The Leadership Quarterly xxx (xxxx) xxx

Contents lists available at ScienceDirect

The Leadership Quarterly

journal homepage: www.elsevier.com/locate/leaqua

Full Length Article

# Physiological arousal variability accompanying relations-oriented behaviors of effective leaders: Triangulating skin conductance, video-based behavior coding and perceived effectiveness

Marcella A.M.G. Hoogeboom [a,*], Aaqib Saeed [b], Matthijs L. Noordzij [a], Celeste P.M. Wilderom [a]

[a] Faculty of Behavioural, Management and Social Sciences, University of Twente, P.O. Box 217, 7500AE Enschede, the Netherlands
[b] Faculty of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600MB Eindhoven, the Netherlands

ARTICLE INFO

*Article history:*
Received 19 July 2018
Received in revised form 16 September 2020
Accepted 4 December 2020
Available online xxxx

*Keywords:*
Physiological arousal
Task- and relations-oriented leader behavior
Healthy variability
Video-based observation
Leader effectiveness

ABSTRACT

With the aim of extending the healthy physiological variability thesis to Leadership Studies, we examined the hypothesized links among leaders' within-person variability in physiological arousal, their task- and relations-oriented behaviors and their overall effectiveness. During regularly-held staff meetings, wristband skin sensors and video cameras captured synchronized physiological and fine-grained behavioral data of 36 leaders within one organization. Perceived leader effectiveness ratings were obtained from their followers. Multi-level log-linear analyses showed no elevated levels of arousal during the task-oriented behaviors of both the highly effective and the less effective leaders. The highly effective leaders showed a significantly greater likelihood of high levels of physiological arousal during positive and negative relations-oriented behaviors. We thus report a physiological correlate of relations-oriented leader behavior; especially among the most effective leaders, higher levels of arousal co-occurred with their positive and negative relations-oriented behavior in the meetings. Having used two high-resolution methods to advance insights about effective organizational leadership, this field study illuminates the importance of capturing the co-occurrence of within-person variability in leaders' bodily responses and their precisely measured behaviors over time in a functional social setting at work.



## Introduction

Because effective leadership is a precursor of high organizational performance, many studies have focused on the underlying processes of leadership. Recently, it has been argued that effective leadership might be associated with within-person *physiological* processes (e.g., Ashkanasy, Becker, & Waldman, 2014; Boyatzis et al., 2012; Damen, van Knippenberg, & van Knippenberg, 2008; Murray & Antonakis, 2019). Leaders' physiological arousal, for instance, is held to inform our understanding of effective leader interactions with followers (Akinola, 2010; Antonakis, Ashkanasy, & Dasborough, 2009; Boyatzis, Rochford, & Taylor, 2015). Despite the growing number of calls to conduct integrative research that combines human physiological and behavioral predictors of workplace outcomes (Arvey & Zhang, 2015; Boyatzis et al., 2012; Heaphy & Dutton, 2008; Zyphur, Narayanan, Koh, & Koh, 2009), empirical studies to date have examined within-leader physiological arousal and leader behavior *separately*. This is remarkable, because earlier research has suggested that physiological and behavioral processes are intertwined (Colarelli & Arvey, 2015; Erez, Misangyi, Johnson, LePine, & Halverson, 2008). An examination of the co-occurrence of level of physiological arousal and specific leader behavior may thus provide insights into the bodily correlates of effective leader behavior. In the present field study, we tested whether distinct leader behaviors are associated with different levels of physiological arousal and whether this co-occurrence of observable behaviors and physiological arousal is different for highly effective vs. less effective leaders.

In pursuing a joint investigation of leaders' within-person variability in physiological arousal and their behaviors, one cannot bypass the so-called healthy variability thesis (Navarro & Rueff-Lopes, 2015). This thesis centers on the predictive value of high fluctuation in physiological arousal in social contexts. Originating in the field of human physiology, the 'healthy' variability thesis holds that to be effective, fluctuation in one's physiological system must be aligned with the fluctuation of cues induced by one's social environment (Navarro & Rueff-Lopes, 2015). Transposed to organizational contexts, Navarro and Rueff-Lopes (2015) argued that organizational environments entail complex social situations in which a host of social cues must be processed at the same time. They implied that if a leader's physiological arousal fluctuation in such situations is high, high performance is more likely to occur. A lack of variability in physiological arousal in such a social setting is then an indicator of a leader's lack of sensitivity to social cues. Navarro and Rueff-Lopes (2015, p. 537) called this latter "unhealthy stability."

* Corresponding author.
*E-mail address:* a.m.g.m.hoogeboom@utwente.nl (M.A.M.G. Hoogeboom).

https://doi.org/10.1016/j.leaqua.2020.101493
1048-9843/© 2020 Published by Elsevier Inc.

Based on this thesis, it can be assumed that within-person variability in physiological arousal (i.e., the fluctuation in levels of arousal in one's physiological system) is associated with particular behaviors (i.e., the cues or social interactions in one's work environment) and effectiveness. To test this crux of the healthy variability thesis in an organizational-leadership context one can examine whether different levels of leader physiological arousal co-occur with specific leader behavior and whether this is associated with leader effectiveness. Testing this thesis among a group of leaders requires a temporal account of both their behavioral display during interactions with their staff and their physiological levels. Hence, when applying the healthy variability thesis to leadership research, particular leader behaviors accompanied by variation in their physiological arousal might thus be an indicator of high leader performance or effectiveness.

Drawing upon the functional approach to organizational leadership, all critical leadership functions are manifested in the behavior of a leader (Burke et al., 2006; McGrath, 1962; Morgeson, DeRue, & Karam, 2010). One of the most well-established leader-behavior categorizations distinguishes task-oriented from relations-oriented behaviors (Behrendt, Matz, & Göritz, 2017; Judge & Piccolo, 2004; Schriesheim & Bird, 1979). In order to cover an even broader range of real-life organizational behaviors of leaders, we also consider positive vs. negative relations-oriented behavior (Meinecke, Kauffeld, & Lehmann-Willenbrock, 2017). Task-oriented and positive and negative relations-oriented leader behaviors fulfill essential leadership functions, especially during social interactions with followers. Due to their different functions, different physiological processes might accompany these three types of behaviors. Despite the wealth of research on the associations between perceptions of task- and relations-oriented leader behaviors and leader effectiveness (DeRue, Nahrgang, Wellman, & Humphrey, 2011; Judge, Piccolo, & Ilies, 2004), the question of how physiological arousal may accompany these leader behaviors has remained underexplored. Thus, capturing the physiological arousal of leaders when interacting with their followers can enhance our knowledge about how our social workplace behaviors are related to our biological system, as intimated, for instance, by Arvey and Zhang (2015).

Building upon the theoretical accounts of healthy variability and functional leadership theory, we aim to advance our understanding of leadership effectiveness by answering the question: During social interactions with followers, how is leaders' level of physiological arousal associated with their task-oriented and (positive and negative) relations-oriented behaviors and their leadership effectiveness? We combine unobtrusive measures of skin conductance to capture leaders' within-person variation in physiological arousal during staff meetings with the simultaneous collection of time-stamped behavioral event data (e.g., Hoogeboom & Wilderom, 2015; Hoogeboom & Wilderom, 2019; Hoogeboom & Wilderom, 2020) and effectiveness ratings. Our time-based field research enables rigorous analyses of synchronized multi-model data, which enables the study of fluctuations in physiological arousal during various expressions of different types of leader behavior. Our precise measurement and systematic coding of fine-grained leader behaviors during real-life workplace interactions in the field, which captures leader behavior as "concrete behavioral acts" (Van Quaquebeke & Felps, 2018, p. 2), is still fairly unique, despite the great demand for direct observational field research instead of merely relying on (introspective) self-reports, hypothetical scenarios, and questionnaire ratings of behavior (e.g., Baumeister, Vohs, & Funder, 2007; Hoffman & Lord, 2013). By taking a temporal, functional perspective on the simultaneous co-occurrence of real-life leader behavior and physiological arousal, thereby bridging two distinct research disciplines, we contribute to research on effective leader behavior and the quest for what it is that makes leaders' interactions with their followers effective, as called for by Lehmann-Willenbrock and Allen (2018). In doing so we also contribute to the literature on the influential role of physiology in the workplace. The tested associations between leaders' level of physiological arousal and their fine-grained behaviors are meant to inform us in triangulated and realistic detail about the social fabric and dynamics of effective leadership.

## Theoretical framework

### *Healthy variability in OB research*

Transposed to the work context, the healthy variability thesis holds that to effectively cope with the social exigencies in the workplace, an individual should flexibly respond to what this environment requires, which is also referred to as "requisite variety" (Ashby, 1956). The thesis thus posits that in a 'healthy' or high-performing individual, the variability within his or her physiological system must match the social cues in the environment (Navarro & Rueff-Lopes, 2015). Building upon this thesis, we focus on a continuous measure of physiological arousal that allows us to track how arousal varies over time. In addition, to examine the thesis, the social cues in the work environment are reflected in observable leader behavior expressed in that situation. The specific behaviors that a leader shows while interacting with followers are presumed to indicate whether he or she understands or is sensitive to the social environment in which he or she operates (Behrendt et al., 2017). To investigate the presumed co-occurrence of physiological variability with the behavior manifested, a dynamic or temporal account of leader behaviors is called for. Examining the dynamic co-occurrence of the level of physiological arousal and leader behavior also fits with the emphasis placed by OB scholars on better understanding the dynamics of leading at work (e.g., McClean, Barnes, Courtright, & Johnson, 2019). Instead of a static account of important workplace variables, such as leader behavior, which has prevailed in OB research to date (Lehmann-Willenbrock & Allen, 2018), adopting a within-person variability perspective—as we do herein—might lead to better understanding of the dynamics of (effective) workplace behavior. During interactions with followers, a leader typically does not present one single type of behavior, but exhibits variability in his or her specific behaviors to accommodate the social exigencies of the workplace. Hence, the specific behaviors that a leader shows can change over time. By testing the healthy variability thesis in an organizational setting, we do not consider behavior here as a static feature, but focus on the variability in specific leader behaviors shown over time.

There is some earlier OB research that also assumed physiological arousal to be associated with behavioral phenomena (Akinola, 2010; Antonakis et al., 2009; Boyatzis et al., 2015), in line with the premises of the healthy variability thesis. This suggests that specific leader behaviors may be accompanied by distinct levels of physiological arousal. However, to date, evidence concerning how possible variations in the degree of physiological arousal are reflected in (effective) behavior is absent in the leadership literature, which restricts our understanding of the effectiveness of leaders' real-time interactions with their followers. Given the premises of the healthy variability thesis, we focus here on the co-occurrence of physiological arousal and leader behavior. In doing so, our research differs in an important way from prior studies that have examined the healthy variability thesis in OB research. Empirical tests of this thesis have focused on the dynamic nature of only *one* phenomenon, for instance, either work motivation or team coordination or performance. They found that there are fluctuations in these work-related variables (Arrieta, Navarro, & Vicente, 2008; Gorman, Cooke, & Amazeen, 2010; Guastello et al., 2014). For instance, when an individual showed performance that fluctuates between high and moderate levels, this led to higher overall performance (Guastello et al., 2014). While these studies have shown that variability in important phenomena, including behavior, can inform our understanding of the workplace as a dynamic, social context, they do not tell us how variability in important workplace phenomena is associated with bodily phenomena that co-occur. In our test of the healthy variability thesis, the dynamically occurring, minute leader behaviors are captured simultaneously with the leaders' subtle fluctuations in physiological arousal.

*Skin conductance as a parameter of physiological arousal*

How people respond (neuro)physiologically to stimuli has been examined with several methods and techniques, including, for instance, functional magnetic resonance imaging (fMRI) for scanning the brain, and electrocardiography (ECG) for measuring heart function. Some leadership studies have applied such cognitive neuroscience and physiological techniques, under the umbrella of what is now termed organizational cognitive neuroscience (Lee, Senior, & Butler, 2012; Senior, Lee, & Butler, 2011). The focus of such studies is on the biological or brain systems and how they are associated with cognition and behavior during social workplace interactions (see, e.g., Balthazard, Waldman, Thatcher, & Hannah, 2012; Spain & Harms, 2014).

A peripheral physiological phenomenon with well-understood couplings to the central nervous system is electrodermal activity (EDA), most often approximated using skin conductance (Boucsein, 2012). To date, EDA is seen as the most precise physiological measure of arousal (Lidberg & Wallin, 1981; Marci, Ham, Moran, & Orr, 2007; Picard, Fedor, & Ayzenberg, 2016). EDA is considered to be "the result of direct mediation by the sympathetic branch of the autonomic nervous system" (Marci et al., 2007, p. 104), especially during social interactions. This means that, unlike heart rate, EDA is not directly influenced by the parasympathetic nervous system. This is important, because neurohormonal influences from the parasympathetic nervous system can potentially confound other physiological variables, such as heart rate (Cacioppo & Tassinary, 2000). Thus, compared to other physiological measures, skin conductance best captures the intensity of emotions experienced during workplace interactions (Akinola, 2010; Figner & Murphy, 2011). EDA has also been used within the context of neuroscience. For instance, neuromarketing studies have approximated the physiological responses in the body during decision making (e.g., Bechara, Damasio, Damasio, & Lee, 1999; Gakhal & Senior, 2008).

Skin conductance is captured through variations in the eccrine sweat glands (present in all bodily parts, with the highest density in the palms and soles) in response to sweat secretion from the skin (e.g., Benedek & Kaernbach, 2010). These changes in eccrine sweating or skin-conductance responses are interconnected with the sympathetic branch of the autonomic nervous system (SNS). The origins of eccrine sweating are tied to brain networks including the amygdala, the hippocampus, the hypothalamus, the brainstem and the prefrontal cortex. It has been shown that these brain areas (such as the amygdala and the prefrontal cortex) are tied to affective processes, and are implicated in skin conductance responses (SCRs) in relation to emotional stimuli, but not in SCRs related to non-emotional stimuli such as taking a deep breath (Naqvi & Bechara, 2006; Tranel & Damasio, 1989, 1994). Furthermore, the SNS is responsible for producing neuronal and hormonal stress responses (e.g., the fight-or-flight response) and has been found to significantly impact emotional processes and people's motivation (Boucsein, 2012; Figner & Murphy, 2011). Changes in skin conductance have been found to be strongly associated with changes in human emotion (Pennebaker, Hughes, & O'Heeron, 1987). Hence, skin conductance is commonly used as an index of general and emotional arousal, attention and intensity of emotions (Akinola, 2010; Figner & Murphy, 2011).

The skin conductance signal provides information about the intensity of physiological arousal, but does not specify the psychological state (e.g., happiness or fear) associated with it (e.g., Akinola, 2010). Hence, one cannot draw inferences from the intensity of arousal about the exact emotional states being experienced (e.g., Boucsein, 2012; Larsen, Diener, & Lucas, 2002). When trying to better understand workplace processes, such as performance or learning, physiological intensity is seen as a crucial biomarker; in a leader development program, the highest amount of learning occurred during critical situations (i.e., when the highest physiological arousal occurred: during a public speaking and difficult conversation exercise; Waller, Reitz, Poole, Riddell, & Muir, 2017). Hence, higher physiological intensity (and not the emotional valence, that is, whether the emotion is positive or negative) was positively associated with long-term learning effects after leader development training. Other empirical work has established that experiencing higher levels of arousal is a more important predictor of recall and forgetting rates than the valence of these emotions (Talarico, LaBar, & Rubin, 2004); when participants were asked to recall an autobiographical event that varied in valence and intensity (i.e., low or high), intensity predicted significantly more variance in memory recall. These results could indicate that a person more strongly experienced those learning or autobiographical events.

More generally, variability in skin conductance captures the bodily signals from the peripheral nervous system that reflect what are termed secondary emotions; the person who experiences such physiological fluctuations is not always conscious of these peripheral (positive or negative) emotions. These biomarkers of human experiences reflect what happens in the social environment (e.g., Bechara, Damasio, Tranel, & Damasio, 2005). As a reflection of that social environment, different within-person physiological variability might be associated with different behaviors that are shown during interactions with others. Physiological responses can thus be considered critical inferential elements that can reflect the diversity in workplace behaviors (Christopoulos, Uy, & Yap, 2019). As such, they "may help to elucidate how physiological processes can underpin and even modulate affective, cognitive, and motivational processes relevant to organizational settings" (Akinola, 2010, p. 204). Extrapolating from that, high physiological intensity or arousal might thus accompany socially sensitive leader behaviors.

*Leaders' task- vs. relation-oriented behaviors*

Only a handful of scholars in the field of Organizational Behavior (OB) have presumed an association between skin conductance levels and important *behaviors* at work (e.g., Ashkanasy et al., 2014; Bormann & Rowold, 2016). In order to empirically examine this association, one must first delineate what leaders "actually do" in terms of their behaviors (i.e., their actions) while interacting with their followers. To operationalize the behavioral repertoire of leaders during interactions with followers, we started out with the well-established Ohio state leadership model (Fleishman, 1973) which includes two broad behavioral dimensions: initiating structure (also referred to as task-oriented behavior) and consideration (also referred to as relations-oriented behavior) (Fleishman, 1973; Judge et al., 2004; Schriesheim & Bird, 1979). We used this classic dichotomy because leader behaviors must "essentially be (1) task-oriented and (2) relations-oriented" (Behrendt et al., 2017, p. 233). This task- vs. relations-oriented dichotomy served as an important foundation for another well-known leader-behavioral model: the transformational-transactional model (Bass, 1985) which is seen by some as "conceptually similar" to relations- and task-oriented leader behavior (DeRue et al., 2011, p. 9; see also Burke et al., 2006). According to Fleishman et al. (1991), between 1940 and 1986 alone, 65 classification systems of leader behavior had been proposed, nearly all with one common theme: that leader behaviors were broken down into one of two categories: those dealing with task accomplishment (i.e., task-focused) and those facilitating team interaction and/or development (i.e., person-focused).

However, leader effectiveness studies have predominantly relied on surveys to assess task- and relations-oriented behaviors. These measures were typically created on the basis of perceptions from experts and others about how leaders behaved (Behrendt et al., 2017; Yukl, 2010). Because survey-based perceptions of behaviors do not necessarily reflect *observable* behaviors or specific behavioral events during moment-to-moment interactions with followers (Collinson, 2005; Day, Gronn, & Salas, 2004; Uhl-Bien, 2006), the survey-based operationalized behaviors based on this theoretically important distinction are unlikely to run entirely parallel to those in a leader-behavior observational coding scheme. Therefore, below we first lay out our theoretical framework, in which we will explain how the classic behavioral descriptions informed a scheme with which one can validly code actual, observable, in-situ behaviors at the event level. In doing so, we see similarity to recent work that took an interaction coding approach (Kauffeld & Lehmann-Willenbrock, 2012; Meinecke et al., 2017), and also to the behavioral scheme developed by Bales (1950), who was one of the first to distinguish between relations- (or person- or social-emotionally) oriented and task-oriented behavioral or communicative events during social interactions, but whose taxonomy has never been extensively validated.

*Task-oriented* leader behavior refers to behaviors that promote the followers' accomplishment of work tasks or goal attainment in an efficient manner, by defining and organizing both one's own role and the role of one or more followers, including the creation of clear patterns and channels of communication (Fleishman, 1973; Judge et al., 2004). Task-oriented or initiating structure leader behaviors thus ensure that employees are well-informed about the work processes (i.e., establishing clear patterns and channels of communication), and provide structure about their roles (i.e., clarifying role expectations) and direction about how to plan activities (i.e., clarifying task objectives and planning short-term activities), thus promoting high task productivity (Bass, 1990; DeRue et al., 2011). When translating this to the behavioral event level of actual observable task-oriented leader behavior during interactions with followers, task-oriented leader behavior comprises the sharing of organizationally-relevant information, guiding followers to ensure that they perform a particular number of tasks in particular ways (i.e., directing) and providing structure (Burke et al., 2006; Hoogeboom & Wilderom, 2019; Pearce et al., 2003; Sims & Manz, 1984).

In addition, transactional behavior and more specifically the management-by-exception behaviors (Sommer, Howell, & Hadley, 2016), has typically been conceptualized as task-oriented leader behavior (Burke et al., 2006; DeRue et al., 2011). During interactions with followers (i.e., at the behavioral event level), transactional management-by-exception behavior entails monitoring of followers' task processes and providing negative task feedback and corrective action if deemed necessary to ensure organizational goal accomplishment (Bass & Avolio, 1995; Ewen et al., 2013; Vecchio, Justin, & Pearce, 2008; Wang, Tsui, & Xin, 2011). If followers are seen to underperform, a leader may thus engage in such behaviors to help them improve their performance (Bass & Avolio, 1995; Sommer et al., 2016). Leaders then typically "specify the standards for compliance, as well as what constitutes ineffective performance" (Bass & Avolio, 1995, p. 97). When a leader provides negative task feedback and corrects followers, the leader behavior involved addresses a discrepancy between actual and aimed for task-based performance (Cianci, Klein, & Seijts, 2010; London & Smither, 2002). Hence, when leaders display task-oriented behavior, they also address task behaviors that the leader does *not* wish to see among followers. Such task-oriented feedback often sets a norm or attempts to sharpen up (i.e., enhance the quality or quantity of) future task output by followers (Morgeson et al., 2010). It can also entail a message from a leader in which he or she ascribes a cause for not reaching a given level of task performance (Gioia & Sims Jr., 1986). Thus, when a leader corrects or provides negative task feedback to one or more followers, information is offered implying that and/or how they could improve their performance (Cianci et al., 2010; London & Smither, 2002). Despite its explicit negative connotation, 'negative *task* feedback' to followers is referred to (and operationalized) here, in conformity with all earlier leader-behavior operationalizations, as stressing followers' *task* behavior with the explicit goal of enhancing task performance.

In addition to the observable task-oriented initiating structure and transactional behaviors, another subset of task-oriented leader behavior can be shown during interactions with followers. This occurs when the leader provides or "gives an opinion" or view about the task objective or requirements for adequately performing a task (Bales, 1950, p. 258; Kauffeld & Lehmann-Willenbrock, 2012). Moreover, creating consensus about task elements is another important task-oriented behavior of a leader (Kauffeld & Lehmann-Willenbrock, 2012), which can translate into agreeing or disagreeing with task directions at the behavioral event level.

A leader who interacts with followers will also engage in socioemotional-, person- or relations-oriented type of interactions (Bales, 1950). Such *relations-oriented behavior* is typically defined as showing concern for followers' (individual) needs, providing support and showing appreciation (Bass & Bass, 2008; Keyton & Beck, 2009). Leaders who display relations-oriented behavior during interactions with followers might show behaviors of (individualized) consideration (i.e., expressing appreciation and emotional support for followers, being concerned for followers' development and looking out for follower wellbeing: see Piccolo et al., 2012). In addition to being considerate, a leader who displays relations-oriented behavior can also engage more specifically in "(1) fostering coordination to synchronize efforts, (2) promoting cooperation to encourage greater individual contribution, and (3) activating resources to expand valuable contributions" (Behrendt et al., 2017, p. 236). Such behaviors are also in line with transformational leader behavior that inspires followers to perform beyond expectations by intellectually stimulating them, providing inspirational motivation through giving positive feedback and articulating a compelling vision through idealized influence behavior (Bass & Avolio, 1995). These behaviors have also been classified as relations-oriented, because they address followers' "developmental and self-actualization aspects," and are therefore considered to be "person-focused" (Burke et al., 2006, p. 293); they elicit or maintain follower engagement "by synchronizing collective efforts and increasing the likelihood of appropriate contributions" (see Behrendt et al., 2017, p. 293, who also classified such behavior as relations-oriented). Thus, relations-oriented leader behavior involves behavior that explicates or stresses elements in the bond between a leader and follower(s) that the leader deems of importance for the follower as a person and/or member of staff (Bales, 1950; Ilgen, Fisher, & Taylor, 1979).

Thus, whereas task-oriented leader behavior involves task-related communication that does not explicitly addresses the person, relations-oriented leader behaviors speak to followers' (personal) interests and can strengthen followers' experience of self-determination and being empowered. Furthermore, relations-oriented leader behaviors can flag that each followers' individual contribution and ideas are indispensable (Behrendt et al., 2017; Burke et al., 2006). Especially in the context of leader-follower interactions at work, humor and building co-operative relationships, based on a shared vision and ideas, including the sharing of *personal* information, often directly contributes to the quality of the (socioemotional sides of) work relations (Lehmann-Willenbrock & Allen, 2014). Thus, relations-oriented leader behavior is often represented by leader expressions conveying a socioemotional aspect of the relationship between a leader and one or more followers (i.e., "person-focused" verbalizations: Burke et al., 2006, p. 293). Whenever leaders express relations-oriented behavior during interactions with their followers, they relate themselves (as a person in the role of leader) explicitly to one or more persons in the role of followers through something the leader deems of value to the followers. A leader's task-oriented behavior, on the other hand, denotes behavior where task-based performance expectations or information and standards are being stressed (including directing, informing, structuring, monitoring, feedback and correcting) to ensure adequate goal accomplishment. Therefore such behaviors strictly regard followers' "task-focused" communication and thus have a different function than relations-oriented behavior (Burke et al., 2006, p. 288). In other words, whenever leaders emphasize task-type behavior during interactions with their followers, they relate themselves (in the role of leader) *not* explicitly to these followers as subjective or personified listeners, but only to (elements of) their tasks (such as projects, budgets, figures, actions, decisions, solutions: i.e., task aspects), rather than to work aspects that the leader deems important for the personal or professional growth of one or more followers.

Previous work has shown that the classic task- vs. relations-oriented behavioral distinction has been very robust (Fleishman & Harris, 1998; Meinecke et al., 2017). Moreover, recent work that has 'translated' earlier (questionnaire-based) leader behavioral taxonomies to actual observational events has also used the distinction between task- vs. relations-oriented behavior as a high-level theoretical organizing frame that covers leading other people's work in interactive ways (Behrendt et al., 2017; Meinecke et al., 2017). Hence, for knowledge accumulation purposes, invoking this behavioral frame in the context of a first test of the dynamic healthy variability thesis has merit.

Whereas the above relations-oriented behaviors are all positive types of behaviors, the current leadership and leader communication literature differentiates positive from negative relations-oriented behaviors (Meinecke et al., 2017). Hence, an important kind of leader behavior that might be observable and influential during interactions with followers is *negative relations-oriented behavior* (Meinecke et al., 2017). This negative strand reflects leader behaviors that are *not* conducive to high job performance (Dalal, 2005). Extreme forms of negative relations-oriented behavior have been discussed in the literature, such as destructive leadership (behavior that disturbs followers: Einarsen, Aasland, & Skogstad, 2007) or abusive supervision (nonphysical aggression by leaders aimed at their followers: Tepper, Henle, Lambert, Giacalone, & Duffy, 2008). These less desirable leader behaviors typically include intense forms, including belittling, (loud) outbursts, malice or tyranny. The present video-observational field study includes coding for three relatively mild forms of such leader behavior during interactions with followers, which are also included in other behavioral classifications (Meinecke et al., 2017): showing disinterest, defending one's own position and interrupting. These three categories of behavior (disinterest, defensiveness and disruptiveness) are labeled here as negative relations-oriented leader behavior at the (mutually exclusive) behavioral event level. Together, they cover the milder forms of counterproductive behavior (Penney & Spector, 2005). Thus, three categories of leader behavior will be considered next, in conjunction with leader physiological arousal and effectiveness.

### *Physiological arousal and positive relations-oriented leader behavior*

Heaphy and Dutton (2008, p. 137) argued that, especially during positive social interactions, "beneficial physiological effects," for example, physical health and work engagement, can enhance human capital. More specifically, they suggested that when employees experience more positive social workplace interactions, their physiological resourcefulness (i.e., healthy heart rate and blood pressure at and after work) leads to higher levels of physical health. Although their review only linked positive social interactions with enduring physiological effects, we might infer from their ideas that, when showing positive leader behavior (such as providing individualized consideration and giving positive feedback), activation of physiological markers accompanies this behavior. In other words, their review made it plausible to assume an association between positive workplace interactions and the physiological effects on the workers, including the leaders.

Prior work in the field of emotions has also found that positive emotions, such as happiness, which usually accompany positive relations-oriented behavior, can be associated with high physiological arousal (Heaphy & Dutton, 2008). For example, studies that related viewing various facial expressions or emotional movie segments to participants' arousal levels found the largest physiological reactions when positive expressions or happy film scenes were shown (Golland, Keissar, & Levit-Binnun, 2014; Vrana & Gross, 2004). Furthermore, when evaluating positive emotional pronoun-noun phrases (e.g., "my happiness") on a computer, people were more aroused than when evaluating neutral or other-related pronoun-noun sentences (e.g., "his happiness") (Weis & Herbert, 2017). This means that positive words related to the self may fuel the highest physiological response. Experiencing positive expressions that are also related to interpersonal or human relations that also automatically link to the self may thus elicit human physiological reactions. Especially in a social context, expressing relations-oriented behavior that is essentially directed towards others is likely to be associated with the experiencing of positive social interactions at work by a leader themself, which might have a physiologically arousing effect (Heaphy & Dutton, 2008). Building upon these ideas, a leader who displays positive relations-oriented behavior might show higher levels of physiological arousal because of the positive emotions associated with their own overt behavior. Damen et al. (2008) argued, along these lines, that the leader's high physiological arousal, coupled with positive affect, leads to attributions of leader charisma and effectiveness. Hence, in addition to expecting higher physiological arousal during positive relations-oriented leader behavior, if leaders pair higher physiological arousal with positive relations-oriented behavior, they might also be seen as more effective.

Related studies that used brain scanning techniques such as quantitative electroencephalography (qEEG) have reached similar insights. When a leader was perceived as transformational, and thus as showing more positive relations-oriented behavior, different regions of the leaders' brain were activated (Balthazard et al., 2012). This evidence suggests that transformational leaders can be differentiated from non-transformational leaders on the basis of stronger activation patterns in the pre-frontal and frontal lobes. On the basis of right frontal coherence, Waldman, Balthazard, and Peterson (2011) were able to differentiate highly inspirational and charismatic leaders from their less inspirational and charismatic counterparts. These studies showed that when different neurological regions were activated in the leaders, they were considered to be transformational. Ample evidence exists about the association between this leadership style and leader effectiveness (for a meta-analysis, see, Judge & Piccolo, 2004). The above offers grounds for suggesting that processes related to high physiological arousal might be activated when positive relations-oriented behavior is displayed, and leader effectiveness is relatively high.

### *Physiological arousal and negative relations-oriented leader behavior*

When psychological threats are encountered in social interactions, distinct patterns of cardiovascular responses are elicited, such as the amount of blood pumped from the heart (Blascovich & Tomaka, 1996). In those situations, higher levels of physiological arousal can be observed (van Prooijen, Ellemers, Van der Lee, & Scheepers, 2018). Such activation of higher arousal when interacting with followers is often paired with negative relations-oriented behavior, in the form of, for instance, protecting one's own position (Scheepers, 2009; Scheepers & Ellemers, 2005). Hence, feelings associated with leader-status threats could trigger higher levels of physiological arousal. Related empirical work has also established that anger is accompanied by higher physiological arousal (Berkowitz, 1990). Negative stimuli, such as pictures of negative facial affect, have been shown to lead to enhanced skin conductance responses (Vrana & Gross, 2004). Another study showed that even when participants were asked to regulate their emotions during the display of negative and neutral pictures (i.e., by distraction or reappraisal), their skin conductance did not decrease as a result (Kinner et al., 2017). This may mean that when a certain emotion is felt, for example, during negative relations-oriented leader behavior, it is difficult to 'manipulate' the co-occurring physiological responses. Furthermore, it was found that when people felt anxiety (when playing a video game), higher arousal was elicited (Ravaja, Turpeinen, Saari, Puttonen, & Keltikangas-Järvinen, 2008). One might argue, therefore, that a leader who displays negative relations-oriented behavior is experiencing higher levels of anger or anxiety, which is likely to be accompanied by relatively higher physiological arousal.

To the best of our knowledge, there is no empirical evidence available on the association between negative relations-oriented behavior, leader effectiveness, and any co-occurring of physiological arousal. However, research in the area of emotional expressiveness and emotional display (see, e.g., Ilies, Curşeu, Dimotakis, & Spitzmuller, 2013) examined both the influence of valence (i.e., positive or negative) and level of activation (i.e., passive vs. active) (Lewis, 2000). Here, it has been argued that an emotional expression is typically effective if it fits with the behavior that a person displays at the same time (also referred to as 'genuine emotion'). Observers or people in a work setting can typically distinguish between genuine and 'acted' emotions; the former behavior is typically more likely to influence other people (Ekman, 1992, 1993). Other research empirically established higher levels of leaders' physiological activation during difficult conversations (Waller et al., 2017), in which negative relations-oriented behaviors are typically displayed; Leaders who experienced these elevated levels of arousal were also found to learn more from this experience. Higher levels of physiological arousal in a work setting with followers can thus be seen as an indicator of a leader's sensitivity to social cues. The noted research implies that when leaders experience higher levels of physiological arousal during the display of negative relations-oriented behavior, they might come across as more genuine or authentic which is positively associated with leader effectiveness.

### *Physiological arousal and task-oriented leader behavior*

Compared to positive and negative relations-oriented behavior, which can be considered as highly arousing or sensitive states, task-oriented behavior can be seen as much more neutral. This is in line with the finding that people respond more neutrally (as captured by facial electromyographic activity) to neutral behavioral expressions shown in pictures (Dimberg, Thunberg, & Elmehed, 2000). Weis and Herbert (2017) extended the idea that spoken language and emotional activation are strongly intertwined (i.e., the so-called embodiment thesis of language, which indicates that language can cause emotional processes in the body and brain). They assessed physiological activation during viewing of positive, negative and neutral word-pairs displayed on a computer screen. Participants showed lower physiological responses (i.e., heart rate and skin conductance) when neutral word pairs (e.g., "no book") were processed. These results fit well with previous findings and assumptions about increased physiological responses (including skin conductance) that are typically evoked by positive or negative behavioral or emotional stimuli and not by neutral events, words or stimuli. Christopoulos et al. (2019) argued that for more resilient performers, physiological elevation quickly returns to the baseline when a decidedly positive or negative stimulus is no longer presented. This assumption is in line with the key assumptions underlying the healthy variability proposition that both physiological and behavioral responses fluctuate in social contexts.

Hence, on the basis of the above, one can expect higher arousal levels during the display of positive as well as negative relations-oriented leader behavior, and lower levels of arousal during decidedly task-oriented leader behavior. Furthermore, we expect that when these leaders shift between task- and relations-oriented behavior, according to the central premises of the healthy variability thesis, highly effective leaders are likely to fluctuate more strongly in their physiological arousal as compared to the less effective leaders, in response to the dynamics of their social environment. In other words, compared to the least effective leaders, the highly effective leaders are expected to show higher levels of physiological arousal during positive and negative relations-oriented behavior, while exhibiting lower levels of arousal during task-oriented behavior. Therefore,

**Hypothesis 1.** Positive relations-oriented leader behavior is associated with higher levels of physiological arousal for highly effective leaders, but less so for less effective leaders.

**Hypothesis 2.** Negative relations-oriented leader behavior is associated with higher levels of physiological arousal for highly effective leaders, but less so for less effective leaders.

**Hypothesis 3.** Task-oriented leader behavior is not associated with higher levels of physiological arousal for either highly effective or less effective leaders.

## Methods

### *Study design*

The present field study examines the relationship between physiological arousal (the DV), video-coded leader behavior during regular staff meetings and follower ratings of leader effectiveness (as the two IVs). A multimethod design was adopted that included three different sources of data: (1) the Empatica E4 wristband to capture physiological arousal of the leaders, (2) video coding of leader behavior, and (3) follower surveys to assess the leaders' effectiveness. The data were collected in a large public-sector organization in the Netherlands, in three of their four divisions. The study was approved by both the ethical review board of the university and the workers' council of the participating organization.

### *Participants*

A total of 101 leaders were randomly selected through stratified random sampling. After selection of these participants, they were invited to information sessions in which the design of the study and its requirements were explained in detail. The teams they led either processed financial-administrative data or were in charge of creating and/or facilitating the ICT infrastructure to increase organizational efficiency. On average, the leaders were 51.9 years old ($SD = 7.5$), had been employed in the organization for 25.1 years ($SD = 13.8$) and had worked with their team for 2.1 years ($SD = 1.5$); 60% were male. The average team was comprised of 12 followers ($SD = 5.7$).

In the 12-month data-collection part of this study, one regular staff meeting, chaired by each of the participating leaders, was video-recorded. In those periodic staff meetings, work-related topics and progress were discussed. As was the custom in this and many other organizations, most teams in the study met periodically, once or twice per month. Most of the leaders' followers, who worked together on the completion of the team's tasks, had to be present. Moreover, the video-recorded meeting had to be a randomly selected regularly held meeting that would have been held even had the video-recording apparatus not been installed in their regular meeting room. The duration of the recorded meetings ranged from 42.2 to 191.2 min ($M = 89.3$, $SD = 37.8$). Directly after each recorded meeting, each attending follower was asked to complete a survey that included ratings of leader effectiveness.

## Measures

### *Leader effectiveness*

Leader effectiveness was measured using follower perceptions. They were assessed with the four leader effectiveness items from the Multifactor Leadership Questionnaire (MLQ: 5X-Short package; Bass & Avolio, 1995). A sample item is: "This leader is effective in meeting organizational requirements." A Likert scale from 1 (*strongly disagree*) to 7 (*strongly agree*) was used. Cronbach's alpha showed good internal reliability: $\alpha = 0.90$. ICCs and $R_{wg}$s were calculated to obtain information about the within-group agreement and group reliability of the scores (i.e., indexing group-level dispersion or diversity in the scores; Newman & Sin, 2009). ICC1 was 0.22 ($p < .01$) and ICC2 was

0.85 ($p < .01$). Within-group agreement (mean $R_{wg} = 0.81$; min $R_{wg} = 0.61$; max $R_{wg} = 0.98$) also indicated that the followers agreed about the relative effectiveness of their team leader (Lance, Butts, & Michels, 2006; LeBreton & Senter, 2008).

*Physiological arousal*

Each leader's skin conductance was assessed during the entire randomly selected regular staff meeting, as a continuous physiological measure, using a special wristband biosensor. This small, unobtrusive wristband allows for the precise capturing of human physiological data. Physiological data from sensors is much more objective compared with self-reports (Blascovich, Mendes, & Seery, 2002; Cacioppo, Tassinary, & Berntson, 2016). Before the meeting started, the biosensor from Empatica (the E4, which uses 8 mm, silver-plated electrodes) was secured around the leader's non-dominant wrist by one of the field researchers. Although previous research has indicated that EDA can be assessed more reliably and validly on the hairless palm of the hand or sole of the foot (Boucsein et al., 2012; van Lier et al., 2017), we chose to use the E4 because of its low obtrusiveness in professional work settings. However, due to technical data-collection issues, valid EDA measurements were not obtained from all of the leaders. Physiological data measured with sensors in field settings tend to be precarious (Sano et al., 2018); in our study, we also encountered many problems where sensors broke down or problems occurred with sensor hardware functionality. Due to these issues, sometimes the leaders' physiological arousal was not recorded at all, or only ‘snippets’ of the meeting were captured. Due to these malfunctions, valid EDA data were obtained from only 46 of the 101 leaders.

The software program ‘Empatica Manager’ uploaded and stored the EDA data for each participating leader. These data were put in an Excel sheet that included 4 data points per second (i.e., 4 Hz). The most important phasic and tonic parameters can be derived from these raw data: Skin Conductance Responses (SCRs, i.e., the number of peaks for certain periods of time: short-term changes in phasic skin conductance activity), Amplitude (i.e., the amplitude of each SCR, with a minimum of 0.03 micro Siemens: μS) and Skin Conductance Level (i.e., SCL: the slowly varying tonic skin conductance activity).

In terms of preprocessing the EDA data, trough-to-peak analysis was conducted with Ledalab, an extension of Matlab, to arrive at the number of SCRs per minute. In addition, the SCL was calculated using Continuous Decomposition Analysis. This type of analysis reduces over-estimation of the SCL by excluding the SCRs from the average SCL (Benedek & Kaernbach, 2010); the Benedek and Kaernbach (2010) formula was used to derive a precise measure of SCL.

The EDA data were manually examined for each participant. Two of the authors independently checked the data for artifacts and non-responsiveness (i.e., flat lines). On that basis, one nonresponsive participant was excluded from the sample, resulting in 45 participants.[1]

*Leader behavior*

Regular staff meetings were videotaped to assess leaders' behavior during actual interactions with their team members. Such meetings are seen as a critical work context (Allen et al., 2015; Baran et al., 2012; Hoogeboom & Wilderom, 2015; Hoogeboom & Wilderom, 2020; Lehmann-Willenbrock et al., 2015). In staff meetings, social interaction patterns occur between leaders and followers (Heaphy & Dutton, 2008). We also checked in the survey responses whether the teams found the meeting to be representative compared to non-videotaped meetings, measured on a Likert scale from 1 to 7 ($M = 5.5$, $SD = 1.4$), whether the leader's behavior was representative of the behavior he or she normally displays ($M = 5.7$, $SD = 1.2$) and whether the team's behavior was similar to that in non-videotaped meetings ($M = 5.9$, $SD = 1.1$).

On the basis of an earlier validated 25-page codebook that was developed in a prior PhD study and later refined on the basis of existing behavioral taxonomies and team communication research (Hoogeboom & Wilderom, 2015), 19 mutually exclusive behaviors were systematically coded using specialized software (“The Observer XT;” Noldus, Trienes, Hendriksen, Jansen, & Jansen, 2000; Spiers, 2004). Based on previous research (e.g., Bass & Avolio, 1995; Behrendt et al., 2017; DeRue et al., 2011; Yukl, 2010), these 19 micro-behaviors can be grouped into 3 meta-categories of verbal behavior (Table 1): task-oriented, positive relations-oriented and negative relations-oriented behavior. Whereas we concur that a verbal statement from a leader that is intended to be task-oriented can be interpreted as relations-oriented by a follower (e.g., every task-oriented message may seem to contain a relational message, see Keyton & Beck, 2009; Watzlawick, Beavin, & Jackson, 1967), the coders were trained to code the leaders' behaviors on the basis of their “surface level communicative function” (Keyton & Beck, 2009, p. 17).

Task-oriented behaviors explicitly deal with task accomplishment (e.g., Burke et al., 2006). Leaders who initiate structure typically engage in various task-oriented behaviors to promote completion of the specific tasks at hand (i.e., by directing, giving information about the tasks and providing structure: codes 4, 5 and 6 in Table 1). Task-oriented behaviors also entail behavior where a leader describes or specifies the standards for task performance, including what would constitute ineffective performance, to make sure that followers are able to improve their task progress and are able to meet the goals (i.e., by providing negative task feedback, task monitoring and correcting: codes 1, 2 and 3 in Table 1). Task-oriented behavior was also coded when leaders were sharing a belief that is relevant for the task execution or when they aimed to create consensus about task elements (codes 7, 8 and 9 in Table 1). Such task-oriented behavior thus “facilitate understanding task requirements” (Burke et al., 2006, p. 291).

Relations-oriented leader behavior is seen as “person-focused behaviors that facilitate the behavioral interactions, cognitive structures, and attitudes that must be developed before members can work effectively as a team” (Burke et al., 2006, p. 291). This means explicitly linking oneself as a person to what one or more followers find important, in an overtly positive or negative way, thereby connecting oneself as a person (in the role of the leader) to one or more specific or identifiable others (followers). Hence, whenever something of (inter-)personal importance was stressed in a leader's message, that is, when it focused on encouraging, participating, supporting or praising, it was classified as relations-oriented. Thus, relations-oriented behavior was coded when a leader specifically addressed the development of a person or follower (i.e., person-focused communication) or offered support, that is, by raising one or more followers' expectation of being helped by the leader (code 10 in Table 1); transforming a person or follower by promoting intellectual contributions and showing that ideas from followers are also valued (code 11 in Table 1); inspiring collective efforts and stimulating engagement (code 12 in Table 1); offering praise to a follower or an explicit (personal) positive evaluation (code 13 in Table 1); using humor as an attempt to liven up the climate (code 14 in Table 1) or sharing *personal* information to bond on a personal level (code 15 in Table 1) to build cooperative relationships, based on a shared vision, that are intended to enhance the quality of socioemotional work relations. Together, these six positive relations-oriented behaviors cover the degree to which a leaders shows concern and respect for his or her followers and their needs, expresses appreciation and support, and fosters mutual trust (Fleishman, 1953; Judge et al., 2004).

With regard to the negative relations-oriented leader messages that we coded, they were either downright impolite or uncivil (code 16 and

[1] When we inspected this data visually, the descriptive plots for each participant, as well as the data overall, showed that leaders were physiologically responsive during regular staff meetings with their followers. This strengthens earlier ideas in the literature that these meetings are good contexts for examining workplace interactions between leaders and followers (e.g., Allen, Yoerger, Lehmann-Willenbrock, & Jones, 2015; Baran, Shanock, Rogelberg, & Scott, 2012; Lehmann-Willenbrock, Meinecke, Rowold, & Kauffeld, 2015; Hoogeboom & Wilderom, 2015: 2020).

**Table 1**
Definitions and examples of categories of video-coded behaviors.

| | Coded behavior | | Definition | Examples of coded behaviors |
|---|---|---|---|---|
| 1 | Providing negative task feedback | Task | Adressing discrepancies in team members' performance-goal accomplishment | "I do not think that this is a good solution"<br>"In August I sent an email with amendments, and I find it regrettable that at least half of the attendees does not know the content of this e-mail" |
| 2 | Task monitoring | Task | Asking team members for clarification and confirmation about (the progress on) their tasks | "How is the project progressing"<br>"Do you also have a specific role in that process, since there might be possibilities for a follow-up project" |
| 3 | Correcting | Task | Imposing disciplinary action; Presenting team members with a "fait accompli" | "Yes, but that is the wrong decision"<br>"Now you are talking about a failure fine, however this is a different type of fine" |
| 4 | Directing | Task | Dividing tasks among team members (without enforcing them); Determining the current direction | "John, I'd like you to take care of that"<br>"Jack, I want you to …" |
| 5 | Informing | Task | Giving factual information | "The budget for this project is…"<br>"The sick-leave figure is relatively low" |
| 6 | Structuring | Task | Structuring the meetings; Changing the topic; Shifting towards the next agenda point | "We will end this meeting at 2 pm"<br>"Maybe, we need to discuss this point after you are finished" |
| 7 | Giving own opinion | Task | Giving one's own opinion about what course of action needs to be followed for the organization, department or the team | "We already discussed this, let's talk especially about how we can avoid these things in the future"<br>"In my opinion, we should..." |
| 8 | Agreeing on task-related matters | Task | Agreeing with something; Consenting to something | "This also reflects how I personally think about the matter"<br>"Yes, I agree with you" |
| 9 | Disagreeing on task-related matters | Task | Contradicting team members | "That is not correct"<br>"I have to disagree with you on this point" |
| 10 | Individualized consideration | Positive relations | Paying attention to each individual's need for achievement and growth by acting as a coach or mentor and creating a supportive climate | "We offer a training course in August, which might be helpful for your career planning"<br>"You can make a note of that request, I am willing to help you with it" |
| 11 | Intellectual stimulation | Positive relations | Asking for ideas, stimulating team members to critically think about team tasks, opportunities and so on, including the questioning of assumptions; Thinking about old situations in new ways | "Yes, if you have any ideas put them together and discuss it with me or Jan"<br>"We will discuss how we can reduce this number together" |
| 12 | Idealized influence behavior | Positive relations | Talking about an important collective sense of vision; Talking about important values and beliefs | "I find it important that we all work in unison towards this shared objective"<br>"Until Vision 2020 is more clearly specified we will be operating under these standards; It is important to follow this agreed line" |
| 13 | Providing positive feedback | Positive relations | Positively evaluating and rewarding the behavior and actions of team members | "How you approach the project is much better than 3 months ago"<br>"I am delighted to see that you did not passively wait, but rather pro-actively came with a proposal" |
| 14 | Humor | Positive relations | Making jokes or funny statements | Often jokes are made within the context of the interaction. When 3 or more members laugh the code 'humor' is assigned |
| 15 | Giving personal information | Positive relations | Sharing personal information (e.g., about the family situation) | "We had a lovely holiday"<br>"My mother is doing better now, thank you" |
| 16 | Showing disinterest | Negative relations | Not taking any action (when expected) | Not listening actively |
| 17 | Defending one's own position | Negative relations | Emphasizing one's leadership position; Emphasizing self-importance | "I am the manager within this organization"<br>"We do it my way, because I am the manager" |
| 18 | Interrupting | Negative relations | Interfering or disturbing when other team members are talking | Disrupting other team members when they did not finish their sentence |
| 19 | Listening | Listening | Active listening | Nodding, paraphrasing |

18 in Table 1) or entailed self-promotion (code 17 in Table 1). These were behaviors in which a leader created more distance between themself and one or more followers, instead of rapproachement, closeness or collaboration.

In addition, the leader-behavioral code 'listening' was assigned when the leaders themselves were not speaking and were not 'showing disinterest.' Even though we did not assess how deeply the leaders listened to what was said by the others in the meetings, because in coding 'listening behavior' we did not distinguish if a leader was engaged in active or attentive listening (Kauffeld & Lehmann-Willenbrock, 2012; Van Quaquebeke & Felps, 2018), we think we have a fairly pure assessment of leader listening.

In order to systematically and reliably code each leader's micro-behaviors, students with a background in either Business Administration, Psychology or Communication studies were selected. Before coding the videos, the students received extensive training, especially in how to properly use the codebook and the video-coding software (Behrendt et al., 2017). Each video was coded in its entirety by two independent coders. Although earlier recommendations point to independent behavioral coding of 15–20% of the videos by two coders (Klonek, Burba, Kauffeld, & Quera, 2016), a rigorous approach was applied here for coding all video-footage (i.e., 100% of the videos), independently by two coders. They had to code the same behavior as occurring within a 2-s time frame. Coding of similar behaviors outside the 2-s time window would result in a disagreement. Overall, an inter-rater reliability of 94.35% was established (Kappa = 0.93), which is considered to be a high level of agreement (Landis & Koch, 1977). The behavioral input for the statistical analyses was the means of the coded frequencies of the behaviors (see Table 1).

*Controls*

A number of variables that were expected to be associated with the level of arousal were controlled for in the analyses. Women are often assumed to have different physiological reactions towards emotional stimuli, compared to men (Poláčková Šolcová & Lačev, 2017). On the basis of social expectations for men and woman, one might expect that females in general show more positive emotions (Fabes & Martin, 1991) and might also show higher arousal during emotion-laden behavior, such as positive or negative relations-oriented behavior. Age can also result in variations in physiological arousal, because of changes in skin thickness, skin elasticity, the number of active eccrine sweat glands and the sweat quantity per gland (Boucsein, 2012). In addition, meeting duration was included as a control variable, because habituation in physiological responses is a physiological mechanism likely to occur during any psychophysiological study (Boucsein, 2012; Figner & Murphy, 2011).

*Analysis plan*

The analysis and synchronization of the EDA and video-coded behavioral data occurred in three phases. In the first phase, the video-coded behaviors were synchronized on a mutual timeline with the EDA data. In the second phase, a Machine Learning (ML) model was used to distinguish low vs. high arousal moments. In the third phase, the associations between arousal, behaviors and leader effectiveness were examined using multi-level log-linear modeling.

*Synchronizing EDA measures and leader behavioral coding*

In order to answer our research question, the physiological recordings and the leaders' video-coded behaviors had to be synchronized. Synchronization of the EDA measures and behavioral coding was done on the basis of the internal clocks in both the EDA and video recording devices, using customized Python and Matlab code. The internal clock time in the Empatica E4 device is represented in Unix time (i.e., seconds from 1-1-1970 in Coordinated Universal Time: UTC). Unix time was converted to UTC. In addition, to ensure precise synchronization, an event marker had been placed in front of the camera by the field researcher at the start of each meeting. The time of this marker was presented in Unix time. Because the video recording device provided a time stamp at the start of each video recording, the number of seconds between the start of each recording and the event marker was calculated. We found the clock times of the Empatica E4 biosensor and those of the video recordings to be equivalent. Using customized Python code, we then synchronized the video-coded behavior with the physiological recordings.

Although several scholars have shown an average delay of 0.8 to 3.0 s between a stimulus and an event-related SCR response (e.g., Dawson, Schell, & Filion, 2007; Weis & Herbert, 2017), we chose not to control for this time-window in the data. The effect of correcting for this small time-window would have resulted in negligible differences. This is because we chose to associate the SCR's with onset and termination of broad categories of behavior (positive and negative relations- and task-oriented behavior) without claiming SCR specificity, and we are relying on a large number of data points (i.e., 20,394).

*Machine learning to assess high vs. low arousal*

Matlab and Python software were used to develop a ML model. We developed a ML classifier for binary arousal detection using the most important EDA parameters: SCR, SCL and amplitude of SCRs (see also Sano et al., 2018, for the application of Machine Learning in classifying high and low arousal on the basis of physiological data). The Random Forest (RF) model (i.e., an ensemble of decision trees) was trained with 25 estimators and evaluated using the Leave-One-participant-Out Cross-Validation (LOOCV) procedure for cross-validation. Performance was evaluated by calculating accuracy and Kappa values for each participant in the dataset.

*Ground-truth generation (high and low arousal labels)*

We defined the ground-truth generation scheme (because supervised ML methods such as RF require labeled training data to learn to differentiate between various categories) for high arousal[2] as well as low arousal, based on the mean and standard deviation of the SCL parameter (or attribute) in the dataset. Below, ***s*** represents the stress label; ***m*** and ***std*** denote the mean and standard deviation of a SCL, respectively, and ***x*** is the mean SCL of an instance (or dataset row). Then high arousal and moderate-to-low arousal labels are specified as follows:

$$s = \begin{cases} no\ stress\ (0), if\ x<m-std \\ stress\ (1), if\ x>m+std \end{cases} \quad (1)$$

An additional 9 participants were discarded because of not having a ground-truth for training the ML model, resulting in a total sample size of 36. The ML models can be evaluated in several different ways, depending on how the problem is specified. Some widely used methods are: stratified cross-validation and randomly splitting data into a training, validation and testing set (e.g., Flach, 2012). We used the LOOCV. With this method, the ML model is trained with all data except the data of one participant. Subsequently, the model is tested against the left-out participant's data. This process is repeated for every participant and the performance metrics are calculated on the validation set. Compared to standard K-fold validation (i.e., randomly splitting the data into training and testing folds), LOOCV reflects model performance better because, during each training cycle, the classifier does not learn from the data of the 'left-out' participant. The model's performance (such as accuracy) on 'left-out' participants is used to validate the model and averaged to get the overall model performance.

*Evaluation metrics*

We evaluated the classifier's performance by using two widely-used metrics, namely accuracy and Cohen's kappa. Brief descriptions of both metrics are given below.

*Accuracy* is expressed as the ratio of the number of correct (or actual) true labels out of all the predictions made by the classifier. Accuracy is the most widely used metric for evaluating the classification performance of ML models (e.g., Flach, 2012). However, it is sometimes also misused and is only suitable when there is an equal number of cases in each class in the dataset or when the dataset is balanced (i.e., when each case has an approximately equal representation). It can be calculated as follows:

$$Accuracy = \frac{Total\ correct\ predictions\ by\ the\ classifier}{Total\ number\ of\ observations}$$

*Cohen's kappa* is a measure of overall agreement between two raters. It classifies items into a given set of ***k*** categories. The formula for kappa is given below, where $\boldsymbol{p_{ii}}$ is the proportion of examples that both raters classify into category ***i***, $\boldsymbol{p_{i+}}$ is the proportion of examples that rater *A* assigns to category ***I*** and $\boldsymbol{p_{+i}}$ is the proportion assigned to category ***i*** by rater ***B***. The denominator is then used as a normalizing factor to make the kappa value (*K*) equal to **1**. A kappa statistic can have a minimum value of **−1**, in case of complete disagreement, and a maximum of **1**, for perfect agreement.

$$K = \frac{\sum p_{ii} - \sum p_{+i} {}^* p_{+i}}{1 - \sum p_{i+} {}^* p_{+i}} \quad (2)$$

[2] It should be noted that in this specific workplace setting, only moderately high levels of arousal are to be expected (see also, e.g., Coughlin, Reimer, & Mehler (2009), who visualized how arousal is associated with performance).

### *Random forest*

RF is an ensemble learning algorithm that generates multiple decision trees, which allows for precise classification of physiological data. The ensemble is a 'divide and conquer technique' that is used to improve the performance of the classification system. The key idea behind this method is that, together, a group of weak learners can produce a strong learner (e.g., Flach, 2012). RF generates many different decision trees. Each decision tree gives a classification or 'tree vote' for the particular class; on the basis of this, the algorithm then selects the classification with the most votes. In contrast to traditional decision trees, which are more likely to suffer from high variance or bias, RF uses the average to find the natural balance between the two extremes. For a detailed description of the RF algorithm, see Breiman (2001).

The ML model generated a mean Cohen's kappa of 0.38 for all participants (mean accuracy = 0.73). According to Landis and Koch (1977, p. 165), this could be termed as "fair agreement." Similar ML studies have found comparable kappas and accuracy. The kappa for each participant provides information about how well the ML model can predict high and low arousal for that specific participant.

### *Multi-level log-linear modeling to test the hypotheses*

To examine the associations between leader arousal, behavior and effectiveness, multi-level log-linear modeling was employed, using the open source platform *R*, while controlling for gender, age and meeting duration. Because the behavioral events are nested at the individual leader level, a multi-level three-way log-linear model was used. Assumptions were checked before conducting the analyses. The residuals were normally distributed and the variance was homogenous across the fitted data. To ensure that the multilevel random-effects model was tenable, a Hausman Test was employed (Antonakis, Bendahan, Jacquart, & Lalive, 2010; McNeish & Kelley, 2019). This test (Hausman, 1978) checks whether the estimator is consistent. The Hausman statistic provides information about the chi-square value (Antonakis et al., 2010; Hausman, 1978). The non-significant chi-square result ($\chi^2 = 7.67$, $df = 6$, $p = .26$) shows a lack of endogeneity, which supports the use of a log linear multi-level model; including group means as level-2 predictors (i.e., following the Mundlak procedure: Antonakis et al., 2010) was therefore not required as a correction for endogeneity issues. In the next section, we report the estimates of the multi-level log-linear model used.

**Table 2**
Parameter estimates for the selected log-linear model: leader arousal proportions per behavior and leader effectiveness[a].

| Behavior | Leader effectiveness | Physiological arousal: *Low arousal* | *High arousal* |
|---|---|---|---|
| Listening | Highly effective | .79 (1436) | .21 (378) |
| | Less effective | .79 (1189) | .21 (319) |
| Task-oriented | Highly effective | .80 (2116) | .20 (517) |
| | Less effective | .79 (1675) | .21 (436) |
| Relation-oriented[a] | Highly effective | .69 (406) | .32 (187) |
| | Less effective | .79 (320) | .21 (85) |
| Counterproductive[a] | Highly effective | .57 (26) | .43 (20) |
| | Less effective | .88 (56) | .12 (8) |

*Note.* Table entries are row-conditional; the sum is 1.0 across rows. Frequency counts are shown in parentheses. The two groups of leaders were formed on the basis of a median split: highly effective leaders ($n = 18$) vs. less effective leaders ($n = 18$). Although median splits have been heavily criticized, as they increase the chance of producing Type I errors and reduce statistical power (e.g., McClelland, Lynch, Irwin, Spiller, & Fitzsimons, 2015), use of a median split in our data is not likely to result in such an error, as it did not suffer from multicollinearity.

[a] Significant difference between the probabilities of high and low arousal for highly effective and less effective leaders for this behavior on the basis of a chi-square test (2-tailed).

## Results

Table 2 depicts both the probabilities and absolute counts of the leaders' behavioral frequencies in relation to leaders' level of arousal, for both the highly effective and less effective leaders. In doing so, two groups of leaders were formed on the basis of a median split (5.71 on a scale of 1–7): highly effective leaders ($n = 18$) vs. less effective leaders ($n = 18$). The probabilities are row-conditional and show that the highly effective leaders displayed higher arousal during both positive relations-oriented ($\chi^2(1) = 13.50$, $p < .001$) and negative relations-oriented behavior ($\chi^2(1) = 13.54$, $p < .001$). The results indicate that high arousal was experienced during positive relations-oriented behavior: this occurred 32% of the time for the highly effective leaders versus 21% for the less effective leaders. This significant difference is even more apparent with negative relations-oriented behavior, where high arousal was felt 43% of the time by highly effective leaders versus 12% of the time by less effective leaders. Hence, highly effective leaders are more likely to experience higher arousal when they display positive or negative relations-oriented behavior. The results also show that compared with the less effective leaders, the highly effective leaders displayed positive relations-oriented behavior more frequently and negative relations-oriented behavior less often. In addition, Table 2 shows that during the display of listening and task-oriented behavior, both highly and less effective leaders were physiologically less aroused (i.e., higher probabilities of low physiological arousal). Table 3 presents the means, standard deviations and intercorrelations between the studied variables. A significant negative association between physiological arousal and meeting duration ($r = -0.38$, $p < .05$) shows that when the meetings lasted longer, fewer moments of high arousal were noticeable.

The hypotheses were tested with the multi-level log linear regression model (see Table 4), which substantiated the preceding results. First of all, when inspecting the results for relations- and task-oriented leader behavior, the results indicate that higher levels of arousal were shown during positive ($\gamma = 1.20$, $p < .001$) and negative relations-oriented behavior ($\gamma = 2.49$, $p < .01$). This physiological-behavioral association was not established for task-oriented leader behaviors, that is, during the display of task-oriented behavior no elevated levels of physiological arousal were experienced ($\gamma = -0.14$, n.s.). Support was found for hypothesis 1, such that high arousal during the display of relations-oriented behavior was experienced significantly more frequently by the highly effective leaders ($\gamma = 0.58$, $p < .01$: see Fig. 1 for the interaction). In support of hypothesis 2, highly effective leaders experienced higher levels of arousal more frequently as compared to less effective leaders during negative-relations oriented behavior ($\gamma = 1.53$, $p < .01$: see Fig. 2 for the interaction). As hypothesized, no association between higher levels of physiological arousal and task-oriented behavior was found for leaders who either scored high or low on effectiveness (hypothesis 3: $\gamma = 0.07$, n.s.).

## Discussion

With this field study, we tested novel hypotheses at the intersection of research on human physiology and organizational leadership. The goal of this study was to examine the associations between leaders' behaviors, their degree of physiological arousal and their effectiveness. Despite the complexity of coupling real-life leader behaviors with skin conductance data (e.g., Arvey & Zhang, 2015), we showed how arousal in the sympathetic branch of the autonomic nervous system may accompany three categories of leader behaviors (positive and negative relations-oriented and task-oriented behavior). Specifically, higher physiological arousal was found when leaders relayed verbal messages of a positive and negative relations-oriented nature to their group of followers. Leaders' task-oriented behaviors were accompanied by low physiological arousal and lacked a significant association with leader effectiveness. Moreover, the highly effective leaders especially exhibited higher physiological arousal during both positive or negative relations-oriented behaviors.

**Table 3**
Means, standard deviations and intercorrelations of study variables.

| | | *M* | *SD* | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1. | Physiological arousal | .25 | .17 | | | | | | | |
| 2. | Leader effectiveness | 1.53 | .51 | .02 | | | | | | |
| 3. | Gender | 1.25 | .44 | −.16 | −.14 | | | | | |
| 4. | Age | 52.61 | 6.85 | −.06 | .15 | −.47** | | | | |
| 5. | Meeting duration | 92.10 | 35.94 | −.38* | −.18 | −.08 | .19 | | | |
| 6. | Task-oriented behavior | 50.68 | 10.49 | .03 | −.08 | .04 | .23 | −.09 | | |
| 7. | Positive relations-oriented behavior | 12.51 | 9.27 | .24 | .36* | −.11 | .04 | −.05 | −.47** | |
| 8. | Negative relations-oriented behavior | 2.44 | 8.60 | −.05 | −.02 | −.34 | −.19 | −.01 | −.10 | −.14 |

*Note.* $n = 36$.
Physiological arousal was classified as 0 (*low arousal*) and 1 (*high arousal*).
Leader effectiveness was classified as 1 (*less effective*) and 2 (*highly effective*).
Gender was coded 1 (*male*) and 2 (*female*).
Meeting duration was measured in minutes.
Task- and relations-oriented behaviors are represented by the relative average frequency per meeting.
* $p < .05$ (2-tailed).
** $p < .01$ (2-tailed).

Thus, among a group of organizational leaders in action, we offer evidence supporting the healthy variability thesis: although a higher level of physiological arousal was associated specifically with positive and negative relations-oriented behavior, the most effective leaders especially showed significantly stronger fluctuations in their physiological arousal when they engaged in relations-oriented behavior.

## *Theoretical implications*

The results have at least three theoretical implications. First, the results from this study extend prior leadership research because we identify a physiological or biological correlate of solidly measured relations-oriented leader behavior. By doing so, we provide an initial answer to the question posed by Heaphy and Dutton (2008, p. 138): "What do we know about the physiological correlates of […] social interactions at work?" Hence, we addressed the similar call to "connect biological materials to behavior and especially to organizational behavior" (Carroll & O'Connor, 2014, p. 16). Hence, although previous studies have focused on biological correlates, that is, on the effect of leader behavior on the physical health of followers (see, e.g., Diebig, Bormann, & Rowold, 2016), there is no prior theory-guided evidence that combines leaders' physiological markers with their own concurrent behavior over time. By bridging the fields of leadership studies and physiology, and extending that empirical work beyond followers' stress and/or physical health (White, Thornhill, & Hampson, 2006), we contribute to the development of biological theories of effective leadership. Such integrative theory can advance our understanding of the micro-processes of effective leadership behavior, thereby coming closer to the reality of (effective) workplace dynamics (e.g., Carroll & O'Connor, 2014). The results reveal that the leaders' physiological responses, as reflected in the activity of the autonomous nervous system (e.g., Golland et al., 2014), co-occur with organizationally relevant, observable behaviors that fulfill different functions (Morgeson et al., 2010). By demonstrating that leaders are significantly more aroused when displaying positive and negative relations-oriented behavior compared to their arousal level during task-oriented behaviors, we add a physiological marker to a functional set of key leader behaviors. Our results underpin the idea that task- and relations-oriented leader behaviors can be separated on the basis of their different functions (Behrendt et al., 2017); lower levels of arousal accompany the behaviors aimed at ensuring task accomplishment and efficiency, whereas different physiological processes are shown to accompany relations-oriented behaviors aimed at supporting followers in a given social setting. Even though there is some debate on the usefulness and content of this behavioral distinction (Keyton & Beck, 2009; Watzlawick et al., 1967), our findings offer support for the idea that the task/relations distinction, which has been prevalent in leadership research for many decades, has a distinctive physiological arousal correlate.

Second, to the best of our knowledge, this is the first study that has yielded an integrative perspective on the co-occurrence of two important workplace phenomena, leaders' physiological arousal and their behaviors. In adopting such a perspective, we take into account the

**Table 4**
Multi-level log-linear results of regression of leader behavior and effectiveness on physiological arousal.

| | Physiological arousal | | |
|---|---|---|---|
| Parameter | γ | SE | 95%CI |
| Intercept | −.26 *** | 1.05 | (−2.317 to 1.806) |
| Gender | −.44 | .24 | (−.901 to .029) |
| Age | .00 | .02 | (−.029 to .034) |
| Meeting duration | −.01 * | .00 | (−.011 to −.000) |
| Leader effectiveness | −.01 | .21 | (−.401 to .405) |
| Task-oriented behavior | −.14 | .18 | (−.496 to .225) |
| Positive relations-oriented behavior | 1.20 *** | .27 | (.668 to 1.734) |
| Negative relations-oriented behavior | 2.49 ** | .79 | (.951 to 4.035) |
| Leader effectiveness * Task-oriented behavior | −.07 | .12 | (−.169 to .308) |
| Leader effectiveness * Positive relations-oriented behavior | .58 ** | .19 | (−.944 to −.214) |
| Leader effectiveness * Negative relations-oriented behavior | 1.53 ** | .54 | (−2.586 to −.486) |

*Note.* $n = 36$ (9174 data points in total, nested in 36 leaders).
CI = Confidence Interval.
* $p < .05$ (2-tailed).
** $p < .01$ (2-tailed).
*** $p < .001$ (2-tailed).

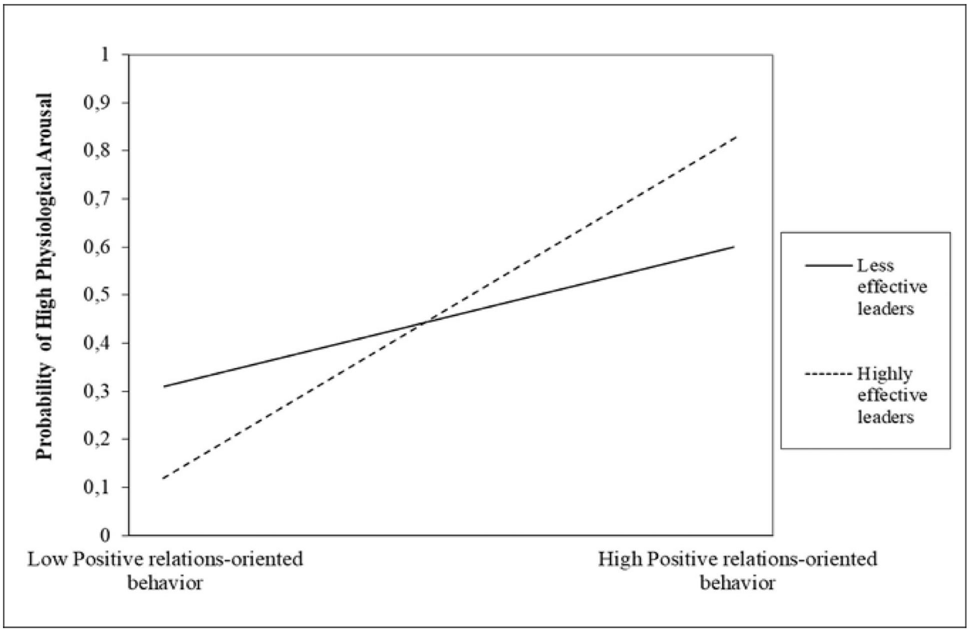


**Fig. 1.** Interaction between leaders' physiological arousal, leaders' positive relations-oriented behavior and leader effectiveness.

fluctuation of multiple variables, which permits an examination of bodily-behavioral correspondence and allows a test of the healthy variability thesis as imported to the field of Leadership. By having also included a measure of leader effectiveness, we furthermore extend prior evidence for the healthy variability thesis in OB research. Our results show that the highly effective leaders experienced higher levels of arousal more often during the display of both negative and positive relations-oriented behavior. Rather than showing a stable pattern of physiological arousal, meaningful elevation was shown during a corresponding set of nine mutually exclusive, observable, relations-oriented micro-behaviors; as hypothesized, this co-occurrence appears to be significantly associated with leader effectiveness. In doing so, we also extend prior empirical work that had already established a significant association of positive relations-oriented (including transformational) behavior with leader effectiveness (for meta-analytic evidence see, e.g., Judge & Piccolo, 2004) by adding a physiological correlate to the mix. Further distinguishing positive from negative relations-oriented behavior allowed us to examine if both relations-oriented behaviors are similarly associated with higher levels of physiological arousal. In comparison with less effective leaders, the highly effective leaders displayed negative relations-oriented behavior less often, but when they showed such behavior, it was more often accompanied by high physiological arousal. The present study is the first to examine the co-occurrence of physiological arousal and different types of leader behavior, including its association with effectiveness. Thus, the dynamic microanalyses conducted offer insight into the fine-grained real-life relations-oriented behaviors shown by leaders who are highly effective.

Third, various scholars have argued that OB scholars commonly neglect “when” a certain phenomenon of interest appears. Our application

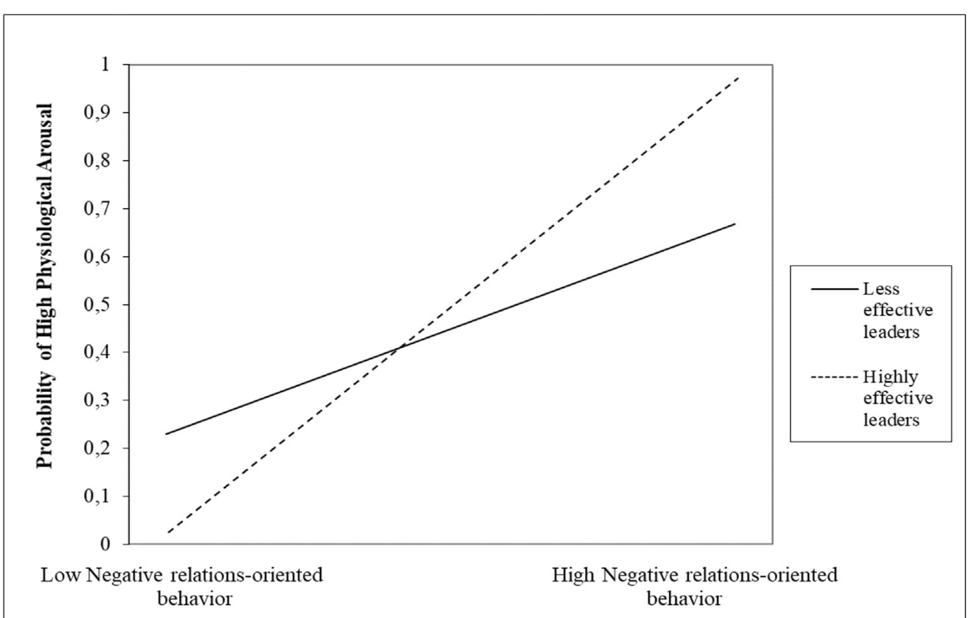


**Fig. 2.** Interaction between leaders' physiological arousal, leaders' negative relations-oriented behavior and leader effectiveness.

of the variability thesis to leadership research illuminates the importance of this gripe (Navarro & Rueff-Lopes, 2015). By pairing a micro-behavioral approach with measuring the continuous physiological data of leaders at the intra-individual level, we are able to answer an essential "when" type of question (see also Endedijk, Hoogeboom, Groenier, de Laat, & Van Sas, 2018). Highly effective leaders are especially likely to show variability in their physiological response "when" they display (both positive and negative) relations-oriented behaviors. Thus, more generically, we show that our dynamic approach, including the capturing of behavioral acts over time, uncovers an important body-behavioral correlate of organizational leadership effectiveness. We expose thereby a more complete picture that comes much closer to the reality of social interactions at work (e.g., McClean et al., 2019). Our study highlights that applying non-linear types of techniques to time-stamped data enriches our understanding of effective leadership within organizations (see also Hoogeboom & Wilderom, 2020). The variability of important entangled phenomena, such as physiological arousal and precise leader micro-behavior, should become a more regular part of future leadership research.

*Strengths, limitations and future research*

The study reported herein is an initial empirical inquiry into how leader physiological arousal is associated with different kinds of leader behaviors in the organizational field. Having collected the physiological data of leaders at work, in combination with minutely-coded, mutually exclusive leader behaviors, is a great strength of this study. Although our sample is somewhat small, other leadership studies that have used neuroscience methods (e.g., Balthazard et al., 2012; Boyatzis et al., 2012; Waldman et al., 2011) have worked with similar sample sizes (Button et al., 2013). Combining video-coded leader behavioral and physiological data is labor-intensive, but allows for detailed investigation of the biological correlates of different leader behaviors, and renders visible some aspects of the otherwise invisible. The research results of this cross-sectional field study, which do not include any common-method bias, show that it is worthwhile to commit resources to similar, fairly nonobtrusive, field research.

Prior studies that have linked physiological reactions to 'felt' emotions have mostly relied upon recalled reflections (see, e.g., Boyatzis et al., 2012). Merely recalling experiences to identify a physiological reaction is problematic, because it is often difficult for individuals to go back and relive that particular emotional state (Mauss & Robinson, 2009). Instead of relying on potentially biased memory recall, we linked actual, in situ leader behavioral displays to physiological arousal at the same moments. Some scholars have already highlighted the positive aspects of obtaining valid physiological measures during real-life interactions. van Prooijen et al. (2018), for example, argued that capturing physiological data during team interactions can show precisely how someone is feeling, "thereby avoiding the possibly biased evaluation of situations in anticipation or in hindsight" (p. 75). Relying on real-time field data from actual leaders in interaction with their followers is more representative than the convenience samples of students in leadership roles, or employing only surveys. Even when EDA data are collected in a laboratory setting, it results in an artificial representation of real workplace behavior. Although collecting EDA data in an organizational field setting is a strength of the study, it also posed several challenges.

First, disruptions of the skin-electrode interface could have been due to the fact that the participants wore the Empatica E4 under their clothes or moved their arms during the meeting (i.e., gross body movements). This affects the stability of the E4. In addition to this practical issue, technical disruptions can be caused by loose electrodes. To minimize such disruptions, the device was checked by a field researcher before the session began. All EDA recordings were also visually inspected to check for segments containing artifacts (Boucsein, 2012). The limitations of collecting data in the field may be outweighed by several advantages; namely, using a wristband (similar to a regular watch) allows for fairly unobtrusive measurement of physiological measures in the workplace, which enables parallel, fine-grained analyses of video-based behavioral data and moment-to-moment physiological responses.

Second, physiological data collection in a field setting poses another limitation. Outside of a laboratory, there are fewer chances to control for a variety of sources of physiological responses, such as recall of emotionally salient events, which can also activate higher levels of arousal (D'Esposito, Zarahn, & Aguirre, 1999). It has been empirically established that EDA activity can also show fluctuations caused by factors such as mental effort, cognitive load, room temperature, general arousal and even body posture (Berntson & Cacioppo, 2000; Stemmler, 2004). Hence, quasi-experimental field studies are especially called for; they would enable the study of causes of leader arousal–behavior linkages.

Third, using the physiological tonic and phasic parameters as dichotomous measures of arousal also means that other dynamic physiological signals have not yet been taken into consideration, such as latency, rise time or recovery time (Boucsein, 2012). Moreover, by using skin conductance as a physiological measure of arousal, we did not empirically take into account the valence of emotions. Distinguishing between different emotional states, such as happiness or fear, on the basis of neural information is difficult (e.g., Lindebaum & Jordan, 2014). Combination with other methods, such as self-reports or non-verbal coding, is then needed to capture such valences. In the present study, leader behaviors and physiological arousal, together with leader effectiveness, were examined without taking any separate measures of leaders' emotional states. One could argue that over and above the bodily response and visible behavior, other factors such as intellectual and social-emotional leader skills can determine the effectiveness of leaders. Thus, future research could be extended with discrete perceptual measures of leaders' (and even followers') emotions or their levels of emotional intelligence (see, e.g., Baker, 2019; Gooty, Connelly, Griffith, & Gupta, 2010). Thus, in such new research, the valence of the emotions and/or the degree of leaders' social-emotional skills, that is, beyond mere arousal fluctuation, could thus be included. In defense of our focus on a core affect variable (i.e., physiological arousal), recent work has challenged the focus of emotion research on discrete categories in relation to coupling such categories to unique and specific (neuro)physiological states (e.g., Siegel et al., 2018). Hence, we theorize that core affect, which can be defined as a basic neurophysiological building block of emotion that can also be experienced as a basic feeling, still remains a viable candidate for finding meaningful and robust associations between behavior, experience and physiology in leadership research.

Fourth, we did not take into account in this study how leaders' behavioral displays may affect followers' physiological processes (Tee, Ashkanasy, & Paulsen, 2011), including their physiological and/or mental health. Some OB colleagues have already argued that to understand leader effectiveness fully, neural or physiological data must be obtained from multiple members of teams (Waldman et al., 2011). Even though it is quite labor-intensive, this type of more comprehensive, triangulated precise behavioral research is indeed urgently needed, so that more evidence-based leadership enrichment interventions can be designed, tested and used, for more sensible organizational leadership and followership.

Fifth, while we controlled for the duration of the meeting, as this might affect the level of arousal from the participants (i.e., so-called habituation effects: Boucsein, 2012; Figner & Murphy, 2011), we did not account for inter-session variability of the effects found. For example, the meeting can be considered as an extended period in which behavior and arousal can vary over time. It could be that the co-occurrence of higher arousal with relations-oriented behavior might be significantly associated with higher effectiveness especially at the beginning and end of a meeting. This is in line with increasing calls from OB scholars for more attention to the temporal dimension of 'when' specific phenomena occur (Cole, Shipp, & Taylor, 2016; Jiang, Yin, & Liu, 2020). Future research should consider if co-occurrence of higher arousal with relations-oriented behavior is more influential during specific moments in a staff meeting.

*Practical implications*

Past organizational leadership research has already shown the importance of both positive relations-oriented and task-oriented behaviors as predictors of high leader effectiveness. The research presented here shows a particular biological correlate of different types of leader behaviors (we differentiated here between positive and negative relations-oriented leader behaviors) and leader effectiveness. The highly effective leaders were especially found to have higher levels of physiological arousal during instances of both positive and negative relations-oriented behavior, but not during their task behavior. Compared to the less effective leaders, the highly effective leaders displayed those negative relations-oriented behaviors less frequently.

Although experiencing higher arousal is a physical process that cannot be controlled by the individual (i.e., individuals cannot deliberately increase or decrease their physiological responses: see, e.g., Kinner et al., 2017), our results might raise awareness about the importance of experiencing higher levels of activation during the display of relations-oriented behavior at work. The joint body-behavior co-occurrences for highly effective leaders during relational-type expressions in social work settings can be interpreted as a higher level of sensitivity shown to social cues. The co-occurrence found of high levels of physiological arousal and positive and negative relations-oriented behavior is suggested as a key element in the social interactions of highly effective leaders. Without denying the importance of the typically more neutral task-type leader behaviors at work (Hoogeboom & Wilderom, 2020), our study suggests that if leaders can become more aware of their bodily expressions, they may become better at expressing relations-oriented behavior. Hence, a practical recommendation coming out of this research, for leaders who feel they lack relational types of behaviors vis-à-vis their followers and wish to become more effective, is to learn how to show their staff (now and then) how they feel about the aspects of the work they supervise that are not directly task-related. Behavioral items 10 through 18 in Table 1 might help them to direct such expressions, but they must feel strongly about such chosen aspects of their work; otherwise their high physiological arousal and their chosen verbalization will never co-occur, and the desired effect will not be achieved.

There are still many related questions that need to be addressed further before more practical application of this study's results can yield significant leader effectiveness enhancements (Lindebaum, Al-Amoudi, & Brown, 2018; Waller et al., 2017). Even though the physiological results of this study were fed back to the participating leaders, more efficient ways of precise, customized leader on-the-job learning could be created. For this purpose, greater technological feedback sophistication would need to be developed. Such easier-to-offer customized feedback must provide leaders with the possibility of physiology-based learning from their own actual work experiences and interactions. Mintzberg (2004, p. 24) called such learning from actual work experiences (other than, for example, role play or generic training) “3rd generation management development.” Once physiological and behavioral leader feedback have been coupled more efficiently, more effective leader coaching is likely to take place. This can potentially advance leadership effectiveness in practice.

## Conclusion

Although an increasingly greater group of scholars has advocated conducting research at the intersection of physiology and leader behavior, empirical developments in understanding the physiological correlates of leader behavior have not up to now made a large, visible impact in OB and leadership research and theory development (Carroll & O'Connor, 2014). This empirical field study is an initial attempt to extend our understanding of physiological correlates of effective leader behavior during workplace interactions. To precisely capture simultaneous behavioral and physiological phenomena, we used *two* different "high-resolution methods" (Klonek, Gerpott, Lehmann-Willenbrock, & Parker, 2019, p. 1). Leaders were asked to wear wrist sensors measuring electrodermal variability during periodic staff meetings and the data were combined with fine-grained, video-based coding of verbal exchanges occurring in the meetings and with leaders' effectiveness ratings. As hypothesized, the highly effective leaders showed more variability in their physiological arousal than the less effective leaders, as evidenced by higher levels of arousal during relations-oriented behavior. Thus, among the highly effective leaders, it was not the task-oriented verbalizations, but rather the relations-oriented leader behaviors that were found to be accompanied by higher levels of physiological arousal. The less effective leaders more frequently experienced lower physiological arousal during both relations- and task-oriented behavior. This study's test of the healthy variability thesis with a sample of leaders operating in their regular work context calls for more integrative analyses or triangulation of physiological, behavioral and effectiveness data in real-life organizational contexts (e.g., Becker & Menges, 2013). Many more such pairings of physiological arousal with theoretically valid micro-behavioral observations in the future are likely to result in new insights about effective organizational leadership.

## Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Acknowledgements

We would like to thank our associate editor, dr. Seth Spain, for the guidance in the review process, and thank the three anonymous reviewers for their valuable feedback.

## References

Akinola, M. (2010). Measuring the pulse of an organization: Integrating physiological measures into the organizational scholar's toolbox. *Research in Organizational Behavior*, *30*(1), 203–223. https://doi.org/10.1016/j.riob.2010.09.003.

Allen, J. A., Yoerger, M. A., Lehmann-Willenbrock, N., & Jones, J. (2015). Would you please stop that!? The relationship between counterproductive meeting behaviors, employee voice, and trust. *Journal of Management Development*, *34*(10), 1272–1287. https://doi.org/10.1108/JMD-02-2015-0032.

Antonakis, J., Ashkanasy, N. M., & Dasborough, M. T. (2009). Does leadership need emotional intelligence? *The Leadership Quarterly*, *20*(2), 247–261. https://doi.org/10.1016/j.leaqua.2009.01.006.

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly*, *21*(6), 1086–1120. https://doi.org/10.1016/j.leaqua.2010.10.010.

Arrieta, C., Navarro, J., & Vicente, S. (2008). Variables related to the emergence of differential patterns in work motivation. *Psicothema*, *20*(4), 745–752.

Arvey, R. D., & Zhang, Z. (2015). Biological factors in organizational behavior and I/o psychology: An introduction to the special section. *Applied Psychology*, *64*(2), 281–285. https://doi.org/10.1111/apps.12044.

Ashby, R. W. (1956). *An introduction to cybernetics.* London, UK: Methuen.

Ashkanasy, N. M., Becker, W. J., & Waldman, D. A. (2014). Neuroscience and organizational behavior: Avoiding both neuro-euphoria and neuro-phobia. *Journal of Organizational Behavior*, *35*(7), 909–919. https://doi.org/10.1002/job.1952.

Baker, W. E. (2019). Emotional energy, relational energy, and organizational energy: Toward a multilevel model. *Annual Review of Organizational Psychology and Organizational Behavior 2019*, *6*(1), 373–395. https://doi.org/10.1146/annurev-orgpsych-012218-015047.

Bales, R. F. (1950). A set of categories for the analysis of small group interaction. *American Sociological Review*, *15*(2), 257–263.

Balthazard, P. A., Waldman, D. A., Thatcher, R. W., & Hannah, S. T. (2012). Differentiating transformational and non-transformational leaders on the basis of neurological imaging. *The Leadership Quarterly*, *23*(2), 244–258. https://doi.org/10.1016/j.leaqua.2011.08.002.

Baran, B. E., Shanock, L. R., Rogelberg, S. G., & Scott, C. W. (2012). Leading group meetings: Supervisors' actions, employee behaviors, and upward perceptions. *Small Group Research*, *43*(3), 330–355. https://doi.org/10.1016/j.jesp.2004.06.002.

Bass, B. M. (1985). *Leadership and performance beyond expectations*. New York, NY: Free Press.

Bass, B. M. (1990). From transactional to transformational leadership: Learning to share the vision. *Organisational Dynamics*, *18*(3), 19–31. https://doi.org/10.1016/0090-2616(90)90061-S.

Bass, B. M., & Avolio, B. J. (1995). *MLQ multifactor leadership questionnaire for research: Permission set*. Redwood City, CA: Mindgarden.

Bass, B. M., & Bass, R. (2008). *The Bass handbook of leadership: Theory, research, and managerial applications* (4th ed.). New York, NY: Free Press.

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on Psychological Science*, *2*(4), 396–403. https://doi.org/10.1111/j.1745-6916.2007.00051.x.

Bechara, A., Damasio, H., Damasio, A. R., & Lee, G. P. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *19*(13), 5473–5481. https://doi.org/10.1523/JNEUROSCI.19-13-05473.1999.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (2005). The Iowa gambling task and the somatic marker hypothesis: Some questions and answers. *Trends in Cognitive Sciences*, *9*(4), 159–162. https://doi.org/10.1016/j.tics.2005.02.002.

Becker, W. J., & Menges, J. I. (2013). Biological implicit measures in HRM and OB: A question of how not if. *Human Resource Management Review*, *23*(3), 219–228. https://doi.org/10.1016/j.hrmr.2012.12.003.

Behrendt, P., Matz, S., & Göritz, A. S. (2017). An integrative model of leadership behavior. *The Leadership Quarterly*, *28*(1), 229–244. https://doi.org/10.1016/j.leaqua.2016.08.002.

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, *190*(1), 80–91. https://doi.org/10.1016/j.jneumeth.2010.04.028.

Berkowitz, L. (1990). On the formation and regulation of anger and aggression: A cognitive-neoassociationistic analysis. *The American Psychologist*, *45*(4), 494–503. https://doi.org/10.1037//0003-066X.45.4.494.

Berntson, G. G., & Cacioppo, J. T. (2000). From homeostasis to allodynamic regulation. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 459–481) (2nd ed.). New York, NY: Cambridge University Press.

Blascovich, J., Mendes, W. B., & Seery, M. D. (2002). Intergroup encounters and threat: A multi-method approach. In D. M. Mackie, & E. R. Smith (Eds.), *From prejudice to intergroup emotions: Differentiated reactions to social groups* (pp. 89–109). New York, NY: Psychology Press.

Blascovich, J., & Tomaka, J. (1996). The biopsychosocial model of arousal regulation. In M. Zanna (Ed.), *Advances in experimental social psychology* (pp. 1–51). New York, NY: Academic Press.

Bormann, K. C., & Rowold, J. (2016). Ethical leadership's potential and boundaries in organizational change: A moderated mediation model of employee silence. *German Journal of Human Resource Management*, *30*(4), 225–245. https://doi.org/10.1177/2397002216649855.

Boucsein, W. (2012). *Electrodermal activity* (2nd ed.). New York, NY: Springer.

Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., Roth, W. T., Dawson, M. E., & Filion, D. L. (2012). Publication recommendations for electrodermal measurements. *Psychophysiology*, *49*(8), 1017–1034. https://doi.org/10.1111/j.1469-8986.2012.01384.x.

Boyatzis, R. E., Passarelli, A. M., Koenig, K., Lowe, M., Mathew, B., Stoller, J. K., & Phillips, M. (2012). Examination of the neural substrates activated in memories of experiences with resonant and dissonant leaders. *The Leadership Quarterly*, *23*(2), 259–272. https://doi.org/10.1016/j.leaqua.2011.08.003.

Boyatzis, R. E., Rochford, K. E., & Taylor, S. (2015). The role of the positive and negative emotional attractors in vision and shared vision: toward effective leadership, relationships and engagement. *Frontiers in Psychology*, *6*, 670–683. https://doi.org/10.3389/fpsyg.2015.00670.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324.

Burke, C. S., Stagl, K. C., Klein, C., Goodwin, G. F., Salas, E., & Halpin, S. M. (2006). What type of leadership behaviors are functional in teams? A meta-analysis. *The Leadership Quarterly*, *17*(3), 288–307. https://doi.org/10.1016/j.leaqua.2006.02.007.

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafo, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Review Neuroscience*, *14*(5), 365–376. https://doi.org/10.1038/nrn3502.

Cacioppo, J. T., & Tassinary, L. G. (2000). *Handbook of psychophysiology*. New York, NY: Cambridge University Press.

Cacioppo, J. T., Tassinary, L. G., Berntson, G. G., Cacioppo, J. T., Tassinary, L. G., & Berntson, G. G. E. (2016). Strong inference in psychophysiological science. *Handbook of Psychophysiology* (pp. 3–15). Cambridge, United Kingdom: Cambridge University Press.

Carroll, G., & O'Connor, K. (2014). Biology, evolution, and organizations: Promises and challenges in building the foundations. In S. Collarelli, & R. Arvey (Eds.), *The Biological Foundations of Organizational Behavior* (pp. 311–342). Chicago, IL: University of Chicago Press.

Christopoulos, G. I., Uy, M. A., & Yap, W. J. (2019). The body and the brain: measuring skin conductance responses to understand the emotional experience. *Organizational Research Methods*, *22*(1), 1–27. https://doi.org/10.1177/1094428116681073.

Cianci, A., Klein, H., & Seijts, G. (2010). The effect of negative feedback on tension and subsequent performance: the main and interactive effects of goal content and conscientiousness. *The Journal of Applied Psychology*, *95*(4), 618–630. https://doi.org/10.1037/a0019130.

Colarelli, S., & Arvey, R. (2015). *The biological foundations of organizational behavior*. Chicago: University of Chicago Press.

Cole, M. S., Shipp, A. J., & Taylor, S. G. (2016). Viewing the interpersonal mistreatment literature through a temporal lens. *Organizational Psychology Review*, *6*(3), 273–302. https://doi.org/10.1177/2041386615607095.

Collinson, D. (2005). Dialectics of leadership. *Human Relations*, *58*(11), 1419–1442. https://doi.org/10.1177/0018726705060902.

Coughlin, J. F., Reimer, B., & Mehler, B. (2009). *Driver Wellness, Safety & the Development of an AwareCar*. Cambridge, MA: MIT AgeLab.

Dalal, R. S. (2005). A meta-analysis of the relationship between organizational citizenship behavior and counterproductive work behavior. *Journal of Applied Psychology*, *90*(6), 1241–1255. https://doi.org/10.1037/0021-9010.90.6.1241.

Damen, F., van Knippenberg, D., & van Knippenberg, B. (2008). Leader affective displays and attributions of charisma: the role of arousal. *Journal of Applied Social Psychology*, *38*(10), 2594–2614. https://doi.org/10.1111/j.1559-1816.2008.00405.x.

Dawson, M. E., Schell, A. M., & Filion, D. L. (2007). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 159–181). Cambridge, United Kingdom: Cambridge University Press.

Day, D., Gronn, P., & Salas, E. (2004). Leadership capacity in teams. *The Leadership Quarterly*, *15*(6), 857–880. https://doi.org/10.1016/j.leaqua.2004.09.001.

DeRue, D. S., Nahrgang, J. D., Wellman, N., & Humphrey, S. E. (2011). Trait and behavioral theories of leadership: an integration and meta-analytic test of their relative validity. *Personnel Psychology*, *64*(1), 7–52. https://doi.org/10.1111/j.1744-6570.2010.01201.x.

D'Esposito, M., Zarahn, E., & Aguirre, G. K. (1999). Event-related functional MRI: implications for cognitive psychology. *Psychological Bulletin*, *125*(1), 155–164. https://doi.org/10.1037/0033-2909.125.1.155.

Diebig, M., Bormann, K. C., & Rowold, J. (2016). A double-edged sword: relationship between full-range leadership behaviors and followers' hair cortisol level. *The Leadership Quarterly*, *27*(4), 684–696. https://doi.org/10.1016/j.leaqua.2016.04.001.

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science*, *11*(1), 86–89. https://doi.org/10.1111/1467-9280.00221.

Einarsen, S., Aasland, M. S., & Skogstad, A. (2007). Destructive leadership behavior: a definition and conceptual model. *The Leadership Quarterly*, *18*(3), 207–216. https://doi.org/10.1016/j.leaqua.2007.03.002.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, *6*(3-4), 169–200. https://doi.org/10.1080/02699939208411068.

Ekman, P. (1993). Facial expression and emotion. *The American Psychologist*, *48*(4), 384–392. https://doi.org/10.1037/0003-066X.48.4.384.

Endedijk, M. D., Hoogeboom, A. M. G. M., Groenier, M., de Laat, S., & Van Sas, J. (2018). Using sensor technology to capture the structure and content of team interactions in medical emergency teams during stressful moments. *Frontline Learning Research*, *6*, 123–147. https://doi.org/10.14786/flr.v6i3.353.

Erez, A., Misangyi, V., Johnson, D., LePine, M., & Halverson, K. (2008). Stirring the hearts of followers: charismatic leadership as the transferal of affect. *The Journal of Applied Psychology*, *93*(3), 602–616. https://doi.org/10.1037/0021-9010.93.3.602.

Ewen, C., Wihler, A., Blickle, G., Oerder, K., Ellen, B., Douglas, C., & Ferris, G. (2013). Further specification of the leader political skill-leadership effectiveness relationships: transformational and transactional leader behavior as mediators. *The Leadership Quarterly*, *24*(4), 516–533. https://doi.org/10.1016/j.leaqua.2013.03.006.

Fabes, R. A., & Martin, C. L. (1991). Gender and age stereotypes of emotionality. *Personality and Social Psychology Bulletin*, *17*(5), 532–540. https://doi.org/10.1177/0146167291175008.

Figner, B., & Murphy, R. O. (2011). Using skin conductance in judgment and decision making research. In M. Schulte-Mecklenbeck, A. Kuehberger, & R. Ranyard (Eds.), *A Handbook of Process Tracing Methods for Decision Research: A Critical Review and User's Guide* (pp. 163–184). New York, NY: Psychology Press.

Flach, P. (2012). *Machine Learning: The Art and Science of Algorithms that Make Sense of Data*. Cambridge, United Kingdom: Cambridge University Press.

Fleishman, E. A. (1953). The measurement of leadership attitudes in industry. *Journal of Applied Psychology*, *37*(3), 153–158. https://doi.org/10.1037/h0063436.

Fleishman, E. A. (1973). Twenty years of consideration and structure. In E. A. Fleishman, & J. G. Hunt (Eds.), *Current Developments in the Study of Leadership* (pp. 1–37). Carbondale, IL: Southern Illinois University Press.

Fleishman, E. A., & Harris, E. F. (1998). Patterns of leadership behavior related to employee grievances and turnover: some post hoc reflections. *Personnel Psychology*, *51*, 825–834. https://doi.org/10.1111/j.1744-6570.1998.tb00740.

Fleishman, E. A., Mumford, M. D., Zaccaro, S. J., Levin, K. Y., Korotkin, A. L., & Hein, M. B. (1991). Taxonomic efforts in the description of leader behavior: a synthesis and functional interpretation. *The Leadership Quarterly*, *2*(4), 245–287. https://doi.org/10.1016/1048-9843(91)90016-U.

Gakhal, B., & Senior, C. (2008). Examining the influence of fame in the presence of beauty: an electrodermal 'neuromarketing' study. *Journal of Consumer Behaviour*, *7*(4), 331–341. https://doi.org/10.1002/cb.255.

Gioia, D., & Sims, H. R., Jr. (1986). Cognition–behavior connections: attribution and verbal behavior in leader–subordinate interactions. *Organizational Behavior and Human Decision Processes*, *37*(2), 197–229. https://doi.org/10.1177/017084068901000403.

Golland, Y., Keissar, K., & Levit-Binnun, N. (2014). Studying the dynamics of autonomic activity during emotional experience. *Psychophysiology*, *51*(11), 1101–1111. https://doi.org/10.1111/psyp.12261.

Gooty, J., Connelly, S., Griffith, J., & Gupta, A. (2010). Leadership, affect and emotions: a state of the science review. *The Leadership Quarterly*, *21*(6), 979–1004. https://doi.org/10.1016/j.leaqua.2010.10.005.

Gorman, J. C., Cooke, N. J., & Amazeen, P. G. (2010). Training adaptive teams. *Human Factors*, *52*(2), 295–307. https://doi.org/10.1177/0018720810371689.

Guastello, S. J., Reiter, K., Shircel, A., Timm, P., Malon, M., & Fabisch, M. (2014). The performance-variability paradox, financial decision making, and the curious case of negative Hurst exponents. *Nonlinear Dynamics, Psychology, and Life Sciences*, *18*(3), 297–328. https://doi.org/10.1007/978-4-431-55312-0_5.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, *46*(6), 1251–1271. https://doi.org/10.2307/1913827.

Heaphy, E. D., & Dutton, J. E. (2008). Positive social interactions and the human body at work: linking organizations and physiology. *The Academy of Management Review*, *33*(1), 137–162. https://doi.org/10.5465/amr.2008.27749365.

Hoffman, E. L., & Lord, R. G. (2013). A taxonomy of event-level dimensions: implications for understanding leadership processes, behavior, and performance. *The Leadership Quarterly*, *24*, 558–571. https://doi.org/10.1016/j.leaqua.2013.03.009.

Hoogeboom, A. M. G. M., & Wilderom, C. P. M. (2015). Effective leader behaviors in regularly held staff meetings: Surveyed vs. videotaped and video-coded observations. In J. A. Allen, N. Lehmann-Willenbrock, & S. G. Rogelberg (Eds.), *The Cambridge handbook of meeting science (pp. 381-412)*. Cambridge, UK: Cambridge University Press.

Hoogeboom, M. A. M. G., & Wilderom, C. P. M. (2019). Advancing the Transformational–Transactional Model of Effective Leadership: Integrating two Classic Leadership Models with a Video-Based Method. *Journal of Leadership Studies*, *13*(2), 23–46. https://doi.org/10.1002/jls.21655.

Hoogeboom, M. A. M. G., & Wilderom, C. P. M. (2020). A complex adaptive systems approach to real-life team interaction patterns, task context, information sharing, and effectiveness. *Group and Organization Management*, *45*(1), 3–42. https://doi.org/10.1177/1059601119854927.

Ilgen, D., Fisher, C., & Taylor, M. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology*, *64*(4), 349–371. https://doi.org/10.1037//0021-9010.64.4.349.

Ilies, R., Curşeu, P. L., Dimotakis, N., & Spitzmuller, M. (2013). Leaders' emotional expressiveness and their behavioural and relational authenticity: effects on followers. *European Journal of Work and Organizational Psychology*, *22*(1), 4–14. https://doi.org/10.1080/1359432X.2011.626199.

Jiang, L., Yin, D., & Liu, D. (2020). Can joy buy you money? The impact of the strength, duration, and phases of an entrepreneur's peak displayed joy on funding performance. *Academy of Management Journal*. https://doi.org/10.5465/amj.2017.1423 Advance online publication.

Judge, T. A., & Piccolo, R. F. (2004). Transformational and transactional leadership: a meta-analytic test of their relative validity. *The Journal of Applied Psychology*, *89*(5), 755–768. https://doi.org/10.1037/0021-9010.89.5.755.

Judge, T. A., Piccolo, R. F., & Ilies, R. (2004). The forgotten ones? The validity of consideration and initiating structure in leadership research. *Journal of Applied Psychology*, *89*(1), 36–51. https://doi.org/10.1037/0021-9010.89.1.36.

Kauffeld, S., & Lehmann-Willenbrock, N. (2012). Meetings matter: effects of team meetings on team and organizational success. *Small Group Research*, *43*(2), 130–158. https://doi.org/10.1177/1046496411429599.

Keyton, J., & Beck, S. J. (2009). The influential role of relational messages in group interaction. *Group Dynamics: Theory, Research, and Practice*, *13*(1), 14–30. https://doi.org/10.1037/a0013495.

Kinner, V. L., Kuchinke, L., Dierolf, A. M., Merz, C. J., Otto, T., & Wolf, O. T. (2017). What our eyes tell us about feelings: tracking pupillary responses during emotion regulation processes. *Psychophysiology*, *54*(4), 508–518. https://doi.org/10.1111/psyp.12816.

Klonek, F., Gerpott, F., Lehmann-Willenbrock, N., & Parker, S. (2019). Time to go wild: how to conceptualize and measure process dynamics in real teams with high-resolution. *Organizational Psychology Review*, *9*(4), 245–275. https://doi.org/10.1177/2041386619886674.

Klonek, F. E., Burba, M., Kauffeld, S., & Quera, V. (2016). Group interactions and time: using sequential analysis to study group dynamics in project meetings. *Group Dynamics*, *20*(3), 209–222. https://doi.org/10.1037/gdn0000052.

Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The sources of four commonly reported cutoff criteria: what did they really say? *Organizational Research Methods*, *9*(2), 202–220. https://doi.org/10.1177/1094428105284919.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.1016/j.ijpsycho.2017.04.009.

Larsen, R. J., Diener, E., & Lucas, R. E. (2002). Emotion: Models, measures, and individual differences. In R. G. Lord, R. J. Klimoski, & R. Kanfer (Eds.), *Emotions in the Workplace: Understanding the Structure and Role of Emotions in Organizational Behavior* (pp. 64–106). San Francisco, CA: Jossey-Bass.

LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods*, *11*(4), 815–852. https://doi.org/10.1177/1094428106296642.

Lee, N., Senior, C., & Butler, M. (2012). The domain of organizational cognitive neuroscience: theoretical and empirical challenges. *Journal of Management*, *38*(4), 921–931. https://doi.org/10.1177/0149206312439471.

Lehmann-Willenbrock, N., & Allen, J. A. (2014). How fun are your meetings? Investigating the relationship between humor patterns in team interactions and team performance. *Journal of Applied Psychology*, *99*(6), 1278–1287. https://doi.org/10.1037/a0038083.

Lehmann-Willenbrock, N., & Allen, J. A. (2018). Modeling temporal interaction dynamics in organizational settings. *Journal of Business and Psychology*, *33*(3), 325–344. https://doi.org/10.1007/s10869-017-9506-9.

Lehmann-Willenbrock, N., Meinecke, A. L., Rowold, J., & Kauffeld, S. (2015). How transformational leadership works during team interactions: a behavioral process analysis. *The Leadership Quarterly*, *26*(6), 1017–1033. https://doi.org/10.1016/j.leaqua.2015.07.003.

Lewis, K. M. (2000). When leaders display emotion: how followers respond to negative emotional expression of male and female leaders. *Journal of Organizational Behavior*, *21*(2), 221–234. https://doi.org/10.1002/(SICI)1099-1379(200003).

Lidberg, L., & Wallin, B. G. (1981). Sympathetic skin nerve discharges in relation to amplitude of skin resistance responses. *Psychophysiology*, *18*(3), 268–270. https://doi.org/10.1111/j.1469-8986.1981.tb03033.x.

van Lier, H. G., Oberhagemann, M., Stroes, J. D., Enewoldsen, N. M., Pieterse, M. E., Schraagen, J. M. C., & Noordzij, M. L. (2017). Design decisions for a real time, alcohol craving study using physio- and psychological measures. In P. de Vries, H. Oinas-Kukkonen, L. Siemons, N. B. -d. Jong, & L. van Gemert-Pijnen (Eds.), *Persuasive Technology: Development and Implementation of Personalized Technologies to Change Attitudes and Behaviors* (pp. 3–16). Enschede, the Netherlands: Springer.

Lindebaum, D., Al-Amoudi, I., & Brown, V. L. (2018). Does leadership development need to care about neuro-ethics? *Academy of Management Learning & Education*, *17*(1), 96–109. https://doi.org/10.5465/amle.2016.0220.

Lindebaum, D., & Jordan, P. J. (2014). When it can be good to feel bad and bad to feel good: exploring asymmetries in workplace emotional outcomes. *Human Relations*, *67*(9), 1037–1050. https://doi.org/10.1177/0018726714535824.

London, M., & Smither, J. (2002). Feedback orientation, feedback culture, and the longitudinal performance management process. *Human Resource Management Review*, *12*(1), 81–100. https://doi.org/10.1016/S1053-4822(01)00043-2.

Marci, C. D., Ham, J., Moran, E., & Orr, S. P. (2007). Physiologic correlates of perceived therapist empathy and social-emotional process during psychotherapy. *The Journal of Nervous and Mental Disease*, *195*(2), 103–111. https://doi.org/10.1097/01.nmd.0000253731.71025.fc.

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: a review. *Cognition and Emotion*, *23*(2), 209–237. https://doi.org/10.1080/02699930802204677.

McClean, S. T., Barnes, C. M., Courtright, S. H., & Johnson, R. E. (2019). Resetting the clock on dynamic leader behaviors: a conceptual integration and agenda for future research. *Academy of Management Annals*, *13*(2), 479–508. https://doi.org/10.5465/annals.2017.0081.

McClelland, G. H., Lynch, J. G., Irwin, J. R., Spiller, S. A., & Fitzsimons, G. J. (2015). Median splits, type II errors, and false-positive consumer psychology: Don't fight the power. *Journal of Consumer Psychology*, *25*(4), 679–689. https://doi.org/10.1016/j.jcps.2015.05.006.

McGrath, J. E. (1962). *Leadership Behavior: Requirements for Leadership Training.* Washington, DC: Prepared for U.S. Civil Service Commission Office of Career Development.

McNeish, D., & Kelley, K. (2019). Fixed effects models versus mixed effects models for clustered data: Reviewing the approaches, disentangling the differences, and making recommendations. *Psychological Methods*, *24*(1), 20–35. https://doi.org/10.1037/met0000182.

Meinecke, A. L., Kauffeld, S., & Lehmann-Willenbrock, N. (2017). What happens during annual appraisal interviews? How leader-follower interactions unfold and impact interview outcomes. *Journal of Applied Psychology*, *102*(7), 1054–1074. https://doi.org/10.1037/apl0000219.

Mintzberg, H. T. D. (2004). Third-generation management development. *Annual Leadership Issue*, *58*(3), 28–38.

Morgeson, F. P., DeRue, D. S., & Karam, E. P. (2010). Leadership in teams: a functional approach to understanding leadership structures and processes. *Journal of Management*, *36*(1), 5–39. https://doi.org/10.1177/0149206309347376.

Murray, M. M., & Antonakis, J. (2019). An introductory guide to organizational neuroscience. *Organizational Research Methods*, *22*, 6–16. https://doi.org/10.1177/1094428118802621.

Naqvi, N. H., & Bechara, A. (2006). Skin conductance: A psychophysiological approach to the study of decision making. In C. Senior, T. Russell, & M. S. Gazzaniga (Eds.), *Methods in Mind* (pp. 103–122). Cambridge, MA: The MIT Press.

Navarro, J., & Rueff-Lopes, R. (2015). Healthy variability in organizational behavior: empirical evidence and new steps for future research. *Nonlinear Dynamics, Psychology, and Life Sciences*, *19*(4), 529–552. https://doi.org/10.2307/258555.

Newman, D., & Sin, H. P. (2009). How do missing data bias estimates of within-group agreement? Sensitivity of sD wG, cVWG, rWG(J), rWG(J), and iCC to systematic nonresponse. *Organizational Research Methods*, *12*(1), 113–147. https://doi.org/10.1177/1094428106298969.

Noldus, L. P. J. J., Trienes, R. J. H., Hendriksen, A. H. M., Jansen, H., & Jansen, R. G. (2000). The observer video-pro: new software for the collection, management, and presentation of time-structured data from videotapes and digital media files. *Behavior Research Methods, Instruments, & Computers*, *32*(1), 197–206. https://doi.org/10.3758/BF03200802.

Pearce, C. L., Sims, H. P., Jr., Cox, J. F., Ball, G., Schnell, E., Smith, K. A., & Trevino, L. (2003). Transactors, transformers and beyond: a multi-method development of a theoretical typology of leadership. *Journal of Management Development*, *22*(4), 273–307. https://doi.org/10.1108/02621710310467587.

Pennebaker, J. W., Hughes, C. F., & O'Heeron, R. C. (1987). The psychophysiology of confession: linking inhibitory and psychosomatic processes. *Journal of Personality and Social Psychology*, *52*(4), 781–793. https://doi.org/10.1037/0022-3514.52.4.781.

Penney, L. M., & Spector, P. E. (2005). Job stress, incivility, and counterproductive work behavior (CWB): the moderating role of negative affectivity. *Journal of Organizational Behavior*, *26*(7), 777–796. https://doi.org/10.1002/job.336.

Picard, R. W., Fedor, S., & Ayzenberg, Y. (2016). Multiple arousal theory and daily-life electrodermal activity asymmetry. *Emotion Review*, *8*(1), 62–75. https://doi.org/10.1177/1754073914565517.

Piccolo, R. F., Bono, J. E., Heinitz, K., Rowold, J., Duehr, E., & Judge, T. A. (2012). The relative impact of complementary leader behaviors: which matter most? *The Leadership Quarterly*, *23*(3), 567–581. https://doi.org/10.1016/j.leaqua.2011.12.008.

Poláčková Šolcová, I., & Lačev, A. (2017). Differences in male and female subjective experience and physiological reactions to emotional stimuli. *International Journal of Psychophysiology*, *117*(4), 75–82. https://doi.org/10.1016/j.ijpsycho.2017.04.009.

van Prooijen, A., Ellemers, N., Van der Lee, R., & Scheepers, D. (2018). What seems attractive may not always work well: evaluative and cardiovascular responses to morality and competence levels in decision-making teams. *Group Processes & Intergroup Relations*, *21*(1), 73–87. https://doi.org/10.1177/1368430216653814.

Ravaja, N., Turpeinen, M., Saari, T., Puttonen, S., & Keltikangas-Järvinen, L. (2008). The psychophysiology of James bond: phasic emotional responses to violent video game events. *Emotion*, *8*(1), 114–120. https://doi.org/10.1037/1528-3542.8.1.114.

Sano, A., Taylor, S., McHill, A. W., Phillips, A. J., Barger, L. K., Klerman, E., & Picard, R. (2018). Identifying objective physiological markers and modifiable behaviors for self-reported stress and mental health status using wearable sensors and mobile phones: observational study. *Journal of Medical Internet Research*, *20*(6), 210. https://doi.org/10.2196/jmir.9410.

Scheepers, D. (2009). Turning social identity threat into challenge: status stability and cardiovascular reactivity during inter-group competition. *Journal of Experimental Social Psychology*, *45*(1), 228–233. https://doi.org/10.1016/j.jesp.2008.09.011.

Scheepers, D., & Ellemers, N. (2005). When the pressure is up: the assessment of social identity threat in low and high status groups. *Journal of Experimental Social Psychology*, *41*(2), 192–200. https://doi.org/10.1016/j.jesp.2004.06.002.

Schriesheim, C. A., & Bird, B. J. (1979). Contributions of the Ohio State studies to the field of leadership. *Journal of Management*, *5*, 135–145. https://doi.org/10.1177/014920637900500204.

Senior, C., Lee, N., & Butler, M. (2011). Organizational cognitive neuroscience. *Organization Science*, *22*(3), 804–815. https://doi.org/10.1287/orsc.1100.0532.

Siegel, E. H., Sands, M. K., Van den Noortgate, W., Condon, P., Chang, Y., Dy, J., & Barrett, L. F. (2018). Emotion fingerprints or emotion populations? A meta-analytic investigation of autonomic features of emotion categories. *Psychological Bulletin*, *144*(4), 343–393. https://doi.org/10.1037/bul0000128.

Sims, H. P., & Manz, C. C. (1984). Observing leader behavior: toward reciprocal determinism in leadership theory. *Journal of Applied Psychology*, *69*(2), 222–232. https://doi.org/10.1037//0021-9010.69.2.222.

Sommer, S. A., Howell, J. M., & Hadley, C. N. (2016). Keeping positive and building strength: the role of affect and team leadership in developing resilience during an organizational crisis. *Group & Organization Management*, *41*(2), 172–202. https://doi.org/10.5465/AMBPP.2015.17642.

Spain, S., & Harms, P. (2014). A sociogenomic perspective on neuroscience in organizational behavior. *Frontiers in Human Neuroscience*, *8*, 1–15. https://doi.org/10.3389/fnhum.2014.00084.

Spiers, J. A. (2004). Tech tips: using video management analysis technology in qualitative research. *International Journal of Qualitative Methods*, *3*(1), 57–61. https://doi.org/10.1177/160940690400300106.

Stemmler, G. (2004). Physiological processes during emotion. In P. Philippot, & R. S. Feldman (Eds.), *Regulation of Emotion* (pp. 33–70). Mahwah, NJ: Erlbaum.

Talarico, J. M., LaBar, K. S., & Rubin, D. C. (2004). Emotional intensity predicts autobiographical memory experience. *Memory & Cognition*, *32*(7), 1118–1132. https://doi.org/10.3758/BF03196886.

Tee, E. Y. J., Ashkanasy, N. M., & Paulsen, N. (2011, April). Upward emotional contagion and implications for leadership: From a cognitive, leader-centric approach to an effective, follower-centric model of leadership. *Paper Presented at Society of Industrial and Organizational Psychology Annual Meeting, Chicago*.

Tepper, B. J., Henle, C. A., Lambert, L. S., Giacalone, R. A., & Duffy, M. K. (2008). Abusive supervision and subordinates' organization deviance. *Journal of Applied Psychology*, *93*(4), 721–732. https://doi.org/10.1016/j.leaqua.2015.07.003.

Tranel, D., & Damasio, H. (1989). Intact electrodermal skin-conductance responses after bilateral amygdala damage. *Neuropsychologia*, *27*(4), 381–390. https://doi.org/10.1016/0028-3932(89)90046-8.

Tranel, D., & Damasio, H. (1994). Neuroanatomical correlates of electrodermal skin-conductance responses. *Psychophysiology*, *31*(5), 427–438. https://doi.org/10.1111/j.1469-8986.1994.tb01046.x.

Uhl-Bien, M. (2006). Relational leadership theory: Exploring the social processes of leadership and organizing. *The Leadership Quarterly*, *17*(6), 654–676. https://doi.org/10.1016/j.leaqua.2006.10.007.

Van Quaquebeke, N., & Felps, W. (2018). Respectful inquiry: a motivational account of leading through asking questions and listening. *Academy of Management Review*, *43*(1), 5–27. https://doi.org/10.5465/amr.2014.0537.

Vecchio, R., Justin, J., & Pearce, C. (2008). The utility of transactional and transformational leadership for predicting performance and satisfaction within a path-goal theory framework. *Journal of Occupational and Organisational Psychology*, *81*(1), 71–82. https://doi.org/10.1348/096317907X202482.

Vrana, S. R., & Gross, D. (2004). Reactions to facial expressions: effects of social context and speech anxiety on responses to neutral, anger, and joy expressions. *Biological Psychology*, *66*(1), 63–78. https://doi.org/10.1016/j.biopsycho.2003.07.004.

Waldman, D. A., Balthazard, P. A., & Peterson, S. J. (2011). Leadership and neuroscience: can we revolutionize the way that inspirational leaders are identified and developed? *The Academy of Management Perspectives*, *25*(1), 60–74. https://doi.org/10.5465/amp.25.1.60.

Waller, L., Reitz, M., Poole, E., Riddell, P. M., & Muir, A. (2017). Experiential learning as preparation for leadership. *Leadership and Organization Development Journal*, *38*(4), 513–529. https://doi.org/10.1108/LODJ-03-2015-0057.

Wang, H., Tsui, A. S., & Xin, K. R. (2011). CEO leadership behaviors, organizational performance, and employees' attitudes. *The Leadership Quarterly*, *22*(1), 92–105. https://doi.org/10.1016/j.leaqua.2010.12.009.

Watzlawick, P., Beavin, J. H., & Jackson, D. D. (1967). *Pragmatics of Human Communication: A Study of Interactional Patterns, Pathologies and Paradoxes*. New York, NY: Norton & Company.

Weis, P. P., & Herbert, C. (2017). Bodily reactions to emotional words referring to own versus other people's emotions. *Frontiers in Psychology*, *8*, 1–14. https://doi.org/10.3389/fpsyg.2017.01277.

White, R. E., Thornhill, S., & Hampson, E. (2006). Entrepreneurs and evolutionary biology: the relationship between testosterone and new venture creation. *Organizational Behavior and Human Decision Processes*, *100*(1), 21–34. https://doi.org/10.1016/j.obhdp.2005.11.001.

Yukl, G. (2010). *Leadership in Organizations* (8th ed.). Englewood Cliffs, NJ: Prentice Hall.

Zyphur, M., Narayanan, J., Koh, G., & Koh, D. (2009). Testosterone-status mismatch lowers collective efficacy in groups: evidence from a slope-as-predictor multilevel structural equation model. *Organizational Behavior and Human Decision Processes*, *110*(2), 70–79. https://doi.org/10.1016/j.obhdp.2009.05.004.