employees to complete their tasks successfully in spite of adverse working conditions. Finally, this study makes a methodological contribution by illustrating how smartphones can be used to capture behavioral data in a naturalistic environment (e.g., Miller, 2012).

2 | THEORETICAL BACKGROUND AND HYPOTHESES

2.1 | High-reliability occupations: a challenging work environment

Employees in high-reliability occupations have to complete complex, highly interdependent, time-limited tasks (e.g., rescue missions, firefighting, and complex surgery), often under adverse conditions, and their failures can have fatal consequences (Baker, Day, & Salas, 2006; Weick et al., 1999). One central characteristic of these occupations concerns the role of teams: Teamwork is considered vital for maintaining performance and safety (Baker et al., 2006; Flin et al., 2002; Wilson et al., 2005). As another central characteristic, performance episodes are an integral part of many high-reliability teams' work routine (Ishak & Ballard, 2012; Sundstrom, McIntyre, Halfhill, & Richards, 2000). Importantly, these performance episodes are recurring; that is, a specific team will conduct several performance episodes during which its members engage in similar patterns of activity (Marks et al., 2001; Mathieu et al., 2008). For example, over the course of a day, an emergency medical team usually deals with multiple incidents, each representing a distinct performance episode.

The fact that high-reliability teams perform their work in recurring performance episodes has several important implications. For one, outcomes of one performance episode can serve as inputs for subsequent episodes (Ilgen et al., 2005; Mathieu et al., 2008): If, for example, firefighters are exhausted from an incident, this might impede their ability to successfully handle a subsequent incident. In addition, performance episodes are characterized by epochality and finality (Ishak & Ballard, 2012). Epochality includes the notion that unlike employees in other occupations, high-reliability teams cannot simply take a break or postpone a task. Finality refers to the irreversibility of work outcomes: If firefighters fail to extinguish a fire in time, they cannot redo their work at a later time.

Research has provided important insights into the functionality of high-reliability teams including the role of leadership (Klein, Ziegert, Knight, & Xiao, 2006), team learning (Ishak & Williams, 2017), and training (Salas & Rosen, 2013). However, explicit efforts to understand the impact of recurring performance episodes are notably absent from the literature. Given the wide-ranging implications of recurring performance episodes noted earlier and their ubiquity in high-reliability occupations, we believe that investigating how performance episodes affect team members' work outcomes constitutes a valuable extension of the literature. In the current study, we focus on team members' appraisal of the specific demands associated with recurring performance episodes, that is, their perceived strain.

2.2 | Job characteristics of high-reliability occupations

The JD-R classifies occupational characteristics into two categories: job demands and job resources (Bakker & Demerouti, 2007; Demerouti et al., 2001). On the one hand, job demands—also termed stressors—include the "physical, social, or organizational aspects of the job that require sustained physical or mental effort and are therefore associated with certain physiological and psychological costs" (Demerouti et al., 2001, p. 501). High work pressure, adverse physical working conditions, or difficult client interactions represent typical job demands (Bakker & Demerouti, 2007). On the other hand, job resources refer to the physical, psychological, social, or organizational aspects of the job that are functional in achieving goals, reduce job demands and the associated costs, or stimulate personal growth and development (Demerouti et al., 2001). This includes a great variety of factors such as career opportunities, support from supervisors, and autonomy (Bakker & Demerouti, 2007; Bakker, Demerouti, & Euwema, 2005). In fact, job resources have been described as "anything perceived by the individual to help attain his or her goals" (Halbesleben, Neveu, Paustian-Underdahl, & Westman, 2014, p. 1338). Given these broad definitions, it seems necessary to specify job demands and resources in high-reliability occupations.

The importance of job demands and resources in a certain occupation depends on the specific job characteristics (Bakker et al., 2005). Whereas job demands vary across high-reliability occupations, many

settings have a significant physical component. For example, physicians and nurses in an emergency department have to lift patients, perform cardiopulmonary resuscitation, and run from one room to another. In this context, firefighting represents an occupation that is especially physically demanding (e.g., Bos, Mol, Visser, & Frings-Dresen, 2004; Mamen, Oseland, & Medbo, 2013). Firefighting tasks involve running, climbing, and lifting (e.g., removing barriers and carrying victims), which require continuous physical activity from the firefighters. In sum, we submit that physical activity constitutes an important job demand in high-reliability occupations, particularly in firefighting.

In terms of job resources of high-reliability teams, trust in teammates represents a potentially important resource because trust constitutes a central determinant of effective teamwork (De Jong et al., 2016; Mathieu et al., 2008; Salas et al., 2005). Trust has been defined as "a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behavior of another" (Rousseau, Sitkin, Burt, & Camerer, 1998, p. 395). With regard to team settings, trust comprises the expectation that one's teammates will perform specific actions that are important to the team, irrespective of one's ability to monitor or control their actions (Breuer, Hüffmeier, & Hertel, 2016). The positive expectations forming the basis of trust can be attributed to both cognitive and affective antecedents (e.g., Colquitt, LePine, Zapata, & Wild, 2011).[1] Whereas cognitive antecedents include beliefs in others' ability and reliability, affective antecedents include beliefs in others' benevolence and concern (Colquitt et al., 2011; Mayer, Davis, & Schoorman, 1995; McAllister, 1995).

Trust is particularly important in high-reliability teams because individual members' tasks are highly interdependent; that is, all members' contributions are necessary for completing the task, and each member must take action for other members to do any part of their work (Wageman, 1995). For example, in a firefighting team, Team Member A, who is tasked with extinguishing the fire, cannot complete his work without the support of Team Member B, who is tasked with securing water supply. Thus, Team Member A has to be confident about Team Member B's ability (i.e., he has the necessary skills) and benevolence (i.e., he would not intentionally do anything that would compromise A). Without trust, task completion would become very inefficient, because Member A would be inclined to closely monitor B (e.g., Langfred, 2004) and second-guess his decisions (e.g., choice of fire hydrant). In line with this argument, research has highlighted the importance of trust in high-reliability teams (Colquitt et al., 2011; Wildman et al., 2012).

2.3 | Negative consequences of high physical demands

The negative consequences of job demands are well documented (e.g., LePine, Podsakoff, & LePine, 2005; Schaufeli & Bakker, 2004; Spector, Dwyer, & Jex, 1988). Extant stress models have specified the process leading from objective job demands to negative work outcomes (e.g., Lazarus & Folkman, 1984). For example, Spector (2002) proposed that job characteristics first have to be perceived as stressors, which in turn leads to negative emotions. These emotions then lead to different types of distal strain reactions—both behavioral (e.g., alcohol abuse) and physical (e.g., health conditions).

In this context, it is necessary to detail how job demands and strain are conceptualized in the context of the current research. Daniels (2006) proposed to differentiate between perceived and enacted job characteristics with the latter including *observable behavioral activities in the job as they happen*. Importantly, these enacted characteristics—and not enduring perceptions or structural features of work—are considered to be the locus of the appraisal processes that influence strain (Daniels, 2006). Moreover, enacted characteristics capture the behavioral dynamics of job demands (e.g., How much actual physical activity did a firefighter exhibit during a specific performance episode?). By contrast, perceived job characteristics refer to stable or typical experiences (e.g., How physically demanding does a firefighter perceive his work in general?). In order to understand the dynamic processes involved in determining how job resources are utilized to manage job demands, calls have been made to focus on enacted job characteristics (Duong, Tuckey, Hayward, & Boyd, 2015). Consequently, we conceptualize physical demands as the amount of objective physical activity during a specific performance episode.

Recently, research has started investigating intraindividual differences in job demands and resources across time—typically in the context of diary studies (see Schaufeli & Taris, 2014, for an overview). These studies suggest that day-level variations in work outcomes can be explained by day-level variations in job demands. However, empirical investigations of the short-term dynamics between enacted job demands and job resources are still scarce (Duong et al., 2015). This neglect is particularly problematic in high-reliability occupations where success and failure is determined at the level of a single performance episode. Each firefighting mission is a discrete event with a separate outcome; failure in one mission can seldom be compensated by success in another. Thus, we need to understand not only how job demands and resources affect employees in general but also how they shape work outcomes in a specific performance episode. To address this issue, we investigate employees' immediate reactions to being exposed to a series of physically demanding tasks and, in particular, the role of trust in shaping these reactions.

As we are interested in employees' immediate reactions, we focus on the most proximal work outcome: perceived strain. Similar to perceived exertion, perceived strain captures the subjective situational appraisal of the task and incorporates several signals, perceptions, and experiences (cf. Borg, 1982). Just as other stress- and strain-related subjective appraisals, perceived strain results from the interplay of external task demands and employees' resources; it is this subjective appraisal that antecedes long-term strain outcomes (Csikszentmihalyi & Csikszentmihalyi, 1992; Lazarus & Folkman, 1984). For this reason, perceived strain has been used frequently as an outcome in research on physically demanding tasks in high-reliability occupations such as firefighting (Lusa, Punakallio, Luukkonen, & Louhevaara, 2006) and police work (Johnsen et al., 2017).

[1]We acknowledge that some researchers regard cognition- and affective-based trust as separate constructs (e.g., McAllister, 1995). Following Colquitt et al. (2011), we conceptualize trust as a unitary construct that has both affective and cognitive antecedents.

Studies on the effects of physical demands using self-reports of typical work experiences suggest negative effects on variables associated with psychological strain (Li, Jiang, Yao, & Li, 2013; van Den Tooren & de Jonge, 2008). Accordingly, we propose an effect of objective physical activity on perceived psychological strain—employees' subjective appraisal of the strain-related psychological aspects of a specific task. In line with research suggesting that day-level variations in work outcomes can be explained by day-level variations in job demands, we propose that variations in perceived psychological strain across performance episodes can be explained by variations in physical activity across performance episodes, and thus, we expect to replicate previous findings regarding the effects of physical demands on the performance episode level.

> **Hypothesis 1.** *Employees' objective physical activity during a performance episode is positively related to their perceived strain after the performance episode.*

2.4 | The buffering effect of job resources

Importantly, we propose that trust in teammates represents a resource capable of buffering the hypothesized negative effects of objective physical activity. Trust can either have a main effect or alternatively moderate the effect of another variable on work outcomes (Dirks & Ferrin, 2001). Focusing on the moderating effect, we argue that high trust in teammates buffers the effect of job demands by affecting how individual members perceive these demands.

Two different literatures support our argument that trust affects the perception of job demands. First, in the occupational stress literature, it is generally assumed that a variable can act as a buffer when it makes job demands more understandable to an employee; that is, the employee appreciates the "reasons for the presence of a stressor and their exposure to it" (Kahn & Byosiere, 1992, p. 622). Building on this notion, the JD-R proposes that job resources can buffer the negative impact of job demands by affecting the perception of these demands (Bakker & Demerouti, 2007). For example, appreciation from a supervisor can make a highly demanding task more bearable for an individual. To this regard, Crawford, LePine, and Rich (2010) have supplemented the JD-R by including appraisal processes of job demands. Appraisal processes play a central role in many stress models by shaping the relationship between objective demands and their subjective perception (e.g., Cohen & Wills, 1985; Lazarus & Folkman, 1984). In this context, trust in management has been proposed to affect the extent to which employees perceive job demands as negative (Mishra & Spreitzer, 1998). In support of this notion, employees with high trust in their management were found to be less negatively affected by work overload, presumably because trust makes the presence of high workload more understandable, thereby mitigating the negative evaluations that usually result from this job demand (Harvey, Kelloway, & Duncan-Leiper, 2003). Similarly, a recent study based on propositions of COR theory found that trust in management represents a resource capable of buffering the negative effects of job insecurity (Jiang & Probst, 2016).

We argue that trust in teammates can have similar effects in team settings. In an individual work setting, management actions are a major factor responsible for employees' job demands because they distribute tasks, assign responsibilities, and set deadlines. As a result, trust in management influences how employees evaluate a highly demanding situation. In a team setting, however, employees' job demands depend on the other team members: If their teammates loaf or are incompetent, employees' workload is likely to increase—in particular if their tasks are highly interdependent as is the case in high-reliability occupations. Consequently, other team members' actions constitute a main driver of workload, and trust in teammates is likely to influence how employees evaluate a highly demanding situation.

Second, in the literature on trust in organizations, Dirks and Ferrin (2001) argue that trust acts as a moderator by affecting a person's interpretation of another's past actions or events related to those actions. Specifically, under high levels of trust, employees are likely to interpret their coworkers' actions more favorably than under low levels of trust (Dirks & Ferrin, 2001). Accordingly, if employees trust their teammates, they should be more likely to interpret their teammates' actions favorably because they are confident that their teammates will perform actions that are beneficial for the team (e.g., Breuer et al., 2016). By contrast, if employees do not trust their teammates, they are not sure that their teammates' actions are beneficial for the team because they are dubious about their teammates' ability or benevolence. As a result, their interpretation of their teammates' actions should be less favorable. As an example, consider a medical team in which the anesthetist insists on double-checking instruments and drugs before starting surgery. If the surgeon trusts her teammate, she will likely interpret her colleague's action as an expression of genuine concern and will not mind double-checking. However, if she does not trust her, she might interpret it as an attempt to undermine her authority because she has doubts about her teammate's benevolence. Consequently, she will be likely to regard the double-checking as unnecessary. In support of our argumentation, trust has been found to affect individuals' perceptions of others' behavior in the team context (Brown, Crossley, & Robinson, 2014).

Building on these two streams of literature, we propose that trust in teammates affects how team members perceive a highly demanding task. Trusting one's teammates implies positive expectations about their actions and motivation (Breuer et al., 2016; De Jong & Elfring, 2010). These expectations are rooted in the belief in one's teammates' ability and benevolence. Thus, if employees trust their teammates, they are confident that (a) their teammates are competent and (b) that their teammates would not intentionally do anything to compromise them. By contrast, if employees do not trust their teammates, they are not sure that their teammates are competent and well meaning. These differences can affect employees' interpretation of physical demands. Because high-trusting employees are confident that their teammates do their part of the job, they will be more likely to interpret high physical demands as a consequence of the task. Low-trusting employees, however, will likely interpret high physical demands as a consequence of their teammates' lack of motivation or competence. As a result, they will regard these demands as unnecessary or unfair burden ("I would not have to do this if my teammates had done their part"). Tasks that are considered unnecessary have been found to be more straining (Semmer et al., 2015). Taken together, we propose that, because of differences in the interpretation of job demands,

employees who do not trust their teammates will perceive a high amount of physical activity as more straining, as compared with employees who do trust their teammates.

> **Hypothesis 2.** *Trust in teammates moderates the relationship between physical activity and perceived strain: Physical activity is more positively related to perceived strain under conditions of low trust as compared with conditions of high trust.*

2.5 | The effect of recurring performance episodes

Turning to the central contribution of the current study, it has been recently suggested that the effectiveness of job resources may vary across task situations (van Woerkom, Bakker, & Nishii, 2016). Interesting questions arise when the hypothesized buffering effect of trust is seen in light of task characteristics that might cause these variations—namely, recurring performance episodes.

We argue that due to the job characteristics described above, employees are regularly forced to operate at their limits. As a consequence, completing a single performance episode is likely to result in a significant depletion of physical resources (Bakker & Demerouti, 2007; Demerouti et al., 2001). Thus, they will start the next performance episode with fewer resources at their disposal. Combining the JD-R with Hobfoll's (1989, 2002) COR theory, it has been assumed that in such a situation of resource loss, other resources can be particularly salient (Bakker, Hakanen, Demerouti, & Xanthopoulou, 2007).

Building on this notion, we propose that trust in teammates should become a more salient resource in later performance episodes. Specifically, we argue that firefighters are motivated to offset the significant loss of physical resources that occurred during the first performance episode. Because direct replacement of physical resources is not possible in this situation, they should—according to COR theory—strive to seek replacement through indirect means (Hobfoll, 1989). Consequently, they should be highly motivated to look for alternative resources that help them in dealing with the demands of the next performance episode. As trust in teammates represents an important resource in the team context, it should become more salient.

In Hypothesis 2, we have argued that employees who do not trust their teammates interpret high demands less favorably, as compared with employees who do trust their teammates. Building on this argumentation, we propose that the increased salience of trust affects the perceptions underlying the associated buffering effect. High-trusting employees are inclined to interpret high demands as consequence of the task situation. Making trust more salient should not change their perception of job demands. Thus, the effect of physical demands on perceived strain should remain stable. By contrast, low-trusting employees are inclined to interpret their high demands as a consequence of their teammates' lack of motivation or competence. Making trust more salient should lead to an even less favorable interpretation, resulting in a stronger effect of physical demands on perceived strain. Thus, when trust becomes more salient in a second performance episode, differences in the interpretation of job demands between high- and low-trusting employees should increase: Employees without an additional resource at their disposal—trust in teammates—should perceive high physical demands as even more straining than employees with this resource. This is again in line with COR theory, which assumes that persons with more resources are "less negatively affected by resource drain or loss that occurs in the face of stressful conditions" (Hobfoll, 2002, p. 318). Accordingly, we formulate the following hypothesis.

> **Hypothesis 3.** *The buffering effect of trust in teammates on the relationship between physical activity and perceived strain becomes more pronounced during later performance episodes.*

3 | METHOD

3.1 | Participants

We conducted the study in cooperation with the fire department of a Western European city. Participants were 45 professional firefighters. All participants were male. We used age cohorts: 8.9% of the participants were between 24 and 29 years, 17.8% were between 30 and 35 years, 13.3% were between 36 and 41 years, 22.2% were between 42 and 47 years, 24.4% were between 48 and 53 years, and 13.3% were over 53 years old.

3.2 | Context

Data collection took place in a burn building—a high-fidelity training facility for naturalistic simulations of fire incidents. The burn building is used for training purposes on a regular basis, so participants were familiar with the facility. Training scenarios in the burn building are highly realistic, as they resemble real-life incidents in many aspects. They involve actual fires, extreme heat, high humidity, restricted visibility, and thick smoke. Moreover, firefighters use their regular equipment including breathing apparatus with pressurized air, protection suit, hose, turntable ladder, and fire engine.

Participants completed a customized training scenario, which we designed in cooperation with two training instructors specifically for the purpose of the current study. The scenario involved a fire on the third floor of an apartment building (i.e., the burn building). The participants arrived in two fire engines. The entrance to the building was blocked, so they had to enter via the roof. This required the use of the turntable ladder. In the building, visibility was limited due to smoke. Thus, it was difficult to locate the source of the fire. The scenario involved an unconscious person (i.e., dummy) in the building. Participants had to rescue the person first and then to extinguish the fire in order to complete their mission.

3.3 | Procedure

On each day of the data collection, participants arrived at the training site with their regular equipment and were informed about the study. They completed a consent form and the trust measure. Each participant was handed a smartphone. Importantly, all participants were required to carry their smartphone in exactly the same manner (i.e., left breast pocket) to ensure both accuracy and comparability of measurements. Next, a training instructor gave a technical briefing about the

training scenario, and the first trial was conducted. After the trial, participants had a break of about 15 min during which they completed the perceived strain measure. After their third trial, participants were thanked and debriefed. The trials had an average duration of 12.45 min (*SD* = 1.90).

Participants completed the trials in seven teams of seven to nine firefighters. The teams consisted of firefighters that worked together on a regular basis. Each team was scheduled to repeat the same, standardized scenario three times. Repeating the same scenario rather than doing different scenarios ensured that the overall physical demands of the task remained constant across performance episodes. To impede learning effects, participants switched their roles after each trial. Team composition remained constant across the three trials, and each team completed all their trials on the same day. Due to organizational issues, three trials had to be canceled. However, the same team of firefighters conducted the scenario at least twice together. The analyses for the current study were based on a sample of 114 observations (i.e., individual firefighters completing the scenario) from the 45 participants, whereby each individual firefighter provided either two or three observations.

3.4 | Measures

The study represents an observation study with three measurement points: Trust was measured once before the first trial, physical activity was measured during each trial, and the outcome variable perceived strain was measured immediately after each trial.

3.4.1 | Trust in teammates

We used the trust scale by Colquitt et al. (2011), who developed this scale specifically for the context of firefighting. The scale assesses the extent to which firefighters trust their colleagues regarding a number of standard tasks. Sample items include "I trust my colleagues to discuss fire tactics before beginning operation" and "I trust my colleagues to study maps of the local territory." Participants indicated their agreement on a 5-point scale (from 1 = *strongly disagree* to 5 = *strongly agree*). The scale was translated in cooperation with a graduate student who was an experienced firefighter. A first version of the scale was pretested in a sample of firefighters. One item was deleted because it did not apply to the current setting. The final German version used in this study included eight items and had a high reliability (α = .87).

3.4.2 | Perceived strain

As the setting prevented us from using lengthy questionnaires after each trial, we employed a short instrument to assess perceived strain. We used the National Aeronautics and Space Administration Task Load Index (TLX)—a multidimensional scale designed to obtain subjective perceptions of work outcomes (Hart & Staveland, 1988). Participants indicate their ratings regarding six items on a scale between 0 and 20. The scale has been validated in many settings (Hart, 2006), and TLX items have been used to operationalize perceived strain in other high-reliability occupations (Collet, Averty, & Dittmar, 2009; Weigl, Müller, Sevdalis, & Angerer, 2013). For the purpose of the current study, we focused on three relevant items of the TLX that tie in closely with perceptions of psychological strain: perceived mental task demands, perceived temporal task demands, and perceived effort.[2] A sample item is, "How hard did you have to work to complete the task?" The three items were aggregated to obtain a single indicator of perceived strain. The scale had good reliability (α = .76).

3.4.3 | Physical activity

We used individual motion activity during the scenario to operationalize physical activity. Motion activity was measured with smartphones (see Appendix A for a detailed description). A number of studies using similar approaches have demonstrated that smartphones' built-in accelerometers can accurately recognize individuals' physical activities and thus reveal, for example, whether people are standing still, walking, running, or climbing a ladder (Kwapisz, Weiss, & Moore, 2011; Miller, 2012; Piwek & Joinson, 2017). Physical activity values can potentially range from 0 (i.e., *person did not move at all during the scenario*) to +1 (i.e., *person was always physically active*).

3.5 | Analysis strategy

We analyzed the data with the packages *lme4* (Bates, Maechler, Bolker, & Walker, 2015) and *multilevel* (Bliese, 2013b) of the statistical software R (R Core Team, 2015). The data had a hierarchical structure spanning three levels: Repeated observations of the dependent variable were nested in individuals, which were again nested in teams. Moreover, trust was measured at the individual level. In other words, whereas the participants completed the trust measure only once (i.e., on the individual level), we elicited the independent variable physical activity and the outcome measure perceived strain for each performance episode (i.e., on the measurement level). Following the recommendations by Bliese (2002, 2013a), we first determined whether the data warranted taking their nested nature into account by investigating the intraclass correlations (ICCs) of the outcome variable perceived strain on the individual and team levels. Nonindependence was present only on the individual level, ICC(1) = .56, $F(44, 69) = 4.23$, $p < .001$, ICC(2) = .76, but not on the team level, ICC(1) = −.02, $F(6, 107) = 0.63$, $p = .70$, meaning that perceived strain values depended on the person from whom they were obtained, but not on the team. This pattern seems reasonable given that all teams conducted the same standardized scenario, and thus, outcomes should be rather affected by individual differences than by team-level factors.

To account for the hierarchical nature of our data, we employed mixed models (i.e., multilevel models) in the analysis (Gelman & Hill, 2006). Mixed models include all variables on their respective levels, thereby providing the most accurate statistical representation of a repeated-measures scenario such as the present one (Galecki & Burzykowski, 2013). Specifically, we employed a two-level random intercept model and tested whether the addition of random slopes explained significant further variance in the dependent variable (Bliese & Ployhart, 2002). To account for effects of performance

[2]Using all six items of the TLX produced a similar pattern of results.

episodes, we used the trial number as an additional predictor. As we were interested in differences between performance episodes, performance episode was operationalized as a categorical predictor. Using difference coding, we calculated two contrasts: one comparing the second trial with the first and one comparing the third trial with the second.

4 | RESULTS

Table 1 provides an overview of descriptive statistics of the study variables. As we measured perceived strain and physical activity for each performance episode, we report them separately for each episode (i.e., trial). Trust was measured only once before the start of the data collection. Notably, the average level of physical activity remained constant across the three performance episodes.

TABLE 1 Means and standard deviations (*SD*) for all study variables

Variables	*N*	Mean	*SD*
Trial 1			
Physical activity	45	0.67	0.22
Perceived strain	45	8.45	2.63
Trial 2			
Physical activity	40	0.68	0.21
Perceived strain	40	8.43	3.61
Trial 3			
Physical activity	29	0.68	0.20
Perceived strain	29	8.93	4.12
Overall			
Physical activity	114	0.67	0.21
Perceived strain	114	8.56	3.38
Trust	45	3.71	0.68

We constructed the initial model as a random intercept model with varying intercepts for perceived strain. An alternative random intercept and slope model did not fit the data better than the random intercept model, $\Delta\chi^2(2) = 0.18$, $p = .92$. We thus employed the random intercept model for hypothesis testing (Model 1). Following conventions for testing interactions, we first added mean-centered trust in teammates as a Level 2 predictor in a second step (Model 2). Then, we added the two-way cross-level interactions between trust and physical activity, and trust and performance episode, to test Hypothesis 2 (Model 3). Finally, we added the three-way interaction between trust, physical activity, and performance episode (Model 4). Please note that in the absence of a random slope as in these models, cross-level interactions pertain only to the fixed effects and are interpreted analogous to interaction effects in ordinary least squares regression; that is, the effect of one fixed effect (i.e., physical activity) on the dependent variable (i.e., perceived strain) depends on another variable (i.e., trust in teammates). The results of the analysis are summarized in Table 2. In addition to *p* values, we also report bootstrap confidence intervals (based on 5,000 resamples) for the parameter estimates. Finally, we estimate the models' total variance explanation with the pseudo-R^2 for generalized mixed models (Nakagawa & Schielzeth, 2013).

4.1 | Hypothesis testing

In support of Hypothesis 1, Model 1 revealed a significant relationship between physical activity, operationalized through the smartphone-based measurement of motion activity, and perceived strain in Model 1 ($\gamma = 6.62$, $p < .001$; 95% CI [4.16, 9.00]). The more physical activity

TABLE 2 Results of the mixed models predicting firefighters' perceived strain

Predictors	Model 1	Model 2	Model 3	Model 4
Level 1				
Physical activity (PA)	6.62*** (1.21)	6.67*** (1.21)	6.99*** (1.19)	7.36*** (1.16)
Trials 2 vs. 1	−0.04 (0.41)	−0.04 (0.41)	−0.09 (0.38)	−0.38 (0.40)
Trials 3 vs. 2	0.37 (0.48)	0.38 (0.48)	0.51 (0.46)	0.75 (0.46)
PA × Trials 2 vs. 1	4.85* (2.24)	4.90* (2.24)	4.69* (2.24)	6.25** (2.30)
PA × Trials 3 vs. 2	−1.43 (2.56)	−1.51 (2.56)	−1.83 (2.48)	−3.39 (2.50)
Level 2				
Trust		0.32 (0.62)	0.46 (0.65)	0.53 (0.65)
Cross-level interactions				
PA × Trust			−4.07* (1.90)	−3.90* (1.95)
Trials 2 vs. 1 × Trust			−0.46 (0.60)	0.05 (0.63)
Trials 3 vs. 2 × Trust			0.25 (0.66)	−0.20 (0.68)
PA × Trials 2 vs. 1 × Trust				−6.71* (3.32)
PA × Trials 3 vs. 2 × Trust				6.99 (4.54)
Log-likelihood	−270.86	−270.72	−268.42	−266.20
~R^2				.18

[a]. $N = 114$ (Level 1), $N = 45$ (Level 2). Values in parentheses are standard errors.

*$p < .05$. **$p < .01$. ***$p < .001$ (two-tailed tests).

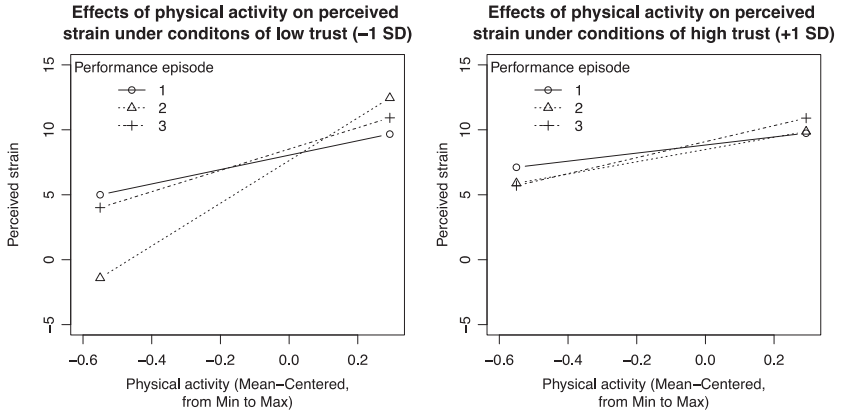

FIGURE 1 Three-way cross-level interaction between physical activity, performance episode, and trust. Please note that in three-way interaction plots, the scale of the *y* axis may be incorrect, but the shape of the interaction is correct (Dawson & Richter, 2006)

firefighters showed during task execution, the more straining they perceived the task. The effect of physical activity remained significant in all models. In Model 2, trust did not have a significant main effect on perceived strain. As a side note, our analysis also revealed a significant interaction between physical activity and performance episode.

Our second hypothesis proposed that trust in teammates would moderate the relationship between physical activity and perceived strain. Model 3 revealed a significant effect of the interaction between physical activity and trust on perceived strain ($\gamma = -4.07$, $p = .034$; 95% CI [−7.76, −0.30]), which remained significant in Model 4. This finding provides support for our second hypothesis. The two-way interaction, however, was expected to be qualified by the three-way interaction predicted by Hypothesis 3.

Regarding our third hypothesis, the final model revealed a significant three-way interaction between physical activity, trust, and the difference between the second and first episodes ($\gamma = -6.71$, $p = .047$; 95% CI [−13.27, −0.31]). By contrast, the three-way interaction between physical activity, trust, and the difference between the third and second episodes was not significant ($\gamma = 6.99$, $p = .13$; 95% CI [−1.97, 15.88]). This finding indicates a difference in the buffering effect of trust between the first and second episodes, but not between the second and third. The plot indicates an overall interaction between physical activity and trust (Figure 1): The relationship between physical activity and perceived strain appears to be weaker under conditions of high trust (+1 *SD*) as compared with low trust (−1 *SD*), which is in line with Hypothesis 2. Moreover, under conditions of low trust, the relationship between physical activity and perceived strain seems to vary across performance episodes, whereas under conditions of high trust, this relationship is similar for all performance episodes. Specifically, the effect of physical activity on perceived strain becomes more pronounced in the second performance episode, but only for firefighters who do not trust their teammates. This pattern is in line with our prediction. At first glance, the strength of this effect appears to become weaker again in the third episode such that it lies between the first and the second. Although the difference between Episodes 2 and 3 is not significant, this could indicate a nonlinear effect of performance episode on the buffering effect of trust. We decided to further explore this possibility in an additional analysis.[3]

[3]We thank the editor for this suggestion.

4.2 | Additional analysis

To further clarify the three-way interaction, we constructed an alternative model using polynomial contrasts: one contrast for the linear effect of performance episode (i.e., 1 < 2 < 3) and one contrast for the nonlinear (i.e., quadratic) effect. This model is identical to Model 4 in most respects including log-likelihood, total variance explanation, and predicted values. The only differences concern estimates for parameters involving performance episode.

The analysis reveals a significant two-way interaction between performance episode and physical activity for the nonlinear effect ($\gamma = -3.94$, $p = .025$; 95% CI [−7.28, −0.50]), but not for the linear effect ($\gamma = 2.02$, $p = .219$; 95% CI [−1.15, 5.25]). The same pattern emerged for the three-way interaction: We found a significant nonlinear effect ($\gamma = 5.60$, $p = .045$; 95% CI [0.43, 10.98]), but no linear effect ($\gamma = 0.20$, $p = .948$; 95% CI [−5.65, 6.14]). This pattern is in line with the previous model and the plot. For low-trusting firefighters (−1 *SD*), the effect of physical activity on perceived strain becomes significantly stronger in the second performance. The effect becomes weaker again in the third performance episode such that it lies between the first and the second, but this change is not significant. By contrast, for high-trusting firefighters (+1 *SD*), the effect of physical activity on perceived strain appears to be stable across performance episodes.

5 | DISCUSSION

The current study adds a temporal perspective to research on the effectiveness of job resources by investigating variations in the buffering effect of trust in teammates. In line with our hypotheses, multilevel modeling revealed a significant positive effect of objective physical activity on the subjective perception of strain. Moreover, our findings suggest that trust in teammates buffered the effect of physical activity on perceived strain. Importantly, and as expected, this buffering effect varied across performance episodes: It was more pronounced in the second performance episode as compared with the first performance episode. Specifically, firefighters with low trust in their teammates were more negatively affected by high physical activity in the second performance, whereas this was not the case for firefighters with high trust in their teammates. This finding suggests that the effectiveness

of job resources in buffering the influence of job demands can vary over time.

Overall, the current study offers several contributions in terms of theory, research methodology, and practice. First, one major contribution of the current research lies in presenting theory and empirical evidence regarding temporal variations in the buffering effect of job resources. We demonstrate that the buffering effect of trust on the relationship between physical activity and perceived strain becomes more pronounced during a second performance episode. Although the stability of the buffering effects of job resources has been recently called into question (van Woerkom et al., 2016), temporal variations in buffering effects have received limited attention in the literature. Empirical evidence suggests that job resources are not consistently effective in buffering the negative effects of physical demands (Bakker et al., 2005). Moreover, diary studies investigating dynamic elements of the JD-R suggest that job demands and resources fluctuate over the course of a week (e.g., Schaufeli & Taris, 2014). The current study extends these lines of research by demonstrating that the buffering effects of job resources can vary across performance episodes. We consider this an important addition to existing research that contributes to an improved understanding of how temporal factors shape the interplay of job demands and resources and, thus, inspires both empirical research and further theories. Of course, we do not provide an exhaustive analysis of all factors that influence the buffering effect of job resources. Our findings, however, do imply that research on job resources should consider temporal factors.

It should be noted that our analysis did only reveal a difference in the buffering effect of trust between the first and second performance episodes, but not between the second and the third. Moreover, our additional analyses suggest that the effect of performance episode is nonlinear. Retrospectively, this resonates with our argumentation. We have argued that a single performance episode results in a significant depletion of physical resources, which increases employees' motivation to look for alternative resources and thus makes trust in teammates more salient (Bakker et al., 2007; Hobfoll, 1989). It is conceivable that the relationship between the loss of one resource and the increased salience of alternative resources is not linear. Instead, this relationship might be better described in terms of a threshold model, such that resource loss in excess of a critical threshold level does not lead to a further increase in salience of alternative resources. Following this interpretation, the depletion of physical resources during the first performance episode might have been enough to reach the threshold and thus render trust a more salient resource. After that, however, additional loss of physical resources did not result in a further increase in the salience of trust, and the buffering effect of trust did therefore not differ between the second and third performance episodes. Alternatively, this finding might be accounted for by the setting of the current study. The participating firefighters knew that the third performance episode would be the final one and that they could replenish their physical resources after that. This knowledge might have decreased their motivation to focus on alternative resources and thus affected the salience of trust in teammates and the respective buffering effect. In sum, findings from our analyses suggest that the detected variations in the buffering effect of trust follow a nonlinear rather than linear pattern. However, more research is needed to fully understand how the effectiveness of job resources can vary over time.

A second contribution of the current study lies in demonstrating how recurring performance episodes affect work outcomes in the team context: Our findings indicate that they are a source of variation in the functionality of trust in teammates. On a conceptual level, researchers have long emphasized the importance of time for team processes and performance (e.g., Gersick, 1988; McGrath, 1991). Yet researchers have hardly addressed variations in team processes and emergent states across performance episodes. In this regard, a recent review of literature stated a notable absence of empirical studies that analyze the dynamic effects predicted by temporal models of teamwork (Mathieu, Tannenbaum, Donsbach, & Alliger, 2014). Against this background, the current study provides initial evidence for how the functionality of team-related variables (i.e., trust in teammates) varies across performance episodes. Although we focused on individual members' work outcomes, our work contributes to an improved understanding of temporal dynamics in the team setting.

Third, our work has implications in terms of research methods. We illustrate how smartphones can be used to quantify job demands in an applied setting. The reliance on subjective, self-report data in the study of job demands has been criticized for a long time (Spector et al., 1988). Self-report data are prone to bias (e.g., Podsakoff, MacKenzie, & Podsakoff, 2012; Spector, 1994), which poses a threat to objectivity. The current study addresses this issue by using smartphone-based motion activity data as an objective measure of physical activity, which also reduces the risk of common method bias. Social sensing research has shown that wearable sensors can be used to measure human behavior in organizational settings (Pentland, 2010; Schmid Mast, Gatica-Perez, Frauendorfer, Nguyen, & Choudhury, 2015). Building on this research, we were able to obtain objective behavioral data in a setting where such data are hardly accessible through conventional methods.

The current study also illustrates how smartphones could promote the study of temporal phenomena in organizations. We have already discussed the lack of empirical work on team dynamics. In this context, innovative measurement techniques such as smartphones have been recently proposed as a means for advancing research (Kozlowski, 2015). Apart from team research, applied psychology in general has been criticized for neglecting the temporal aspects of behavior (Roe, 2008). Smartphones can help address these issues as they are capable of generating behavioral data at a much higher temporal resolution than conventional methods (Miller, 2012). Of course, for the purpose of the current research, we used a relatively simple approach to conceptualize physical activity that did not make full use of smartphones' functionality. For example, we did not use location data (e.g., Global Positioning System, which did not work inside the building) that would have potentially allowed us to determine covered distances more precisely. Future research should further capitalize on the different sensors to obtain even more precise measures of physical activity. Besides, because social sensing is still a new field, there are no standards regarding the use of smartphone-based measures. As a result, studies require substantial preparation work in developing and testing their measures, which is complicated by the large amount of data produced by this approach (Miller, 2012). On a general level, however, we think that the present study

illustrates the usefulness of smartphone-based metrics and provides an example of how smartphones can be used to collect behavioral data in applied settings. In doing so, we hope to further promote the use of behavioral data in organizational psychology.

5.1 | Practical implications

Another goal of our study was to identify psychological resources that help members of high-reliability teams in dealing with the effects of highly demanding working conditions. Our findings indicate that trust in teammates buffers the negative effects of physical activity. Notably, trust seems to play an important role even in a predominantly physical work setting. The practical implications of this finding become apparent if one considers that physical demands constitute an integral part of many high-reliability occupations. Physical demands, in turn, are associated with high levels of strain. Because physical demands often cannot be reduced, it is important to identify job resources that help employees deal with these demands. Against this background, the current study makes a practical contribution by demonstrating that trust in teammates buffers some of their negative effects.

Importantly, as opposed to the physical activity required to complete a task, trust can be targeted by interventions. Two basic strategies are available to improve trust. First, team composition could be modified to maximize trust among team members. This strategy, however, is unfeasible in applied settings where team composition is determined by shift schedules and staffing requirement. The second strategy—interventions to promote trust—seems more promising. Research has provided a variety of specific interventions that can be used to promote trust in interpersonal relations (Abrams, Cross, Lesser, & Levin, 2003). For example, leaders and command personnel should give employees the opportunity to voice concerns and present suggestions for improvements. Everyone's contribution should be valued in order to create a climate of open and collaborative communication, which in turn should increase interpersonal trust (Abrams et al., 2003). In general, we think trainers and human resource managers should pay increased attention to psychological resources. Some high-reliability occupations such as health care have long acknowledged the role of human factors (Baker et al., 2006), and training programs addressing these issues have been designed (e.g., Wilson et al., 2005). In view of our findings, similar training approaches to promote trust could be beneficial for employees in high-reliability occupations.

5.2 | Limitations and future research

As this is the first study addressing temporal variations in the buffering effect of job resources, there are several limitations and a number of avenues for future research. First, our study focused on investigating variations in the buffering effect of a single job resource (i.e., trust in teammates). As a consequence, we cannot be sure to what extent our findings can be generalized, that is, whether similar variations can be found in the buffering effect of other job resources. As the mechanism underlying the observed variations in the buffering effect of trust—namely, that other resources can become particularly salient in the context of resource loss—is well established in the literature (e.g., Bakker et al., 2007), we are confident that this finding is not specific to the current setting.

Similarly, our study can be criticized for focusing on the physical aspects of firefighting. While we believe that firefighting is characterized by its physical demands, we acknowledge there are instances in which firefighting is also mentally demanding (e.g., decision making under uncertainty). Moreover, in other high-reliability occupations such as emergency medicine, job demands are not primarily physical but also involve emotionally demanding interactions with patients. Thus, future research should investigate the buffering effect of trust in the context of multiple types of demands including physical, mental, and socio-emotional demands. For example, it would be interesting to investigate the buffering effect of trust in emergency care. In general, future research should try to validate our findings using different combinations of job demands and resources.

Another limitation of the current study concerns our focus on the short-term dynamics of the strain processes. We only investigated variations in the buffering effects of trust with regard to the most proximal variable in that process—perceived strain. Although the relationship between a job demand and the perception of said demand constitutes a vital part of the strain process, future studies should investigate other strain outcomes.

Our study featured a limited number of similar teams (i.e., all members were highly trained professionals who worked together regularly), which could explain the absence of a group-level effect on perceived strain. Thus, future research should investigate the buffering of trust in teammates in a larger, more diverse set of teams. A potential study could compare multiple teams performing different tasks.

We also measured trust only once at the beginning of the study. Although we think it is rather unlikely that trust in teammates could have changed over the course of the three scenarios, we cannot exclude this possibility. Consequently, future research should measure trust in teammates before each performance episode to account for potential variations.

Another avenue for future research would be to focus on negative emotions, which constitute an important part of the strain process (Spector, 2002). It would be worthwhile to investigate the role of negative emotions in the context of recurring performance episodes. For example, if emergency physicians are frustrated from dealing with a first incident, they might sustain this negative emotion, which could impede their ability to successfully handle subsequent emergencies. A potential question would be whether trust in teammates could prevent "carryover effects" of negative emotions from one performance episode to the next.

6 | CONCLUSIONS

The current study offers three main contributions. First, we add a temporal perspective to research on the effectiveness of job resources by investigating variations in the buffering effect of trust in teammates. Our findings indicate that this buffering effect varies across performance episodes: Trust in teammates was more effective in buffering the negative effects of physical activity during a second performance episode. Second, we show that trust in teammates is an important psychological resource for employees in high-risk occupations that is capable of buffering some of the negative effects of objective physical job

demands. Third, the current study illustrates how smartphones can be used for obtaining objective behavioral data in an applied setting.

ACKNOWLEDGEMENT

The authors thank the members of the City of Zurich Fire Department, in particular Jan Bauke and Pascal Eichmann, for their support throughout the study. We also thank Anna-Lena Köng, Laura Fischer, and Nadja Ott for their help with the data collection; Bert Arnrich and Paul Langraf for their help in preparing this study; and Maike Debus and Joachim Hüffmeier for their helpful comments on an earlier version of this manuscript.

ORCID

Michael J. Burtscher http://orcid.org/0000-0002-6686-3859

REFERENCES

Abrams, L. C., Cross, R., Lesser, E., & Levin, D. Z. (2003). Nurturing interpersonal trust in knowledge-sharing networks. *The Academy of Management Executive*, *17*, 64–77. https://doi.org/10.5465/ame.2003.11851845

Baker, D. P., Day, R., & Salas, E. (2006). Teamwork as an essential component of high-reliability organizations. *Health Services Research*, *41*, 1576–1598. https://doi.org/10.1111/j.1475-6773.2006.00566.x

Bakker, A. B., & Demerouti, E. (2007). The job demands–resources model: State of the art. *Journal of Managerial Psychology*, *22*, 309–328. https://doi.org/10.1108/02683940710733115

Bakker, A. B., Demerouti, E., & Euwema, M. C. (2005). Job resources buffer the impact of job demands on burnout. *Journal of Occupational Health Psychology*, *10*, 170–180. https://doi.org/10.1037/1076-8998.10.2.170

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*, 274–284. https://doi.org/10.1037/0022-0663.99.2.274

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Baumann, M. R., Gohm, C. L., & Bonner, B. L. (2011). Phased training for high-reliability occupations live-fire exercises for civilian firefighters. *Human Factors*, *53*, 548–557. https://doi.org/10.1177/0018720811418224

Bliese, P. D. (2002). Multilevel random coefficient modeling in organizational research: Examples using SAS and S-PLUS. In F. Drasgow, & N. Schmitt (Eds.), *Modeling in organizational research: Measuring and analyzing behavior in organizations* (pp. 401–445). San Francisco, CA: Jossey-Bass.

Bliese, P. D. (2013a). Multilevel modeling in R (2.5). Retrieved 12 March, 2014, from http://cran.r-project.org/doc/contrib/Bliese_Multilevel.pdf

Bliese, P. D. (2013b). Multilevel: Multilevel functions (version R package version 2.5). Retrieved from http://cran.r-project.org/package=multileve

Bliese, P. D., & Ployhart, R. E. (2002). Growth modeling using random coefficient models: Model building, testing, and illustrations. *Organizational Research Methods*, *5*, 362–387. https://doi.org/10.1177/109442802237116

Borg, G. A. (1982). Psychophysical bases of perceived exertion. *Medicine and Science in Sports and Exercise*, *14*, 377-381.

Bos, J., Mol, E., Visser, B., & Frings-Dresen, M. H. (2004). The physical demands upon (Dutch) fire-fighters in relation to the maximum acceptable energetic workload. *Ergonomics*, *47*, 446–460. https://doi.org/10.1080/00140130310001643283

Breuer, C., Hüffmeier, J., & Hertel, G. (2016). Does trust matter more in virtual teams? A meta-analysis of trust and team effectiveness considering virtuality and documentation as moderators. *Journal of Applied Psychology*, *101*, 1151–1177. https://doi.org/10.1037/apl0000113

Brown, G., Crossley, C., & Robinson, S. L. (2014). Psychological ownership, territorial behavior, and being perceived as a team contributor: The critical role of trust in the work environment. *Personnel Psychology*, *67*, 463–485. https://doi.org/10.1111/peps.12048

Bulling, A., Blanke, U., & Schiele, B. (2014). A tutorial on human activity recognition using body-worn inertial sensors. *ACM Computing Surveys*, *46*, 1–33. https://doi.org/10.1145/2499621

Butte, N. F., Ekelund, U., & Westerterp, K. R. (2012). Assessing physical activity using wearable monitors: Measures of physical activity. *Medicine & Science in Sports & Exercise*, *44*, S5–S12. https://doi.org/10.1249/MSS.0b013e3182399c0e

Cohen, S., & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin*, *98*, 310–357. https://doi.org/10.1037/0033-2909.98.2.310

Collet, C., Averty, P., & Dittmar, A. (2009). Autonomic nervous system and subjective ratings of strain in air-traffic control. *Applied Ergonomics*, *40*, 23–32. https://doi.org/10.1016/j.apergo.2008.01.019

Colquitt, J. A., LePine, J. A., Zapata, C. P., & Wild, R. E. (2011). Trust in typical and high-reliability contexts: Building and reacting to trust among firefighters. *Academy of Management Journal*, *54*, 999–1015. https://doi.org/10.5465/amj.2006.0241

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*, 834–848. https://doi.org/10.1037/a0019364

Csikszentmihalyi, M., & Csikszentmihalyi, I. S. (1992). *Optimal experience: Psychological studies of flow in consciousness*. Cambridge: Cambridge University Press.

Daniels, K. (2006). Rethinking job characteristics in work stress research. *Human Relations*, *59*, 267–290. https://doi.org/10.1177/0018726706064171

Dawson, J., & Richter, A. (2006). Probing three-way interactions in moderated multiple regression: Development and application of a slope difference test. *Journal of Applied Psychology*, *91*, 917–926. https://doi.org/10.1037/0021-9010.91.4.917

De Jong, B. A., Dirks, K. T., & Gillespie, N. (2016). Trust and team performance: A meta-analysis of main effects, moderators, and covariates. *Journal of Applied Psychology*, *101*, 1134–1150. https://doi.org/10.1037/apl0000110

De Jong, B. A., & Elfring, T. (2010). How does trust affect the performance of ongoing teams? The mediating role of reflexivity, monitoring, and effort. *Academy of Management Journal*, *53*, 535–549. https://doi.org/10.5465/AMJ.2010.51468649

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

Dirks, K. T., & Ferrin, D. L. (2001). The role of trust in organizational settings. *Organization Science*, *12*, 450–467. https://doi.org/10.1287/orsc.12.4.450.10640

Duong, D., Tuckey, M. R., Hayward, R. M., & Boyd, C. M. (2015). Work–family conflict: The importance of differentiating between different facets of job characteristics. *Work & Stress*, *29*, 230–245. https://doi.org/10.1080/02678373.2015.1074629

Feese, S., Arnrich, B., Tröster, G., Burtscher, M. J., Meyer, B., & Jonas, K. (2013). CoenoFire: Monitoring performance indicators of firefighters in real-world missions using smartphones. *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 83–92. https://doi.org/10.1145/2493432.2493450

Flin, R., O'Connor, P., & Mearns, K. (2002). Crew resource management: Improving team work in high reliability industries. *Team Performance Management: An International Journal*, *8*, 68–78.

Galecki, A., & Burzykowski, T. (2013). *Linear mixed effects models using R: A step-by-step approach*. New York, NY: Springer.

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge, MA: Cambridge University Press.

Gersick, C. J. G. (1988). Time and transition in work teams: Toward a new model of group development. *Academy of Management Journal*, *31*, 9–41. https://doi.org/10.2307/256496

Halbesleben, J. R. B., Neveu, J.-P., Paustian-Underdahl, S. C., & Westman, M. (2014). Getting to the "COR" understanding the role of resources in conservation of resources theory. *Journal of Management*, *40*, 1334–1364. https://doi.org/10.1177/0149206314527130

Hart, S. G. (2006). NASA-task load index (NASA-TLX); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *50*, 904–908. https://doi.org/10.1177/15419312060500090

Hart, S. G., & Staveland, L. E. (1988). Development of a multi-dimensional workload rating scale: Results of empirical theoretical research. In P. A. Hancock, & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). Amsterdam, The Netherlands: North Holland Press.

Harvey, S., Kelloway, E. K., & Duncan-Leiper, L. (2003). Trust in management as a buffer of the relationships between overload and strain. *Journal of Occupational Health Psychology*, *8*, 306–315. https://doi.org/10.1037/1076-8998.8.4.306

Hobfoll, S. E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist*, *44*, 513–524. https://doi.org/10.1037/0003-066X.44.3.513

Hobfoll, S. E. (2002). Social and psychological resources and adaptation. *Review of General Psychology*, *6*, 307–324. https://doi.org/10.1037/1089-2680.6.4.307

Ilgen, D. R., Hollenbeck, J. R., Johnson, M., & Jundt, D. (2005). Teams in organizations: From input-process-output models to IMOI models. *Annual Review of Psychology*, *56*, 517–543. https://doi.org/10.1146/annurev.psych.56.091103.070250

Ishak, A. W., & Ballard, D. I. (2012). Time to re-group: A typology and nested phase model for action teams. *Small Group Research*, *43*, 3–29. https://doi.org/10.1177/1046496411425250

Ishak, A. W., & Williams, E. A. (2017). Slides in the tray: How fire crews enable members to borrow experiences. *Small Group Research*, *48*, 336–364. https://doi.org/10.1177/1046496417697148

Jiang, L., & Probst, T. M. (2016). The moderating effect of trust in management on consequences of job insecurity. *Economic and Industrial Democracy*, 1–25. https://doi.org/10.1177/0143831X16652945

Johnsen, B. H., Espevik, R., Saus, E.-R., Sanden, S., Olsen, O. K., & Hystad, S. W. (2017). Hardiness as a moderator and motivation for operational duties as mediator: The relation between operational self-efficacy, performance satisfaction, and perceived strain in a simulated police training scenario. [Advance online publication]. *Journal of Police and Criminal Psychology*. doi:https://doi.org/10.1007/s11896-017-9225-1

Kahn, R. L., & Byosiere, P. (1992). Stress in organizations. In M. D. Dunnette, & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 3) (pp. 571–650). Palo Alto, CA: Consulting Psychologists Press.

Karantonis, D. M., Narayanan, M. R., Mathie, M., Lovell, N. H., & Celler, B. G. (2006). Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. *IEEE Transactions on Information Technology in Biomedicine*, *10*, 156–167. https://doi.org/10.1109/titb.2005.856864

Klein, K. J., Ziegert, J. C., Knight, A. P., & Xiao, Y. (2006). Dynamic delegation: Shared, hierarchical, and deindividualized leadership in extreme action teams. *Administrative Science Quarterly*, *51*, 590–621. https://doi.org/10.2189/asqu.51.4.590

Kozlowski, S. W. (2015). Advancing research on team process dynamics: Theoretical, methodological, and measurement considerations. *Organizational Psychology Review*, *5*, 270–299. https://doi.org/10.1177/2041386614533586

Kwapisz, J. R., Weiss, G. M., & Moore, S. A. (2011). Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter*, *12*, 74–82.

Langfred, C. W. (2004). Too much of a good thing? Negative effects of high trust and individual autonomy in self-managing teams. *Academy of Management Journal*, *47*, 385–399. https://doi.org/10.2307/20159588

Lazarus, R. S., & Folkman, S. (1984). *Stress appraisal and coping*. New York, NY: Springer.

LePine, J. A., Podsakoff, N. P., & LePine, M. A. (2005). A meta-analytic test of the challenge stressor–hindrance stressor framework: An explanation for inconsistent relationships among stressors and performance. *Academy of Management Journal*, *48*, 764–775. https://doi.org/10.5465/amj.2005.18803921

Li, F., Jiang, L., Yao, X., & Li, Y. (2013). Job demands, job resources and safety outcomes: The roles of emotional exhaustion and safety compliance. *Accident Analysis and Prevention*, *51*, 243–251. https://doi.org/10.1016/j.aap.2012.11.029

Lusa, S., Punakallio, A., Luukkonen, R., & Louhevaara, V. (2006). Factors associated with changes in perceived strain at work among fire-fighters: A 3-year follow-up study. *International Archives of Occupational and Environmental Health*, *79*, 419–426. https://doi.org/10.1007/s00420-005-0059-2

Mamen, A., Oseland, H., & Medbo, J. I. (2013). A comparison of two physical ability tests for firefighters. *Ergonomics*, *56*, 1558–1568. https://doi.org/10.1080/00140139.2013.821171

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *Academy of Management Review*, *26*, 356–376. https://doi.org/10.5465/AMR.2001.4845785

Mathieu, J. E., Maynard, M. T., Rapp, T., & Gilson, L. (2008). Team effectiveness 1997–2007: A review of recent advancements and a glimpse into the future. *Journal of Management*, *34*, 410–476. https://doi.org/10.1177/0149206308316061

Mathieu, J. E., Tannenbaum, S. I., Donsbach, J. S., & Alliger, G. M. (2014). A review and integration of team composition models moving toward a dynamic and temporal framework. *Journal of Management*, *40*, 130–160. https://doi.org/10.1177/0149206313503014

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

McAllister, D. J. (1995). Affect- and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of Management Journal*, *38*, 24–59. https://doi.org/10.2307/256727

McGrath, J. E. (1991). Time, interaction, and performance (TIP): A theory of groups. *Small Group Research*, *22*, 147–174. https://doi.org/10.1177/1046496491222001

Miller, G. (2012). The smartphone psychology manifesto. *Perspectives on Psychological Science*, *7*, 221–237. https://doi.org/10.1177/1745691612441215

Mishra, A. K., & Spreitzer, G. M. (1998). Explaining how survivors respond to downsizing: The roles of trust, empowerment, justice, and work redesign. *Academy of Management Review*, *23*, 567–588. https://doi.org/10.5465/amr.1998.926627

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*, 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Olguin, D. O., Waber, B. N., Kim, T., Mohan, A., Ara, K., & Pentland, A. (2009). Sensible organizations: Technology and methodology for automatically measuring organizational behavior. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *39*, 43–55. https://doi.org/10.1109/tsmcb.2008.2006638

Pentland, A. (2010). To signal is human. *American Scientist*, *98*, 204–210. https://doi.org/10.1511/2010.84.203

Piwek, L., & Joinson, A. (2017). Automatic tracking of behavior with smartphones: Potential for behavior change interventions. In L. Little,

E. Sillence, & A. Joinson (Eds.), *Behavior change research and theory: Psychological and technological perspectives* (pp. 137–165). San Diego, CA: Academic Press.

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, *63*, 539–569. https://doi.org/10.1146/annurev-psych-120710-100452

R Core Team. (2015). R: A language and environment for statistical computing. Vienna, Austria. Retrieved from http://www.r-project.org/

Roe, R. A. (2008). Time in applied psychology: The study of "what happens" rather than "what is". *European Psychologist*, *13*, 37–52. https://doi.org/10.1027/1016-9040.13.1.37

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review*, *23*, 393–404. https://doi.org/10.5465/AMR.1998.926617

Salas, E., & Rosen, M. A. (2013). Building high reliability teams: Progress and some reflections on teamwork training. *BMJ Quality & Safety*, *22*, 369–373. https://doi.org/10.1136/bmjqs-2013-002015

Salas, E., Sims, D. E., & Burke, C. S. (2005). Is there a "Big Five" in teamwork? *Small Group Research*, *36*, 555–599. https://doi.org/10.1177/1046496405277134

Schaufeli, W. B., & Bakker, A. B. (2004). Job demands, job resources, and their relationship with burnout and engagement: A multi-sample study. *Journal of Organizational Behavior*, *25*, 293–315. https://doi.org/10.2307/4093692

Schaufeli, W. B., & Taris, T. W. (2014). A critical review of the job demands–resources model: Implications for improving work and health. In G. F. Bauer, & O. Hämmig (Eds.), *Bridging occupational, organizational and public health: A transdisciplinary approach* (pp. 43–68). New York, NY: Springer Science + Business Media.

Schmid Mast, M., Gatica-Perez, D., Frauendorfer, D., Nguyen, L., & Choudhury, T. (2015). Social sensing for psychology: Automated interpersonal behavior assessment. *Current Directions in Psychological Science*, *24*, 154–160. https://doi.org/10.1177/0963721414560811

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

Spector, P. E. (1994). Using self-report questionnaires in OB research: A comment on the use of a controversial method. *Journal of Organizational Behavior*, *15*, 385–392. https://doi.org/10.1002/job.4030150503

Spector, P. E. (2002). Employee control and occupational stress. *Current Directions in Psychological Science*, *11*, 133–136. https://doi.org/10.1111/1467-8721.00185

Spector, P. E., Dwyer, D. J., & Jex, S. M. (1988). Relation of job stressors to affective, health, and performance outcomes: A comparison of multiple data sources. *Journal of Applied Psychology*, *73*, 11–19. https://doi.org/10.1037/0021-9010.73.1.11

Sundstrom, E., McIntyre, M., Halfhill, T., & Richards, H. (2000). Work groups: From the Hawthorne studies to the work teams of the 1990s. *Group Dynamics: Theory, Research, and Practice*, *4*, 44–67. https://doi.org/10.1037//1089-2699.4.1.44

van Den Tooren, M., & de Jonge, J. (2008). Managing job stress in nursing: What kind of resources do we need? *Journal of Advanced Nursing*, *63*, 75–84. https://doi.org/10.1111/j.1365-2648.2008.04657.x

Wageman, R. (1995). Interdependence and group effectiveness. *Administrative Science Quarterly*, *40*, 145–180. https://doi.org/10.2307/2393703

Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (1999). Organizing for high reliability: Processes of collective mindfulness. In R. I. Sutton, & B. M. Staw (Eds.), *Research in Organizational Behavior* (Vol. 21) (pp. 81–123). Stanford, CA: JAI Press.

Weigl, M., Müller, A., Sevdalis, N., & Angerer, P. (2013). Relationships of multitasking, physicians' strain, and performance: An observational study in ward physicians. *Journal of Patient Safety*, *9*, 18–23. https://doi.org/10.1097/PTS.0b013e31826b7b87

Wildman, J. L., Shuffler, M. L., Lazzara, E. H., Fiore, S. M., Burke, C. S., Salas, E., et al. (2012). Trust development in swift starting action teams: A multilevel framework. *Group & Organization Management.*, *37*, 137–170. https://doi.org/10.1177/1059601111434202

Wilson, K. A., Burke, C. S., Priest, H. A., & Salas, E. (2005). Promoting health care safety through training high reliability teams. *Quality and Safety in Health Care*, *14*, 303–309. https://doi.org/10.1136/qshc.2004.010090

van Woerkom, M., Bakker, A. B., & Nishii, L. H. (2016). Accumulative job demands and support for strength use: Fine-tuning the job demands–resources model using conservation of resources theory. *Journal of Applied Psychology*, *101*, 141–150. https://doi.org/10.1037/apl0000033

Michael J. Burtscher is a senior lecturer and researcher in the Department of Psychology at the University of Zurich. He obtained his PhD in Work and Organizational Psychology from ETH Zurich. His research interests include team cognition and adaptation, social sensing of interaction processes, and job resources in high-reliability industries.

Bertolt Meyer is a Professor of Organizational and Economic Psychology at Technische Universität Chemnitz. He obtained his PhD from Humboldt University Berlin. His research focuses on team diversity, leadership, and stereotypes. He serves as an associate editor for Small Group Research.

Klaus Jonas is a Professor of Social and Business Psychology at the University of Zurich. He obtained his PhD from the University of Tübingen. He is a coeditor of an international social psychological textbook (together with M. Hewstone and W. Stroebe, sixth English edition, translation into eight additional languages).

Sebastian Feese is a senior researcher at the Wearable Computing Lab of ETH Zürich. He received his PhD in Electrical Engineering from ETH Zürich in 2014. In his doctoral thesis, he investigated methods to automatically observe social interaction and coordination in teams using wearable sensors. His research interests are team and crowd behavior sensing, social signal processing, and machine learning. In 2015, he cofounded antavi GmbH.

Gerhard Tröster is a Professor and Chair of the Electronics Laboratory at the ETH Zurich. He obtained his doctorate from the Technical University of Darmstadt about the design of analog integrated circuits. At ETH Zurich, his fields of research include wearable computing, flexible electronics, human–computer interaction, and detection of the physical, mental, and social contexts of the user aiming at applications in health care and music.

How to cite this article: Burtscher MJ, Meyer B, Jonas K, Feese S, Tröster G. A time to trust? The buffering effect of trust and its temporal variations in the context of high-reliability teams. *J Organ Behav*. 2018;39:1099–1112. https://doi.org/10.1002/job.2271

APPENDIX A

SMARTPHONE-BASED MEASUREMENT OF PHYSICAL ACTIVITY

We chose the Sony Xperia Active Smartphone because of its built-in sensors and its robustness (i.e., dust and water resistance). We developed an Android™ app to sample the phone's built-in sensors: Acceleration sensors were used to measure participants' body movements (please see also Feese et al., 2013 for a description of the technical background).

To quantify motion activity during the scenario, we used a threshold-based approach: The motion data were divided into active segments (i.e., acceleration is above the activity threshold) and nonactive segments (i.e., acceleration is below active threshold). A segment had a duration of 1 s, which means each performance episode consisted on average of 745 segments. For each 1-s segment, the smartphone's triaxial accelerometer provided 50 measurements of the changes in *X*, *Y*, and *Z* directions. Thus, on average, more than 100,000 individual data points were used for the calculation of each physical activity score. The data were aggregated in a two-step procedure. First, using a two-component Gaussian mixture model, the activity threshold was learned from all available data (>10 million data points). Second, the threshold was used to decide for each segment (i.e., each second) whether or not the firefighter was active. Similar procedures have been used in other studies (Bulling, Blanke, & Schiele, 2014; Karantonis, Narayanan, Mathie, Lovell, & Celler, 2006; Olguin et al., 2009). Overall motion activity was then calculated as the ratio of active segments during the scenario. Values can potentially range from 0 (i.e., *person did not move at all during the scenario*) to +1 (i.e., *person was always physically active*).

A recent review of the literature indicates that triaxial accelerometer data—as were used in the present study—are strongly correlated to indicators of physical activity such as oxygen consumption and energy expenditure (Butte, Ekelund, & Westerterp, 2012). This suggests that our approach of operationalizing physical activity constitutes a valid measurement strategy. To assess construct validity more directly, we related our measure to the TLX item capturing physical strain.[4] Multilevel regression revealed a strong positive relationship between smartphone data and perceived physical strain ($\gamma = 12.43$, $p < .001$; 95% CI [8.84, 16.09]), which provides further evidence for the validity of our measure.

[4]We thank the editor for this suggestion.