



Faking It with the Boss's Jokes? Leader Humor Quantity, Follower Surface Acting, and Power Distance

Xiaoran Hu

The London School of Economics and Political Science

X.Hu16@lse.ac.uk

Michael R. Parke

University of Pennsylvania

mrparke@wharton.upenn.edu

Randall S. Peterson

London Business School rpeterson@london.edu

Grace Simon

University of Pennsylvania

gmsimon@wharton.upenn.edu

Faking It with the Boss's Jokes? Leader Humor Quantity, Follower Surface Acting, and Power Distance

ABSTRACT

Most scholarly work concludes that leader humor positively affects emotional outcomes for followers. However, two interrelated issues with past research make this view incomplete: (1) studies often conflate the act of expressing humor with the humor's effectiveness, and (2) existing research overlooks follower emotion regulation. In addressing these issues, we generate and test new theory that challenges the theoretical consensus, highlighting why and when leader humor expression has emotional costs for followers. Specifically, we theorize that leader humor *quantity*, irrespective of its *quality*, puts pressure on followers to engage in surface acting to fake or exaggerate positive emotions. We further propose that these surface acting responses are more likely to occur when followers hold high power distance values, such that followers more willingly comply with expectations to display positive emotions in response to leader humor expression. This increase in surface acting then leads to more emotional exhaustion and less job satisfaction for followers. Results from three studies—including a field experiment, a laboratory experiment, and a multi-wave field study—provide support for these hypotheses. We close with a discussion of how our findings provide a unique counterpoint regarding the effects of leader humor on follower emotional outcomes.

Keywords: Leader humor, surface acting, power distance, emotional labor, wellbeing

Leader humor expression, or the extent to which a leader communicates or shares stimuli intended to be amusing to a follower (Cooper, 2005; Cooper, Kong, & Crossley, 2018), can be an effective tool for leaders (Romero & Cruthirds, 2006). Current theory suggests that leaders who use humor to bring amusement, happiness, or joviality to followers are likely to induce positive emotions in them, which alleviates negative work experiences, such as exhaustion or strain, and makes work more enjoyable (e.g., Cooper et al., 2018; Mesmer-Magnus, Glew, & Viswesvaran, 2012; Robert & Wilbanks, 2012; Romero & Cruthirds, 2006). Thus, a central tenet of leader humor research is that leader humor generally benefits, and certainly does not harm, follower wellbeing outcomes like emotional exhaustion and job satisfaction (e.g., Decker, 1987; Hughes & Avey, 2009; Kong, Cooper, Sosik, 2019; Vecchio, Justin, & Pearce, 2009). Because employee wellbeing outcomes help drive organizational success by enhancing employee performance and reducing employee attrition (Judge, Thoresen, Bono, & Patton, 2001; Rubenstein, Eberly, Lee, & Mitchell, 2018; Swider & Zimmerman, 2010), it is not surprising as to why leaders widely believe that humor is key for their career advancement, and as a result, prevalent in the modern workplace (2017 Robert Half International Survey).

However, two distinct concerns about the leader humor literature suggest that these claimed benefits are incomplete. First, most evidence supporting leader humor's beneficial effects on follower wellbeing comes from studies that use measures conflating leader humor expression with positive outcomes (e.g., Hughes & Avery, 2009; Karakowsky, Podolsky, & Elangovan, 2020; Sobral & Islam, 2015; Vecchio, Justin, & Pearce, 2009). Such studies combine leader humor expression, a behavior representing the act or *quantity* of sharing something intended to be amusing to followers, with its positive outcomes—i.e., the *quality* of leader humor in which it is perceived to be funny, induce amusement, or relieve stress. For example, the most common measures to capture leader humor include items such as my leader “uses humor to take

the edge off during stressful periods” (Avolio, Howell, & Sosik, 1999), “enjoys making people laugh” (scale by Martin et al., 2003 adapted by Pundt & Hermann, 2015), and “has a good sense of humor” (Decker, 1987; see also quantitative reviews by Kong, Cooper, & Sosik, 2019 and Mesmer-Magnus et al., 2012 that illustrate measures combining leader humor quantity and quality). We argue that such a conflation can imprecisely inflate the relationship between leaders expressing humor and wellbeing outcomes. Moreover, this conflation has likely inhibited theoretical and empirical specification of leader humor quantity’s potential adverse effects on follower wellbeing. By isolating leader humor quantity as a behavior distinct from its comedic quality, we can begin to understand why and when leader humor expression might have emotional costs.

The second reason leader humor’s wellbeing benefits might be overstated is because existing research overlooks a critical theoretical perspective. In contrast to the genuine positive emotions theorized to occur in response to leader humor expression, we posit that employees may also be prompted to *act* positively. That is, emotion-eliciting events like leader humor expression do not just trigger emotional reactions but also emotion regulation (Elfenbein, 2007). Research on emotional labor (e.g., Menges, Kilduff, Kern, & Bruch, 2015; Shumski Thomas, Olien, Allen, Rogelberg & Kello, 2018) and the experience of power (e.g., Magee & Galinsky, 2008) highlight how followers, as lower-power actors, tend to regulate their emotions in the presence of higher-power leaders. They do so to display appropriate emotions, such as showing positive emotions like smiling while hiding negative feelings (Hecht & LaFrance, 1998). Thus, although leader humor expression can increase follower positive emotions, as is currently theorized, it is also likely to increase follower *surface acting*—or the extent to which an employee changes their public display of feelings by masking, faking, or suppressing what they

feel internally (Côté, 2005; Grandey & Gabriel, 2015)—which impairs rather than helps wellbeing (Hülshager & Schewe, 2011; Kammeyer-Mueller et al., 2013).

Combining these insights, we aim to challenge the overly positive view connecting leader humor and follower emotional wellbeing. We develop and test new theory as to why and when leader humor *quantity* (hereafter LH quantity), irrespective of leader humor *quality* (hereafter LH quality), can increase follower surface acting, and as a result, indirectly lead to increased *emotional exhaustion* (i.e., feeling overextended and depleted; Taris et al., 2004) and reduced *job satisfaction* (i.e., the positive appraisals of one's job; Locke, 1976). Applying theory on emotional regulation and the experience of power, we propose that when leaders express humor, which signals an expectation to respond with the display of positive emotions such as laughing (Chapman, 1983), followers will likely feel pressure to show, exaggerate, or even fake positive emotional reactions to their higher-power bosses, regardless of how they actually feel (i.e., surface act). Further, although the inherent power differential between leaders and followers sets the stage for follower surface acting from leader humor expression, how psychologically attuned and sensitive followers are to the power differential will determine the degree to which surface acting occurs. In this context, we propose that employees' *power distance*, or the extent to which followers accept the unequal distribution of power in institutions and organizations (Kirkman, Chen, Farh, Chen, & Lowe, 2009), plays a crucial role in determining when followers will more likely surface act in response to leader humor expression. Figure 1 illustrates our theoretical model. We test this model and find supporting evidence across three studies—a field experiment, a laboratory experiment, and a multi-wave field study.

Our paper aims to make three contributions. First, we challenge the dominant narrative on leader humor and follower emotional wellbeing outcomes. Existing theory positions leader humor as beneficial for follower wellbeing, either positively influencing wellbeing outcomes or,

at worst, having no impact (Decker, 1987; Cooper et al., 2018; Guenzi, Rangarajan, Chaker, & Sajtos, 2019; Hughes & Avey, 2009; Robert, Dunne, & Iun, 2016; Vecchio, Justin, & Pearce, 2009). However, our research introduces a complication and counterpoint by highlighting the adverse effects of LH quantity on followers through increased surface acting. Second, we demonstrate the value of theoretically and empirically distinguishing LH quantity from LH quality. We show how work that explicitly or implicitly examines LH quality likely produces incomplete estimates for the effects of leader humor because it fails to account for LH quantity's emotional costs. Additionally, our exploratory analyses of LH quantity and LH quality illustrate how separating them can lead to novel findings, implications, and predictions, such as how funny leaders, because they express humor more frequently, can have the unintended effect of increased surface acting in followers. Third, our research generates new knowledge for leaders on the risks of expressing humor at work. Moving beyond the established fact that humor is risky based on *what* leaders say (i.e., humor quality), we demonstrate how it is also risky based on *how much* they say (i.e., humor quantity) and to *whom* they say it (i.e., follower power distance).

THEORY DEVELOPMENT AND HYPOTHESES

Leader Humor and Follower Wellbeing Outcomes

Research connecting leader humor and wellbeing falls into one of two theoretical domains—the (1) stress relief hypothesis, and (2) the inducing positive affective states hypothesis (Cooper et al., 2018). Each of these perspectives aligns with an employee wellbeing outcome that humor aims to improve – decreasing burnout-based experiences (i.e., exhaustion) and increasing job satisfaction. The most common explanatory mechanism for how leader humor affects these outcomes is through follower positive affect (Cooper et al., 2018) – a generalized and diffused emotional state characterized by pleasant emotions such as happiness, interest, and amusement (Barsade & Gibson, 2007). By increasing follower positive affect, leader humor

helps alleviate followers' feelings of exhaustion and "can lighten the mood within organizational environments and make work life more enjoyable" (Romero & Cruthirds, 2006: 58).

However, we question whether this positive view is the whole story. Several theoretical and empirical issues in studies on leader humor make current inferences about its effects on follower wellbeing uncertain. Crucially, how scholars conceptualize and operationalize leader humor matters. Humor is defined as "any event shared by an agent...with another individual (i.e., a target) that is intended to be amusing to the target and that the target perceives as an intentional act" (Cooper, 2005: 766–767; see also Pundt & Herrmann, 2015). Leader humor is when the agent occupies a leadership position and expresses humor to a follower (i.e., subordinate) (Cooper et al., 2018). Leader humor is thus an expressive behavior—the *quantity*, frequency, or extent that a leader tells jokes, stories, puns, uses wit, or shares stimuli that intend to entertain or create an amused reaction in a follower. Leader humor expression (i.e., LH quantity), therefore, is distinct from the manifestation of its intended positive outcomes.

The problem, however, is that many studies have treated leader humor not as an expressive behavior but as an effectiveness attribute whereby the leader's humor results in positive outcomes. For instance, the leader's humor is deemed typically funny, good, or positive (e.g., Decker, 1987; Goswami et al., 2016; Karakowsky et al., 2020); it makes people laugh (e.g., Carnevale, Huang, Yam, & Wang, 2022; Pundt & Hermann, 2015; Robert et al., 2016); it generates amused reactions (Hurren, 2006); or it relieves tension, stress, or conflict (e.g., Avolio et al., 1999; Hughes & Avery, 2009; Sobral & Islam, 2015; Vecchio, Justin, & Pearce, 2009). In other words, studies often investigate leader humor *quality* in which they combine into one construct both the leader's behavior (i.e., expression of humor) and the mechanism of inducing positive emotions in followers. Therefore, it is not surprising that LH quality consistently

improves job satisfaction and reduces exhaustion (e.g., Hughes & Avey, 2009; Kong et al., 2019; Mesmer-Magnus et al., 2012; Sobral & Islam, 2015; Vecchio et al., 2009). Similarly, work that defines humor by its negative outcomes (i.e., low-quality humor), such that the leader's teasing, mocking, or put-downs make followers feel offended, inferior, or demeaned (e.g., aggressive or superiority humor), shows it results in negative effects on wellbeing (e.g., Guenzi et al., 2019; Huo, Lam, & Chen, 2012; Susa, 2002). Because most studies have implicitly or explicitly focused on LH quality, the relationship between LH quantity and follower emotional outcomes is more uncertain. To this point, Cooper et al. (2018) examined leader humor as a behavior and found a nonsignificant effect with emotional exhaustion. Moreover, the focus on LH quality, which by definition is good, has hindered investigation into how the act of leader humor expression (e.g., LH quantity) can unintentionally hurt follower wellbeing.

In this paper, we isolate leader humor as a behavior (i.e., LH quantity) and integrate research on emotional labor and power to specify its emotional costs to followers. We theorize why and when LH quantity, irrespective of a leader's humor quality¹, increases follower surface acting, which then subsequently worsens emotional exhaustion and job satisfaction. Our focus on LH quantity also aligns with calls in the broader leadership literature to keep leader behaviors distinct from their intended effectiveness outcomes, which enables greater precision in causal

predictions and estimates of leader behavior effects (van Knippenberg & Sitkin, 2013).

Furthermore, understanding the effects of LH quantity has practical value as it provides

¹ Because our theoretical focus is on the unintended emotional costs of LH quantity, we make predictions for LH quantity while holding constant (i.e., controlling for) LH quality. Furthermore, given the lack of studies that have examined both LH quantity and LH quality simultaneously, we lack theory or evidence to make a priori predictions about their relationship or how they may have joint effects on follower emotional outcomes. That said, at the end of our studies, we explore in a post-hoc fashion their possible inter-relationships and effects on follower outcomes.

knowledge on how often leaders should attempt humor and to whom; by contrast, the takeaway from LH quality-wellbeing findings—that a leader’s humor should be genuinely funny and make people feel amused or less stressed—is arguably less useful. Figure 1 illustrates our model.

-----Insert Figure 1 about here-----

Leader Humor Quantity and Follower Surface Acting

Theory on emotional labor outlines a three-part process for when employees are likely to surface act (Diefendorff & Gosserand, 2003; Grandey, 2000; Grandey & Gabriel, 2015; Morris & Feldman, 1996). First, an emotional display standard is created or made salient such that an employee perceives the need to display a specific emotion in a certain situation (e.g., smile at the leader). Second, an emotional event creates a discrepancy between an employee’s naturally occurring *outward* emotional appearance and the display standard. Third, the employee is more likely to use surface acting to meet the emotional display standard, as opposed to deep acting, when the event is unexpected and interpersonal acceptance is instrumental to the employee. Below, we unpack this emotional labor process for LH quantity’s effect on surface acting.

To begin, we propose that the more leaders express humor, irrespective of the humor’s quality, the more followers will come to perceive expectations for displaying positive emotions (e.g., happiness or amusement) and concealing negative emotions. Expressing humor is a unique form of social exchange in that it requires the target to respond with a specific emotional reaction for the encounter to be successful. The initiator of humor needs the target to display positive emotions (e.g., laugh, smile, appear amused, or show appreciation) for the social interaction to be complete and successful (Hay, 2001). Humor is so universal and ubiquitous that the target and the initiator are both aware of this expectation of the appropriate responses to humor (Bennett & Lengacher, 2006; Chapman, 1983, Hay, 2001). As a result, LH quantity meets the conditions specified by theory on emotional labor that make emotional display rules strong (Diefendorff &

Gosserand, 2003; Morris & Feldman, 1996): the display rule is salient and repeated (frequent humor expression), the display expectation is specific and unambiguous (a concrete display expectation for followers to laugh, smile, or show amusement), and the source of the display standard has authority and legitimacy (leaders have higher power than followers).

Given that LH quantity likely increases positive display expectations, the next step in the emotional labor process is whether the leader humor expression creates a discrepancy between these display expectations and followers' naturally occurring outward display of their current affective state (Diefendorff & Gosserand, 2003). We argue that LH quantity will frequently create such discrepancies in two ways. The first is that as LH quantity increases, the greater the chances that some of the leaders' humor expression will not generate the intended positive feelings of amusement in the follower. This possibility is likely since reactions to humor are uniquely individualized (Romero & Cruthirds, 2006), and individuals differ from each other in their sense of humor (i.e., what makes them genuinely amused; Robert et al., 2016). As such, as LH quantity increases, it increases the likelihood of discrepancies occurring where followers do not genuinely feel, but must act, amused.

Second, even if the leader's humor expression produces genuine amusement, it will still likely create a discrepancy between followers' naturally occurring outward display and the expected display. To this point, Morris and Feldman (1996: 988) describe such emotional labor:

“even in situations in which there is congruence between the individual's felt emotion and the...desired emotion, there will still be some degree of effort (or “labor”) required in expressing emotions...individuals will still have to exert some effort to ensure that what is felt will be displayed in...appropriate ways (i.e., that the feeling of happiness is displayed in an appropriate smile or greeting).”

For leader humor expression, followers will likely need to display more exaggerated positivity to ensure that the leader knows unambiguously that they find the humor amusing. Also, leader humor expression likely necessitates followers to manufacture the outward display of

strong, intense, or exaggerated positivity because such humor is typically an affective shock—the display expectation of leader humor expression differs from default affective states, which tend to be lower in energy (low arousal) and vary between slightly negative, neutral, or slightly positive (Seo, Barrett, & Jin, 2008). Even if followers think the leader’s humor expression is amusing, they may not automatically display their amusement with outward laughing or smiling. For example, past work indicates how individuals who feel happy do not always smile or show it to others (Diefendorff et al., 2010). Alternatively, followers may prefer a subdued response to the leader’s humor expression (e.g., they might be tired or stressed from an unrelated source). However, not displaying amusement or showing subdued reactions in response to leader humor expression are less socially appropriate. In short, LH quantity, regardless of whether the humor is genuinely amusing or not, can increase the demand for emotional labor by increasing the number of discrepancies between the expected emotional display standard and followers’ naturally occurring outward displays.

The last step in the emotional labor process helps explain why employees will surface act in response to leader humor expression to meet the emotional display standard. Employees are especially likely to surface act (a) when displaying the correct emotion is instrumental to their interpersonal or resource goals and (b) when the emotion-eliciting event is unexpected (Diefendorff & Gosserand, 2003; Ozcelik, 2013). Starting with the former, employees surface act more with organizational members when they desire to fulfill either or both of two motives: (a) “to maintain interpersonal acceptance” and (b) “to obtain or secure valued resources and outcomes” (Ozcelik, 2013: 292). Leaders have the power to control valuable resources (e.g., social recognition and financial rewards) and influence a follower’s career advancement opportunities (Magee & Galinsky, 2008). In general, followers are motivated to provide positive upward feedback to leaders, even if they feel otherwise, to ensure they are accepted and to avoid

any potential punishment or backlash that may result from making leaders feel incompetent or inadequate (Farmer & Aguinis, 2005; Morrison & Milliken, 2000). Providing insincere or exaggerated positive feedback is more likely in situations when doing so will not harm work goals or performance (Levine, 2022), which is typically the case in response to leader humor expression (e.g., exaggerated laughing at the boss's jokes is unlikely to harm performance). Hence, followers are motivated to display positive emotional responses to leader humor expression to affirm the warmth and competence of the leader (Bitterly, 2022). Given the discrepancy between the display standard and followers' affective appearance outlined above, it is probable that followers will resort to surface acting to produce the expected positive emotional response. By contrast, if followers do not surface act, the absence of positive emotional responses to leader humor expression can cause leaders to lose face, appear incompetent, and potentially jeopardize followers' relationship with, or resources from, the leader, which are outcomes followers generally want to avoid.

Adding to followers' interpersonal and resource motives to surface act is the notion that leader humor expression evokes status dynamics. Humor expression is a high-risk, high-reward action such that successful humor can enhance the expressers' status while failed attempts can undermine it (Bitterly, Brooks, & Schweitzer, 2017). Leaders, as higher-power actors who desire to preserve their superior status (e.g., Anderson, Hildreth, & Sharps, 2020), likely pay special attention to whether followers enjoy their humor. This heightened leaders' attention on followers' reactions to humor likely adds further pressure for followers to surface act in order to be accepted and not challenge the leaders' status.

The second reason why followers are more likely to surface act in response to leader humor expression is because humor expression is an unexpected event. An unexpected

discrepancy between an employee's naturally occurring emotional display and the display standard results in surface acting because surface acting is a response-focused emotion regulation strategy (Diefendorff & Gosserand, 2003). When a discrepancy is unexpected, employees face a limited choice to either surface act (change the outward display but not the felt emotion) or express their authentic feeling. By contrast, when the discrepancy can be anticipated, such as flight attendants who know they will need to smile when greeting passengers, employees can use proactive emotion regulation strategies such as deep acting, whereby they change the situation or the cognitive meaning of the situation to produce the genuinely felt emotion (Grandey, 2000).

Humor, by definition, is unexpected because it involves an unanticipated incongruity or violation of something known, true, or expected (Warren, Barsky, & McGraw, 2021). Although it is possible to predict whether someone might express humor, predicting when the humor will occur or the specific content of the humor would be nearly impossible (unless it is a repeated joke to a similar situation, which, in such a case, a violation is unlikely to occur because the joke is known and expected, and therefore, it is unlikely to produce genuine amusement). Thus, followers likely resort to surface acting as the available emotion regulation strategy to meet the display rule from leader humor expression.

Taken together, when leaders express humor, irrespective of the humor's quality, followers find themselves in a situation where they need to suppress negative feelings and express, exaggerate, or fake positive emotional responses to comply with positive display rules. As such, leaders who attempt humor more frequently will trigger more situations where employees feel compelled to "put on an act" to display positive emotions. Thus, we hypothesize:

Hypothesis 1: Leader humor quantity increases follower surface acting.

Leader Humor Quantity Impacts Follower Emotional Wellbeing via Surface Acting

Surface acting in response to LH quantity will likely bring adverse outcomes for followers' emotional exhaustion and job satisfaction. Emotional exhaustion is feeling "overextended and depleted of one's emotional resources" (Taris et al., 2004), whereas job satisfaction refers to "a pleasurable or positive emotional state resulting from the appraisal of one's jobs or job experience" (Locke, 1976, p. 1300).

Past research has found that surface acting increases emotional exhaustion and reduces job satisfaction. First, surface acting increases employee tension due to the emotional dissonance and inauthenticity associated with it (Grandey, 2003; Hülshager & Schewe, 2011). In general, people desire to feel and act in an authentic manner (Lehman, O'Connor, Kovács, & Newman, 2019); any behavioral act or situation that impedes authenticity can add tension and strain and reduce people's ability to satisfy their needs. Second, surface acting is depleting as it requires cognitive resources and self-control to suppress or mask inward feelings in order to display externally inconsistent emotions (Grandey & Gabriel, 2015). Heightened surface acting is likely to drain employees, leaving them feeling more exhausted and less able to fulfill their needs and act in accordance with their values at work. Meta-analyses find that surface acting has a positive association with emotional exhaustion and a negative association with job satisfaction (Hülshager & Schewe, 2011; Kammeyer-Mueller et al., 2013). Thus, we hypothesize:

Hypothesis 2: Leader humor quantity has a positive indirect effect on follower emotional exhaustion mediated by surface acting.

Hypothesis 3: Leader humor quantity has a negative indirect effect on follower job satisfaction mediated by surface acting.

The Moderating Role of Follower Power Distance

Although humor expression and power are intricately intertwined (Bitterly, 2022), most research on leader humor ignores how followers are likely affected by their relatively lower power compared to leaders (for an exception in terms of power differences based on gender, see Gloor, Cooper, Bowes-Sperry, & Chawla, 2021). This oversight is problematic because humor expression, by intending to make the target feel amused or happy, fundamentally functions in a way to increase warmth and reduce social (i.e., power) distance between the humor expresser and the target (Romero & Cruthirds, 2006). Such interpersonal functions of humor can be highly valued and desirable among individuals of similar power or rank. However, when stark power differences exist, as is the case with leaders and followers, followers are likely to react differently to this reduced social distance depending on their personal beliefs and values—specifically, their power distance values. Thus, we incorporate followers' individual power distance—the extent to which followers accept the unequal distribution of power in organizations (Kirkman et al., 2009)—as a critical moderating factor of LH quantity's effects.

We propose that the reasons for followers to surface act in response to LH quantity (as discussed leading up to Hypothesis 1) will be more (less) pronounced for followers with higher (lower) power distance. Followers with high power distance will see obeying the display standards of LH quantity as more instrumental to their goals than followers with lower power distance. Past research demonstrates that high power distance followers are more likely to accept their leader as superior (Kirkman et al., 2009), defer to their leader, and meet leader expectations (Schaubroeck, Lam, & Cha, 2007), compared to low-power distance followers. Followers with high power distance enact these behaviors because they believe they are more dependent on their leaders for resources and guidance than those holding low power distance values (Wang, Mao, Wu, & Liu, 2012). Hence, followers with higher power distance feel greater pressure and

motivation to ensure their leaders accept them and provide them with valuable resources. This is why high power distance followers tend to provide more positive and respectful upward feedback to leaders compared to low power distance followers (Bond, Wan, Leung, & Giacalone, 1985; Ng, Koh, Ang, Kennedy, & Chan, 2011). It follows that LH quantity will more likely cause greater surface acting for followers with high (versus low) power distance.

We also propose that the unexpected nature of leader humor expression will be more pronounced for followers with high power distance, thus leading to greater surface acting by them. For followers with high power distance, the shock of leader humor expression should be greater because, in their view, it is more counter-normative. Followers with high power distance accept the boundaries of power between them and their leaders, and any actions or behaviors that muddle such boundaries tend to be counter-normative and unexpected (Alves et al., 2006; Kirkman et al., 2009). For example, followers with high power distance do not expect their leaders to share power by letting them make decisions, nor do they expect them to socialize because these actions can blur the power boundaries (Bochner & Hesketh, 1994; Hofstede, 1980; Smith, Peterson, & Schwartz, 2002). Thus, LH quantity will be more unexpected by followers with high power distance because it will be surprising each time their leader chooses to break their value of power distance. By contrast, low power distance followers are more likely to expect, rather than be surprised by, a social distance-reducing behavior like LH quantity.

In sum, we propose that LH quantity will increase surface acting significantly more among followers with higher power distance compared to lower power distance. As a result of increased surface acting, LH quantity will indirectly increase emotional exhaustion and decrease job satisfaction more so for followers with high power distance values.

Hypothesis 4: Follower power distance moderates the positive effect of leader humor quantity on follower surface acting such that the effect is stronger for high power distance followers and weaker for low power distance followers.

Hypothesis 5: Leader humor quantity's (a) positive indirect effect on follower emotional exhaustion and (b) negative indirect effect on follower job satisfaction, via surface acting, are moderated by follower power distance such that the indirect effects are stronger (weaker) for followers with high (low) power distance.

OVERVIEW OF STUDIES

We tested our hypotheses across three studies. Study 1 addresses internal and external validity by utilizing a field experiment where leaders were randomly assigned to a LH quantity intervention or control condition. Study 2 further improves internal validity with a laboratory experiment where we manipulate both LH quantity and power distance. Study 3 extends generalizability and provides an ecologically valid test of our theoretical model using a multiwave field study. After Study 3, we report exploratory analyses examining the independent and joint effects of LH quantity and LH quality across our studies. Data, syntax, and supplementary files for each study can be accessed in our online supplement:

https://osf.io/49sk7/?view_only=54c9e8e8a38c4b84ada20f41738441b5. In the online supplement, we also report two initial studies that show further support of our hypotheses.

STUDY 1 METHOD

Sample and Procedure

We designed a field experiment where we manipulated LH quantity and randomly assigned leaders, and their followers, to either a LH quantity intervention or a control condition. With the help of a large research university in Southern China, we recruited 100 part-time professional degree students who were managers at their respective organizations, as well as one of their subordinates, to participate in this study. The managers provided a list of their subordinates and how frequently they interacted with each follower. The researchers selected one

subordinate with whom the leaders interacted with the most. All leaders and their subordinates came from different organizations.

The leaders participating in this study were enrolled in a semester-long leadership course and participated in a class workshop that aims to improve leadership effectiveness. Classes were held once per week. The experiment was embedded in between two weeks of the course focused on improving leaders' ability to interact with followers (e.g., building relationships and soliciting input). Prior to the class that began the experiment, leaders completed a survey that included general leadership measures. They were told that the results from this survey would be used to provide them with instructions regarding their goal for the week to improve their leadership. They were also told that at the end of the week, their followers would complete a survey for feedback purposes. This design served as a guise for our experiment.

Leaders were randomly assigned to one of two conditions. Every leader was sent a standardized message with a goal and instructions ostensibly as feedback regarding their leadership effectiveness from the survey. In both the LH quantity and control condition, leaders were given the same goal to focus on improving the way they interact with their followers for the following week, and that they would be evaluated at the end of the week on their improvements. In the LH quantity condition only, we included additional instructions for leaders to engage in more humor expression because leadership research demonstrated the effectiveness of leader humor expression. Therefore, the messages to leaders in the control and humor conditions were kept the same, except that the control condition leaders were not given the instructions regarding humor. We purposely chose not to include any additional instructions in the control condition because we wanted those leaders to choose the strategies and lessons from the course that they

believed would be most effective or appropriate for improving interactions with their followers.²

In the leader humor quantity condition, leaders received the following instructions:

“Based on your results, we suggest that you improve the way you interact with your followers. The best way to do that is to express humor. In the following week, we suggest that you engage more frequently in humor expression when interacting with your followers. Leadership research has established the effectiveness of leader humor expression as one of the most effective ways to improve your leadership effectiveness. Please make sure you follow our suggestion, and we will be assessing the effectiveness of your improvement towards the end of the week.”

In the control condition, leaders received the following instructions:

“Based on your results, we suggest that you improve the way you interact with your followers. Please make sure you follow our suggestion, and we will be assessing the effectiveness of your improvement towards the end of the week.”

To encourage participation and salience of the manipulations, we sent text messages to leaders via mobile phones with their specific condition instructions three times: the night before the experiment started, at the beginning of the experiment, and in the middle of the week.

At Time 2, which was one week after the experiment began, the selected follower of each leader was asked to report their surface acting, emotional exhaustion, job satisfaction, positive affect, and deep acting. We measured LH quantity as a manipulation check from leaders to separate the source and to avoid informing followers about the nature of the experiment.

In this study, we accounted for LH quality through random assignment—that is, leaders’ pre-existing level of humor quality would be randomly distributed across conditions. This means that a leader with pre-existing high-quality humor is just as likely to be assigned to the LH quantity condition as a leader with pre-existing low-quality humor. We also conducted two

² We also conducted a supplemental study examining our LH quantity and control condition instructions (see [online supplement](#)). This supplemental study ensured that our LH quantity and control instructions did *not* create differences in (a) leaders’ motivation to follow the instructions or (b) leaders’ perceived efficacy of the instructions. This supplemental study also compared the Study 1 control condition instructions to two alternative and hypothetical control conditions that paralleled the language of our LH quantity condition but instead focused on alternative leadership behaviors, specifically leader inspiration and leader support (as opposed to leader humor). Our supplemental study showed that our Study 1 control condition had equivalence across key metrics (most critically, how much LH quantity leaders planned to enact) as these alternative control conditions.

additional checks to ensure that our manipulation did not increase low-quality humor. First, after the experiment, we contacted a few leaders in the LH quantity condition to ask how they followed the instructions to increase humor expression. These leaders commented on how they planned to tell amusing stories during meetings or how they researched jokes. Second, we conducted another supplemental study (full details in the [online supplement](#)) with a separate sample of leaders to see how leaders who were given these humor instructions would implement them. We recruited 100 leaders on Prolific (average age of 38.9, 33% female, 89% White) and gave them the LH quantity prompt from our field experiment. We coded the responses, and all but four leaders (who said they would do nothing) indicated intentional and effortful ways to increase and improve their humor expression. For example, leaders commented on how they would only use humor when appropriate and would not force it (17%), share humorous stories (17%), research and prepare jokes ahead of time (15%), and use specific types of humor (15%). Thus, leaders in our field experiment likely incorporated more humor expression (and higher quality humor) at work for the week. This is further likely since leaders in the field experiment were participating in a class workshop to improve their leadership and were being evaluated.

The final sample included 88 leader-follower dyads, of which 42 were in the control condition and 46 were in the LH quantity condition (Leader: $M_{\text{age}} = 31.92$, $SD_{\text{age}} = 4.09$, 60.5% male; Followers: $M_{\text{age}} = 30.83$, $SD_{\text{age}} = 5.10$, 48.86% male). All leaders had a 4-year university or higher-level degree, and 85.2% of the followers held a 4-year university or higher-level degree. Since participants spoke Chinese, all the questionnaires used in the study were backtranslated (Brislin, 1970). A native speaker first translated the materials from English to Chinese. After that, another native speaker translated the Chinese version back to English to confirm consistency and clarity.

Measures

All items for measures are listed in the [online supplement](#). Unless otherwise specified, we used a Likert scale from 1 = “*strongly disagree*” to 5 = “*strongly agree*.” All measures, except for power distance and wellbeing outcomes, referred to the previous week of work.

Leader humor quantity. Followers whose leader was assigned to the LH quantity (control) condition were coded as “1” (“0”). For the manipulation check, we used a slightly modified measure of the Cooper et al. (2018) three-item scale (sample item: “How frequently did you attempt to express humor with your followers at work, overall?”), which was rated by leaders ($\alpha = 0.90$).

Power distance. We used 3 items from the power distance values scale developed by Dorfman and Howell (1988; $\alpha = 0.62$). This scale has been widely used by scholars to measure power distance at the individual level (e.g., Cole et al., 2013; Ng et al., 2011).

Surface acting. We used an adapted version of the Discrete Emotions Emotional Labor Scale (Hu & Shi, 2015) to capture surface acting based on how often employees express three specific positive emotions – amusement, happiness, and interest – to leaders when they do not really feel that way (1 = “*almost never*” to 5 = “*very frequently*”; $\alpha = 0.91$). We chose these 3 emotion items because they are common emotional responses to humor (Mesmer-Magnus et al., 2012) and because employees experience and/or display these emotions frequently in the workplace (Basch & Fisher, 2000; Glasø & Einarsen, 2008).

Wellbeing outcomes. Followers reported their general *emotional exhaustion* using an established six-item scale (Wharton, 1993; $\alpha = 0.94$) and indicated their general *job satisfaction* with a modified three-item scale used in previous research (Tepper, 2000; $\alpha = 0.92$).

Control and alternative process variable. Because leaders and their followers were

randomly assigned to differing conditions of LH quantity, this helped account for concerns regarding leader characteristics (e.g., LH quality), their relationship with their subordinate (e.g., leader-member exchange), or demographic characteristics (e.g., gender or age) as omitted variables. Given this, we focused on controlling for follower *positive affect*. Past research has demonstrated that leader humor can have beneficial effects on wellbeing outcomes by increasing followers' positive affect. To account for this alternative mechanism, we measured followers' positive affect with three items: employees reported the frequency (1 = "*almost never*" to 5 = "*very frequently*") of which they felt the same positive emotions included in the surface acting scale ($\alpha = 0.87$). Results of hypothesis tests excluding positive affect do not substantially differ, and the regression results excluding this control are also reported in the main table. Lastly, although we have theoretical rationale to focus on surface acting, many emotional labor studies investigate both surface acting and deep acting (Grandey & Gabriel, 2015). Thus, we examined *deep acting* as an alternative mechanism with a three-item scale (1 = "*almost never*" to 5 = "*very frequently*") adapted from Brotheridge and Lee (2003) ($\alpha = 0.93$).

Analytical Approach

To test H1, we used analysis of variance (ANOVA) as well as a single-level regression analysis in Mplus 8.5 using maximum likelihood with robust standard errors (Muthén & Muthén, 2017) in which surface acting was regressed on the control variable and independent variables (see Model 1a of Table 3). To test H4, we used the same model as H1 but added the interaction term between LH quantity and power distance (see Model 2a of Table 3), and followed recommendations by Aiken and West (1991) to plot interactions and test simple slopes. To test H2, H3, and H5, we used a bias-corrected bootstrapping method with 10,000 resamples in Mplus 8.5 to test the unconditional and moderated indirect effects (Edwards & Lambert, 2007). For the unconditional indirect effects, we specified the path model as diagramed by Model 4 from Hayes

(2017) and included the control variable predicting both our mediator (surface acting) and two outcome variables (exhaustion and job satisfaction). For the moderated indirect effects, we tested the path model diagramed by Model 7 from Hayes (2017) with the inclusion of the control on the mediator and both outcome variables as specified for the unconditional indirect effects.

STUDY 1 RESULTS AND DISCUSSION

Descriptive statistics and correlations among variables are displayed in Table 1.

-----Insert Tables 1-4 and Figure 2 about here-----

We conducted confirmatory factor analyses (CFAs) to ensure the distinct factor structure of the variables. Given the larger number of items for exhaustion and because this is an established scale, we used item parceling (3 parcels) for this measure (Little et al., 2002, Little et al., 2013). A CFA of the six-factor model—power distance, surface acting, emotional exhaustion (3 parcels), job satisfaction, deep acting, and positive affect—demonstrated acceptable fit ($\chi^2(120) = 179.80$; CFI = .95; SRMR = .063; RMSEA = .075). It also had a better fit than all alternative models where we set different pairs of factors to correlate at [1.0]. For example, setting the correlation between surface acting and emotional exhaustion to 1.0 ($\chi^2(121) = 324.54$; CFI = .82; SRMR = .098; RMSEA = .138) resulted in a worse fit.

We conducted a manipulation check to assess the efficacy of our experimental manipulation of LH quantity. An ANOVA revealed a significant main effect of the LH quantity condition on reported LH quantity, $F(1, 86) = 12.55, p < .001, \eta_p^2 = .13$: leaders in the LH quantity condition expressed humor more frequently than their counterparts in the control condition ($M = 3.01, SD = .75$ versus $M = 2.46, SD = .69$).

Hypotheses tests. Means, standard deviations, and ANOVAs of the outcome variables are provided in Table 2. To test Hypothesis 1, an ANOVA revealed a significant main effect of LH

quantity on surface acting, $F(1, 86) = 5.50, p = .021, \eta_p^2 = .06$. Hypothesis 1 was supported as followers in the LH quantity condition reported higher surface acting during the week than followers in the control condition ($M = 2.15, SD = .88$ versus $M = 1.72, SD = .83$).

Regression results are presented in Table 3. LH quantity predicted surface acting ($\beta = .27$ [SE = .10], $p = .005$, Model 1a; $\Delta R^2 = .071$), supporting Hypothesis 1. Surface acting in turn positively predicted emotional exhaustion ($\beta = .43$ [SE = .08], $p < .001$, Model 4a; $\Delta R^2 = .162$) and negatively predicted job satisfaction ($\beta = -.24$ [SE = .10], $p = .013$, Model 6a; $\Delta R^2 = .052$). Indirect effect analyses indicated that LH quantity had unconditional indirect effects on follower emotional exhaustion and job satisfaction via follower surface acting (see Table 4). Hypotheses 2 and 3 were supported.

Regarding Hypothesis 4, an interaction between LH quantity and follower power distance predicted surface acting ($\beta = .31$ [SE = .13], $p = .016$, Model 2a; $\Delta R^2 = .053$). Simple slopes (see Figure 2) indicated that LH quantity condition positively related to surface acting when follower power distance was high (+1 SD: unstandardized B = .88 [SE = .25], $p < .001$) but was unrelated to surface acting when follower power distance was low (−1 SD: unstandardized B = .06 [SE = .24], $p = .79$). Hypothesis 4 was supported. Indirect effect analyses also provide support for Hypothesis 5a and 5b (see Table 4) such that LH quantity increased (decreased) emotional exhaustion (job satisfaction) via surface acting at high levels of power distance (+1 SD), LH quantity had no indirect effects at low levels of power distance (−1 SD), and the difference in those effects were significant.

Supplemental analyses. Results (see [online supplement](#)) showed the interaction of LH quantity \times power distance did not predict emotional exhaustion or job satisfaction. Also, positive affect and deep acting did not mediate the effects of LH quantity on wellbeing outcomes.

Study 1 Discussion

Using a field experiment with random assignment, which is often discussed as the gold standard of research studies because it strengthens both internal and external validity of the findings (Shadish, Cook, & Campbell, 2002), Study 1 found that LH quantity indirectly increased emotional exhaustion and reduced job satisfaction through surface acting, especially for followers with high power distance. LH quantity did not have main effects on wellbeing outcomes, however, for emotional exhaustion there was a direct increase of LH quantity, which may not be significant due to the lower sample size of the study (i.e., Study 1 may be underpowered). Despite Study 1's strengths, causal inference could be improved by exercising more control over the LH quantity manipulation (e.g., greater parallelism between the high and low conditions as well as each follower being exposed to the same humor stimuli), manipulating power distance, empirically modeling LH quality, and increasing the sample size. To this end, we conducted Study 2.

STUDY 2 METHOD

Participants and Procedure

This experiment was pre-registered: https://aspredicted.org/BNR_34X. We recruited participants from a behavioral lab subject pool at a business school in the northeastern United States. Participants participated in a ~30-minute study in exchange for \$12. We conducted a 2 (LH quantity: low vs. high) \times 2 (power distance: low vs. high) between-subjects design. We aimed to recruit 230-260 participants, which was the estimated maximum number of participants available for our study. At the time of the pre-registration, 238 participants had signed up, and a total of 212 participants showed up and completed the experimental study.

Participants registered online for an hour-long experimental session timeslot. When participants arrived at the building, they were directed into a computer lab to complete a consent

form. Participants were told that they would be participating in a focus group for the university's bookstore, and that they would participate in groups of 3-5 participants. At this point, the experimenter randomly selected 3-5 participants and escorted them to a separate breakout room to participate in the focus group. Then, after the first focus group finished, these participants were brought back to the computer lab to complete a survey, while the next group of participants (who were waiting in the computer lab) were then escorted to the breakout room to participate in the focus group. This happened three times within the hour timeslot such that all participants participated in the focus group and completed a survey afterwards. While participants waited in the computer lab for their focus groups, they were told they could use their smartphones or the lab computer. We aimed to have the focus group size be 5 participants, but we varied the number between 3-5 depending on the number of participants that registered and the number of noshows. The final sample included 2 groups of two, 14 groups of three, 13 groups of four, and 20 groups of five participants. As we note below, controlling for group size did not alter our results.

Upon entering the breakout room for the focus group, participants were greeted by the focus group leader (i.e., the focus group moderator). The leader introduced himself as the Vice President of Sales for the university bookstore. However, he was a professional actor (age between 50-60) trained to deliver our LH quantity manipulation and part of the power distance manipulation. We used the guise of a focus group to test our hypotheses for several reasons: (a) it allowed us to have a leader-led group in a short amount of time; (b) it facilitated face-to-face interaction for our phenomenon to occur; (c) it allowed the leader to use a script in real time (it is common for focus group moderators to use scripts); (d) it added mundane realism; (e) it was believable to participants; and (f) it allowed us to more naturally manipulate power distance.

At the start of the focus group, participants in all conditions were asked to read a task overview sheet, which described the purpose of the focus group and provided participants with the focus group norms. It also contained the major component of our power distance manipulation. To manipulate a cultural value like power distance, we relied on the group norms approach, which has been used to manipulate other cultural values (individualism versus collectivism) in lab settings (Goncalo & Staw, 2006; Goncalo & Duguid, 2012). We used the key features of Goncalo and colleagues' manipulation: (1) we established a clear group norm; and (2) we reinforced the norm with descriptive (e.g., hundreds of focus groups have followed these norms) and prescriptive information. The two versions of the task overview sheets are found in the [online supplement](#). Consistent with its conceptualization, in the high-power distance condition (high PD), the task overview sheet instructed participants as follows:

It is critical to treat the moderator as a respected authority figure who is leading you in the session. When the moderator moves the conversation to a new topic, follow their lead and do not return to a prior topic. Please only speak up or provide feedback when you are called on. It is completely inappropriate to challenge or question the moderator during the session.

The low-power distance condition (low PD) task overview sheet instructed participants:

It is critical to treat the moderator as an equal who is participating in the session with you. When the moderator moves the conversation to a new topic, feel free to change the subject and return to a prior topic. Please speak up or provide feedback at any point. It is completely appropriate to challenge or question the moderator during the session.

To further emphasize the power distance manipulation, we used the focus group leader because leaders have a highly visible and strong influence on group norms (Feldman, 1984). First, we used the visual appearance of the leader to signal high versus low power distance. For the high PD, the moderator wore formal business attire (a navy-blue suit and tie); by contrast, the leader wore business casual attire (navy-blue polo shirt) in the low PD. Second, we modified the leader's script slightly across the two conditions to reinforce the differences in power distance norms. The full scripts are provided in Study 2 Materials in the [online supplement](#). In the high

PD, the leader introduced himself as “Mr. last name,” reiterated that he was “leading this session,” instructed participants to “talk to me like you’re talking to a formal representative from the bookstore,” and thanked the group “for their respect” at the halfway point. In the low PD, the leader introduced himself with his first name, reiterated that “we’re all equals here,” instructed participants to “talk to me like you’re talking to a friend from the bookstore,” and thanked the group “for their participation” at the halfway point.

We developed our LH quantity manipulation based on existing humor research (Bitterly & Schweitzer, 2019; Evans, Slaughter, Ellis, & Rivin, 2019; Gloor et al., 2021). We altered the focus group leader’s script to contain jokes (high LH quantity) or no jokes (low LH quantity). In line with humor manipulations (Bitterly & Schweitzer, 2019; Gloor et al., 2021), we used puns as jokes due to their simplicity, affiliative nature, and quality. The differences in the two LH quantity scripts are displayed in Appendix A, and the full scripts for each condition are in the Study 2 Materials in the [online supplement](#). To ensure that participants recognized the humor expression, we had the actor either use two punchlines in a row or repeat each joke (we learned in pilot testing that most participants did not recognize that humor was being expressed when the actor only said the pun once). We also instructed the leader to laugh briefly at his joke to emphasize the humor (e.g., see Evans et al., 2019). These elements of repeating punchlines, using two punchlines in a row, and/or using humor meta-communication, such as laughing when expressing humor or using language to notify humor is being expressed (e.g., “get it?”), are not only fundamental elements of expressing humor, but they are also used by people to signal or bring emphasis to the fact that humor is being attempted (Attardo et al., 1994; Canestrari, 2010; Norrick, 2003; Pickering et al., 2009). In other words, their use is more likely in situations when it is more uncertain or ambiguous to an audience whether humor is occurring, such as when the audience is not expecting humor, does not know the humor expresser, and the jokes are subtle.

All of these situational aspects were present in our experiment; hence, we incorporated these elements that emphasize humor expression in our study (we should also note that, as discussed below, that our LH quantity manipulation significantly increased perceptions of LH quality).

Across all four conditions, the leader maintained the same neutral demeanor and tone.

Furthermore, we gave the leader rules to follow when asking the focus group questions and listening to participants' answers. For the first question of the focus group ("Has anyone been to the bookstore, and if so, what did you purchase?"), the leader let each participant speak if they desired. For all other scripted questions, he let no more than two participants speak and varied who the participants were for each question. We instructed the leader to call on individual participants to answer the focus group questions. In addition, the leader was trained to respond to each answer from participants with one of three short acknowledgments: "mmhmm," "ok," and "very well." We chose these expressions because they were expressions that the actor used naturally. Besides these three responses, the actor did not provide any other non-verbal feedback to participants (e.g., he kept a neutral tone and reaction and did not smile or nod). Given the task and emotional labor demands for the actor, we kept the focus group sessions very short (approximately 5 minutes). We also conducted three separate training sessions in advance of the experiment where the actor could practice each script and all rules with live participants.

After completing the focus group, participants returned to the computer lab to complete a survey of follow-up questions (this survey contained the measures listed below). In total, we had 212 participants show up and participate in the experiment. A research assistant watched the focus groups live from CCTV to ensure the actor followed the script and to note any issues that arose in the sessions. From this monitoring, three focus groups were invalid (out of 52), and we removed 14 participants' data from the study. Two of the focus groups were invalid because the actor deviated significantly from the script in which he mixed conflicting conditions, and the

third group was invalid because a participant took over the focus group session and facilitated the conversation even though the actor tried to stop this participant. We note that including the full sample of participants does not substantially change the results of hypotheses tests.

In our survey, we included two attention checks plus an open-ended question. As we noted in our pre-registration, we planned to only remove participants if they failed an attention check *and* failed the open-ended response question. Although 21 participants missed an attention check, none of them failed the open-ended response question; thus, we retained these participants (removing these participants from the sample does not substantially change the results of our hypotheses tests). In total, we had full data for 198 participants across 49 focus groups ($M_{\text{age}} = 22.22$, $SD_{\text{age}} = 6.92$; Gender: 24.8% male, 73.7% female, 0.5 other, 1% prefer not to say; Race and Ethnicity: 42.9% Asian, 11.6% Black or African American, 7.6% Hispanic or Latino, 29.3% White, 7.6% Mixed, and 1.0% Other; 87.9% were students and 12.1% were local employees).

Measures

Unless otherwise noted, measures were rated on a Likert scale ranging from 1 = “*very much disagree*” to 7 = “*very much agree*.” All measures are listed in the [online supplement](#).

Follower surface acting. For generalizability and its relevance, we used the seven-item Diefendorff et al. (2005) measure of surface acting and adapted it to our focus group context (1 = “*none*” to 7 = “*completely*”; $\alpha = 0.94$).

Follower wellbeing outcomes. We adapted the scales for *emotional exhaustion* ($\alpha = 0.90$) and *work satisfaction* ($\alpha = 0.95$) that we used in Study 1 to this study’s context.

Controls and testing alternative explanations. To test whether our manipulations influenced other variables that could drive outcomes in our model, we measured several additional variables. Critically, we examined *leader humor quality* by having participants report

how funny they found the group leader to be using three items that were based on existing scales that capture the quality (or perceived funniness) of a person (e.g., “The moderator had a good sense of humor”; Decker, 1987; see also Wanzer, Booth-Butterfield, & Booth-Butterfield, 1995; $\alpha = 0.95$). We also examined the robustness of our hypothesized relationships examining the quality of participants’ perceived relationship with the group leader. To do so, we captured followers’ ratings of their *leader-member exchange* (LMX) using five items from the LMX scale developed by Scandura, Graen, & Novak, (1986) ($\alpha = 0.88$) adapted to our study’s context. Not only does LMX relate to leader humor (Cooper et al., 2018), but LMX also correlates with surface acting, emotional exhaustion, job satisfaction, as well as a number of other favorable work attitudes (e.g., Dulebohn et al., 2012; Fisk & Friesen, 2012; Schermuly & Meyer, 2016). Similar to Study 1, we captured positive affect (1 = “*not at all*” to 5 = “*a lot*”; $\alpha = 0.83$) and deep acting (1 = “*not at all*” to 5 = “*a lot*”; $\alpha = .90$) using the same items from Study 1 adapted to Study 2’s context. Lastly, to examine whether the hypothesized effects are unique to LH quantity, as opposed to any positive lead behavior, we measured charismatic leadership with three items from the vision and articulation dimension (Conger & Kanungo, 1994; $\alpha = 0.83$). We focused on charismatic leadership due to its conceptual relevance to our model (Goswami et al., 2016; Hughes & Avey, 2009; Menges et al., 2015).

STUDY 2 RESULTS AND DISCUSSION

Descriptive statistics and correlations among variables are displayed in Table 5.

-----Insert Tables 5-7 and Figure 3 about here-----

A CFA of the hypothesized five-factor model composed of surface acting (3 parcels), exhaustion (3 parcels), satisfaction, positive affect, and LH quality demonstrated good fit to the data ($\chi^2 (80) = 157.67$; CFI = .97; SRMR = .043; RMSEA = .070). It also had better fit than all

alternative models that set different pairs of factors to correlate at [1.0]. For example, setting the correlation between emotional exhaustion and satisfaction to -1.0 ($\chi^2(81) = 401.99$; CFI = .89; SRMR = .080; RMSEA = .141) resulted in worse fit.

Manipulation checks. We adapted the Cooper et al. (2018) LH quantity scale to this study's context ($\alpha = 0.96$) as a manipulation check (sample item: "The moderator expressed humor with you"). An ANOVA revealed our LH quantity manipulation was effective as participants in the high LH quantity condition reported that the leader used humor more frequently than the low LH quantity condition ($M = 4.10$, $SD = .87$ versus $M = 1.78$, $SD = .92$; $F(1, 194) = 339.87$, $p < .001$, $\eta_p^2 = .64$). To assess the efficacy of our power distance manipulation, we adapted seven items from the Kirkman et al. (2009) scale to refer to the focus group norms ($\alpha = 0.87$). An ANOVA also indicated the efficacy of this manipulation as participants in the high PD condition reported higher power distance ($M = 3.42$, $SD = .76$) than in the low PD condition ($M = 2.28$, $SD = .63$; $F(1, 194) = 130.83$, $p < .001$, $\eta_p^2 = .40$). There were no other direct or interaction effects on the manipulation checks, except for one: the power distance manipulation had a small effect on perceptions of humor frequency ($F(1, 194) = 6.64$, $p = .011$, $\eta_p^2 = .03$) in which low PD participants reported more LH quantity ($M = 3.12$, $SD = 1.50$) than high PD participants ($M = 2.77$, $SD = 1.42$).

We also examined whether our experimental conditions created any significant differences in the alternative explanation variables collected (i.e., LH quality, LMX, charismatic leadership, positive affect, and deep acting). The only variable where there was large significant differences across conditions was LH quality, which the LH quantity condition had a positive effect (difference: 1.54; $F(1, 194) = 55.23$, $p < .001$, $\eta_p^2 = .22$) and power distance condition had a negative effect (difference: -0.74; $F(1, 194) = 12.75$, $p < .001$, $\eta_p^2 = .06$). The only other

differences of experimental conditions on any of these variables were small effects of power distance on LMX ($F(1, 194) = 4.81, p = .03, \eta_p^2 = .02$) and on charismatic leadership ($F(1, 194) = 8.36, p = .004, \eta_p^2 = .04$) as well as a small, positive effect of LH quantity condition on positive affect ($F(1, 194) = 4.91, p = .03, \eta_p^2 = .025$). Thus, for both theoretical and empirical reasons, we report the results controlling for LH quality and positive affect. As an additional check (see supplemental analyses below), we re-tested all hypotheses controlling for LMX and leader charisma, and the results remain substantially unchanged.

Hypotheses tests. The means and standard deviations of the outcome variables are provided in Table 6. Individuals were nested within focus groups. To account for this nesting, we analyzed the data in Mplus 8.5 using cluster-robust standard errors, which is a technique recommended to analyze a single-level model while fully addressing the clustered nature of the data (McNeish, Stapleton, & Silverman, 2017). The results of these analyses are reported in Table 7. For each hypothesis, we specified the model paths as outlined in Study 1, and continuous predictor variables were grand-mean centered.

Hypothesis 1 was supported as LH quantity condition predicted surface acting ($\beta = .30$ [SE = .07], $p < .001$, Model 1a; $\Delta R^2 = .068$). Surface acting positively predicted emotional exhaustion ($\beta = .34$ [SE = .08], $p < .001$, Model 4a; $\Delta R^2 = .096$) and negatively predicted satisfaction ($\beta = -.15$ [SE = .07], $p = .029$, Model 6a; $\Delta R^2 = .019$). Indirect effect analyses using the same approach as Study 1, but also accounting for the clustered standard errors (i.e., accounting for the nesting), indicated that LH quantity had unconditional indirect effects on follower emotional exhaustion and job satisfaction via follower surface acting (see Table 4).

Hypotheses 2 and 3 were supported.

There was a significant interaction between LH quantity and power distance conditions predicting surface acting ($\beta = .36$ [SE = .10], $p < .001$, Model 2a; $\Delta R^2 = .043$). Tests of simple slopes (e.g., see Figure 3) indicated that LH quantity increased surface acting for followers in the high PD condition (unstandardized B = 1.35 [SE = .25], $p < .001$), but it did not significantly relate to surface acting for followers in the low PD condition (unstandardized B = .24 [SE = .15], $p = .11$). Thus, Hypothesis 4 was supported. Moderated indirect effect results (see Table 4) supported Hypothesis 5a and 5b: LH quantity increased and decreased emotional exhaustion and job satisfaction, respectively, via surface acting at high levels of power distance, LH quantity had no indirect effects at low levels of power distance, and the differences in these effects were significant.

Supplemental analyses. We conducted several supplemental analyses for Study 2 (see [online supplement](#)). First, we tested the LH quantity \times power distance interaction directly predicting emotional exhaustion and job satisfaction variables. None of these interaction effects were significant. Second, adding controls of group size or LMX and charismatic leadership, or removing positive affect or LH quality as controls, do not substantially alter the results of hypotheses tests. Third, we explored positive affect and deep acting as alternative mechanisms of LH quantity's effects, and neither of these variables mediated LH quantity's effects on outcomes. Fourth, we examined whether charismatic leadership produced similar effects as LH quantity. However, charismatic leadership, or its interaction with power distance, did not significantly predict surface acting. Finally, LMX did not moderate the effects of LH quantity condition on surface acting or wellbeing outcomes.

Study 2 Discussion

Study 2 complemented the findings from Study 1. In an experiment that modeled real interactions among a leader and followers as well as manipulated LH quantity and power

distance, we found support that LH quantity increases surface acting, which subsequently impairs follower wellbeing outcomes. We also found that these effects occur for followers with high power distance values, but not low power distance values. Our LH quantity condition had a significant main effect on emotional exhaustion, but not job satisfaction. Interestingly, we found these harmful effects of LH quantity on surface acting and downstream wellbeing outcomes even though this LH quantity manipulation also had a strong positive effect on LH quality. This finding suggests that even when followers perceive a leader as genuinely funny, frequent leader humor expression heightens positive display expectations that drive surface acting and aggravate wellbeing, especially for followers with higher power distance (we delve deeper into the interplay of LH quantity and LH quality in our exploratory analyses section after Study 3). In Study 3, we looked to further increase the external validity of the findings from Studies 1 and 2 by testing our model in a field setting that captured our focal variables over time in general (rather than the shorter time frames of one week in Study 1 and one meeting in Study 2).

STUDY 3 METHOD

Sample and Procedure

We conducted a pre-registered field study (https://aspredicted.org/DMG_B44) in a technology firm based in Southern China. The data were collected in three waves to help minimize concerns regarding common method variance (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). This organization had a project-based structure where employees worked with multiple managers and coworkers that changed across projects. Given this, team membership was more fluid, but each focal project had a respective manager who reviewed performance, gave feedback, and offered general coaching to employees who were on that project.

At time 1, employees completed a paper survey on LH quantity, LH quality, follower power distance, leader control variables, and demographics (age and gender). We had followers

complete survey questions about leaders whom they considered to be their focal manager at the point of the study. At time 2, 5 workdays later, employees completed another paper survey on their surface acting, positive affect, emotional exhaustion, job satisfaction, and deep acting. At time 3 (about a week after time 2), the company conducted its own “pulse survey.” Specifically, the company periodically assesses employees’ wellbeing and satisfaction by having employees complete short pulse surveys. Two of their pulse survey questions aligned with our two dependent variables: employees’ overall energy level (which we reverse coded for exhaustion) and satisfaction with the job (which we used for job satisfaction). We timed our data collection in relation to this company pulse survey such that our time 2 survey occurred at least one week before the pulse survey. To preserve the anonymity of employees, company HR personnel collected the paper surveys from the earlier two time points, matched them with the pulse survey data (i.e., company recorded data), and then provided the anonymized data to the research team. Then, two research assistants entered the data into a spreadsheet and checked for accuracy. We aimed to collect data from as many of the ~150 employees possible at the company. In total, 126 participants provided completed responses across all time points and sources (no attrition across time points). The average age of employees was 32.83 ($SD = 5.06$) and 43% were female. Since participants spoke Chinese, all the questionnaires used in the study were back-translated using the same procedure outlined in Study 1.

Measures

The complete list of Study 3 measures and items is in the [online supplement](#). Unless otherwise specified, all measures were rated on a five-point Likert scale ranging from 1 = “*very much disagree*” to 5 = “*very much agree*.”

Leader humor quantity. We measured LH quantity using the three-item scale developed by Cooper et al. (2018) (Sample item: “How frequently does your leader express humor with you

at work, overall?"; 1 = "*almost never*" to 5 = "*very frequently*"; $\alpha = 0.92$). We used a lead-in that communicated our desire to capture the *quantity* of humor expression by leaders regardless of whether followers think the leader's humor is funny or not.

Power distance. We captured power distance with the six-item scale developed by Dorfman & Howell (1988) ($\alpha = 0.63$).

Surface acting. We used a slightly modified version of the surface acting scale from Study 1 that asked employees to report how much they express or exaggerate three different feelings (amusement, happiness, and interest) in interactions with their leader when they do not really feel that way (1 = "*almost never*" to 5 = "*very frequently*"; $\alpha = 0.96$).

Wellbeing outcomes. We obtained two sources of wellbeing outcomes. The primary source came from the company's records (pulse survey). In time 3, the organization had employees report their overall energy level at work using one item that ranged from 1 = "*very exhausted*" to 10 = "*very energetic*." We reverse coded this item for emotional exhaustion. The organization also had employees report their job satisfaction at time 3 using one item (1 = "*not at all satisfied*" to 10 = "*very satisfied*"). As another source, we had employees report in the time 2 survey their *emotional exhaustion* ($\alpha = 0.92$) and *job satisfaction* ($\alpha = 0.92$) using the same scales as Study 1.

Control variables. For the rationale provided in Study 2, we controlled for *leader humor quality* using four items (e.g., "My leader has a good sense of humor"; $\alpha = 0.92$), LMX using the seven-item LMX scale developed by Scandura, Graen, & Novak, (1986) ($\alpha = 0.93$), and followers' positive affect using the same three items as Studies 1 and 2 (i.e., "amusement," "happiness," and "interest"; 1 = "*almost never*" to 5 = "*very frequently*"; $\alpha = 0.93$). Consistent with past research (Yam, Christian, Wei, Liao, & Nai, 2018), we also controlled for

demographics of follower age and gender as well as leader gender. Results of hypothesis tests excluding controls do not substantially differ (e.g., see the main regression table).

Testing alternative explanations. Consistent with Study 2, we examined deep acting and charismatic leadership to rule out alternative explanations (see supplemental analyses below). We examined *deep acting* as an alternative mechanism with the same scale as Study 1 (1 = “almost never” to 5 = “very frequently”; $\alpha = 0.95$). We measured charismatic leadership with the same scale as Study 2 adapted to this study’s context ($\alpha = 0.90$).

STUDY 3 RESULTS AND DISCUSSION

-----Insert Tables 8 & 9 and Figure 4 about here-----

Table 8 presents the means, standard deviations, and correlations among variables. We first conducted a CFA. Given the larger number of items for exhaustion, power distance, and LMX, and because these are established scales, we used item parceling (3 parcels each) for these measures (Little et al., 2002, Little et al., 2013). The hypothesized nine-factor model (LH quantity, LH quality, power distance, surface acting, exhaustion, job satisfaction, LMX, deep acting, and positive affect) demonstrated adequate fit to the data, ($\chi^2 (314) = 545.46$; CFI = .93; SRMR = .054; RMSEA = .076). When testing alternative models in which we constrained all possible pairs of factors to correlate at [1.0], all such models had worse fit. For example, the model setting the correlation between positive affect and job satisfaction to 1.0 ($\chi^2 (315) = 741.43$; CFI = .87; SRMR = .063; RMSEA = .104) had worse fit.

Hypotheses tests. To test our hypotheses, we used the same analytical approach and path models as outlined in Study 1. Regression results (see Table 9) suggested that LH quantity was positively associated with follower surface acting ($\beta = .32$ [SE = .11], $p = .002$, Model 1a; $\Delta R^2 = .079$), providing support for Hypothesis 1. Follower surface acting positively related to the

company-recorded emotional exhaustion ($\beta = .16$ [SE = .08], $p = .037$, Model 4a; $\Delta R^2 = .018$) but did not relate to the company-recorded job satisfaction ($\beta = .04$ [SE = .08], $p = .59$, Model 6a; $\Delta R^2 = .001$). Results of indirect effects analyses (see Table 4) indicated that LH quantity had an unconditional positive indirect effect on emotional exhaustion but not job satisfaction. We replicated these results for the employee survey measures of emotional exhaustion and job satisfaction (see Table 4). Thus, only Hypothesis 2 (but not Hypothesis 3) was supported.

Next, results indicated that the interaction term between LH quantity and follower power distance predicted surface acting ($\beta = .16$ [SE = .07], $p = .015$, Model 2a; $\Delta R^2 = .023$). As illustrated in Figure 4, simple slope tests showed that LH quantity had a positive effect on follower surface acting at high levels (+1 SD; unstandardized B = .49 [SE = .14], $p < .001$), but no effect at low levels (−1 SD; unstandardized B = .12, [SE = .13], $p = .385$) of follower power distance, supporting Hypothesis 4.

As summarized in Table 4, indirect effects analyses indicated support for Hypothesis 5a for both the company recorded and employee survey measures of emotional exhaustion: LH quantity had (a) positive indirect effects at high levels of power distance (b) no significant indirect effects at low levels of power distance, and (c) the differences among the effects at high and low levels were significant. However, Hypothesis 5b was not supported for job satisfaction.

Supplemental analyses. The results of our supplemental analyses are in the [online supplement](#). First, the LH quantity \times power distance interaction did not directly predict emotional exhaustion nor job satisfaction variables. Second, hypotheses tests without control variables remained substantially unchanged. We also ran a separate analysis that retested hypotheses excluding only positive affect but including the other controls, and results were unchanged. Third, we explored positive affect and deep acting as alternative mechanism of LH quantity's

effects. LH quantity negatively predicted positive affect, and LH quantity had harmful, unconditional indirect effects on both wellbeing outcomes via positive affect. Deep acting did not mediate LH quantity's effects. Fourth, charismatic leadership, or its interaction with power distance, did not predict surface acting. Fifth, because employees reported all variables in our model, we conducted two tests to examine bias due to common method, and results of these tests are inconsistent with common method bias.

Study 3 Discussion

Study 3 provides support for most hypotheses. Although LH quantity did not have main effects on follower wellbeing outcomes³, it directly increased surface acting and indirectly increased emotional exhaustion (two different measures of it), and these effects were significantly stronger for followers with high power distance. By contrast, surface acting did not predict job satisfaction. Also, and interestingly, LH quantity significantly and positively related to LH quality, a theme we return to in our following section that explores the interplay of LH quantity and LH quality. Overall, Study 3 findings, which include a robust set of controls and supplementary analyses, provide support that LH quantity can increase follower surface acting and subsequently harm follower wellbeing, especially for followers with high power distance.

Study 3 also complements findings from Studies 1 and 2 by enhancing external validity. Despite its strengths, Study 3 is limited in several ways (e.g., correlational and all variables rated by employees) that we acknowledge and address in our limitations section of our general discussion.

³ LH quantity predicting job satisfaction appears to be an inflation effect created by controlling for positive affect; this direct effect of LH quantity on job satisfaction disappears when excluding positive affect as a control.

Exploration of LH Quantity and LH Quality Effects on Follower Emotional Outcomes

An important and open question from our investigation is how LH quantity and LH quality relate to one another and, in combination, influence follower emotional and wellbeing outcomes. We examine this question across three studies that have both LH quantity and LH quality measures: Study 2, Study 3, and Supplemental Study B. The full details of these results are reported in the “Exploratory Analyses of LH Quantity and LH Quality” in the [online supplement](#). To begin, we used a mini meta-analysis (Goh, Hall, & Rosenthal, 2016) to examine the correlations for LH quantity versus LH quality. The mini meta-analytic correlations showed how LH quality has consistent emotional and wellbeing benefits for followers; by contrast, LH quantity mostly harms followers: LH quantity had significant positive mini meta-analytic correlations with follower emotional exhaustion and surface acting, and it had non-significant mini meta-analytic correlations with job satisfaction and positive affect. This mini meta-analytic summary helps support our argument that studies only investigating LH quality can be problematic because they fail to consider how LH quantity has a downside to followers’ emotional outcomes.

Next, we examined how LH quantity and LH quality, in combination, might impact follower emotional and wellbeing outcomes. In our exploratory analyses, we tested three theoretical possibilities: (a) an interaction model, (b) a mediation model, and (c) an additive model. First, LH quality may moderate the relationships among LH quantity and outcomes such that LH quantity benefits outcomes when quality is high, but worsens outcomes when quality is low. We tested this interaction effect of LH quantity and LH quality on all possible outcomes (surface acting, emotional exhaustion, and job satisfaction). All these interaction effects were insignificant. We also investigated three-way interactions of LH quantity, LH quality, and power

distance. Again, we found no consistent support for any three-way interaction effects. Thus, we find no support for LH quality as a moderator of the effects of LH quantity.

Second, we examined whether LH quantity mediates the effects of LH quality.

Theoretically, a person's humor quality, that is, a person who is seen as funny/amusing and has a good sense of humor, is an antecedent to how much that person expresses humor. People who are funny and have a good sense of humor will express humor more frequently because they have historical evidence and feedback of their funniness and comedic quality (e.g., past experiences where people laugh at their humor); thus, they will be confident and willing to express humor more frequently. Although not explicitly studied, there is some evidence that funny people express humor more often (Wanzer, Booth-Butterfield, & Booth-Butterfield, 1995). We find some support for partial mediation such that LH quality indirectly increases surface acting via LH quantity (Study 3 and Supplemental Study B) and indirectly increases emotional exhaustion via LH quantity (Supplemental Study B). Furthermore, we find that LH quality has a serially mediated, positive indirect effect on emotional exhaustion (LH quality → LH quantity → surface acting → emotional exhaustion) in Study 3 and Supplemental Study B. Altogether, this evidence suggests that there may be an unintended downside of humorous leaders—funny leaders may express humor more frequently, which causes increased surface acting and exhaustion.

Third, a final theoretical possibility is that LH quantity and LH quality may exert additive (i.e., independent) effects on follower emotional and wellbeing outcomes. We also find support for this model such that (a) in each study, we find the hypothesized, harmful effects of LH quantity; and (b) we find the following beneficial independent effects of LH quality: LH quality directly reduces surface acting (Study 2 and Supplemental Study B) and emotional exhaustion (Study 2) as well as directly increases positive affect (Study 2) and job satisfaction

(Supplemental Study B). LH quality also has beneficial indirect effects via surface acting on emotional exhaustion (all studies) and job satisfaction (Study 2 and Supplemental Study B) and via positive affect (Study 2). Therefore, results indicate that accounting for both dimensions of leader humor in an empirical model – quantity and quality – produces independent harmful effects (quantity) and beneficial effects (quality) on follower wellbeing outcomes. We discuss the conclusions and implications of these findings for existing and future research below.

GENERAL DISCUSSION

What do a field experiment, lab experiment, and multi-wave field study all have in common? This is not the start to a bad joke but to important findings from our studies: LH quantity increases follower surface acting and indirectly worsens their emotional exhaustion and job satisfaction, and these effects are stronger for followers with high (compared to low) power distance. As such, our studies offer several meaningful theoretical contributions.

Theoretical Implications

Our core theoretical contribution lies in challenging the general consensus that leader humor is a good thing for follower emotional wellbeing (Cooper et al., 2018; Decker, 1987; Decker & Rotondo, 2001; Mesmer-Magnus et al., 2012). We demonstrate how (a) focusing on LH quantity (as opposed to LH quality), (b) specifying surface acting as a mechanism, and (c) accounting for follower power distance help provide a counterpoint to leader humor's wellbeing effects. As such, we detail why and when LH quantity can have unintended *negative* consequences for follower emotional exhaustion and job satisfaction.

Importantly, our research also demonstrates the value of distinguishing LH quantity from LH quality. Our studies and exploratory analyses of LH quantity and LH quality find generally consistent (and independent) harmful effects from LH quantity and beneficial effects from LH quality on follower wellbeing outcomes. These results help to explain why most past studies,

which have focused primarily on LH quality variables, find leader humor to increase follower positive affect (e.g., Goswami et al., 2016; Wijewardena et al., 2017) and wellbeing outcomes (e.g., Hughes & Avey, 2009; Kong et al., 2019; Mesmer-Magnus et al., 2012; Robert & Wilbanks, 2012; Vecchio et al., 2009). In fact, our findings suggest that studies that only examine LH quality without differentiating it from LH quantity may produce positively inflated estimates for the beneficial effects of leader humor expression. Instead, research examining both LH quantity and LH quality simultaneously (or research that explicitly and clearly focuses on one dimension) can help specify more accurate estimates and theoretical predictions.

To this end, our results consistently demonstrate across studies that LH quantity increases surface acting and worsens follower wellbeing outcomes indirectly through surface acting even while controlling for LH quality's effects. Therefore, an important contribution of our work is the finding that LH quantity harms followers by increasing their surface acting. This finding is even more noteworthy in light of the positive association between LH quantity and LH quality in our studies. As our exploratory analyses indicated, funny or humorous leaders joke more, which can have an unintended and harmful effect – LH quality increases LH quantity, which then subsequently increases follower surface acting and downstream emotional exhaustion. An implication of these results is that leaders likely experience primarily positive reactions from followers in response to their humor, whether it be through surface acting or genuine amusement. This consistent positive feedback could encourage leaders to tell even more jokes, triggering a potential negative cycle where leaders express more humor, which backfires and causes more surface acting. Instead, our research suggests that leaders should focus on fewer, higher-reward attempts to maximize the benefits of humor quality and minimize the harm of humor quantity.

We also contribute by introducing an emotion regulation perspective to leader humor research. We identify surface acting as a novel mechanism of LH quantity's effects on follower

wellbeing. Not only does past research generally assume that leader humor expression induces genuine positive emotions in followers (Kong et al., 2019), but it also overlooks the possibility that followers are likely to fake or exaggerate positive emotions in response to leader humor expression. Our studies demonstrate the problem with this oversight. LH quantity consistently increased surface acting, and surface acting mediated LH quantity's effects on follower emotional exhaustion and job satisfaction, even while controlling for positive affect. By contrast, when accounting for both LH quantity and LH quality separately, LH quantity only predicted follower positive affect in Study 3, and the association was negative.

Another contribution to leader humor research involves identifying power distance as a moderator of LH quantity's effects on surface acting and downstream follower wellbeing outcomes. Our studies' findings suggest that high power distance followers find leader humor expression unexpected and counter-normative, and they surface act to affirm their leaders and uphold positive display expectations. By contrast, leader humor expression is consistent with low power distance followers' normative expectations, which leads to less surface acting.

Our final contribution to the leader humor literature, and to the humor literature more broadly, is that we encourage an expanded discussion of the potential risks of humor. Existing research recognizes that humor, in general, is risky because the target may not find the *content* of the joke or stimuli amusing (i.e., the humor is of low quality), focusing in particular on what kinds or types of humor is riskier for target reactions (Bitterly, 2022; Robert et al., 2016; Romero & Cruthirds, 2006; Warren et al., 2021). We expand this discussion by identifying *how much* people say (humor quantity) and *to whom* they say it (follower power distance) as additional risks that determine target affective reactions.

Beyond leader humor research, we also contribute to the emotional regulation literature by identifying how a positively intended leadership behavior can lead to follower surface acting.

A core theoretical viewpoint of emotional regulation research is that, in the context of positive display rules, surface acting most often occurs when a negative event or interaction makes employees feel bad, but they are expected or required to display positive emotions. For example, negative interactions, such as an employee being mistreated by a customer or boss (e.g., Carlson, Ferguson, Hunter, & Whitten, 2012; Rupp & Spencer, 2006), increase the discrepancy between what an employee feels (e.g., negative) and the positive display standard (i.e., displaying positive emotions) and hence increase surface acting. Grandey and Melloy (2017: 413) summarize this view of emotional labor research: “surface acting seems to be a way to cope with negative interpersonal situations.” In this context, we provide theory and empirical results as to why and when a positive behavior—specifically, leaders expressing humor to bring amusement to a follower—can elevate levels of surface acting because it increases the positive display demands that employees need to uphold, especially for followers with high power distance.

Practical Implications

Our studies inform leaders to be more mindful of how often they express humor. We find that frequent leader humor expressions, even when they are deemed funny, can cause followers to surface act, worsening their emotional exhaustion and job satisfaction. The implication is that many leaders likely need to scale back on how frequently they tell jokes or express humor. Instead, leaders should focus on fewer, higher-reward attempts to utilize the benefits of humor quality while avoiding the harm of excess humor quantity. This advice is particularly important given leaders’ superior power over followers; not only are leaders less inhibited and more likely to express humor frequently at work (Coser, 1960; Robinson & Smith-Lovin, 2001), but also followers are more likely to surface act to appear as though they enjoy the humor even when they do not. To avoid this problematic cycle of humor expression and surface acting, leaders would be wise to adopt the adage of “less is more” and not overuse humor.

Additionally, our research demonstrates that directing humor at followers who hold high power distance values is risky and more likely to backfire, while expressing humor among followers with low power distance has fewer costs and more upside. With this insight in mind, managers should learn about their followers' power distance values before expressing humor. To do so, leaders can incorporate power distance measures into the existing training and evaluation programs or survey their followers (Kirkman et al., 2009).

Strengths, Limitations, and Future Research Directions

Our empirical strategy aimed to address each study's main limitations by the strengths of the other studies. Study 1 addressed both external and internal validity considerations by testing our model within a field experiment. Study 2, which was limited in external validity, enhanced internal validity further by using a behavioral experiment that manipulated both LH quantity and power distance. Study 3 further improved external validity and provided an ecologically valid test of our theoretical model. Despite these complementary strengths of our studies, several limitations and questions can be addressed by future research.

One limitation across our studies is the modest sample size, in particular Study 1 ($N = 88$) and Study 3 ($N = 126$). Such lower sample sizes can potentially raise concerns about adequate power to detect the hypothesized effects and reduce measure reliability and model fit (Little et al., 2013). For example, the lack of direct effects of LH quantity on wellbeing outcomes in Studies 1 and 3 may be due in part to low sample size. Also, the reliability of our power distance measure in Studies 1 and 3 were below the typical threshold of 0.70 for Cronbach's alpha (Cortina, 1993). Although low, the original study used to validate this power distance measure achieved a reliability of .63 (Dorfman & Howell, 1988), and the range of reliability across 18 studies using this power distance measure was .55 to .94 (average reliability was .75). Thus, although the reliability of our power distance measure was low in Study 1 and Study 3, they do

fall within the range of reliabilities found in past work. Furthermore, we found support for LH quantity \times power distance interactions across three studies, including Study 2 which manipulated power distance and used a different measure of power distance (Kirkman et al., 2009) as a manipulation check. That said, future research could test our hypotheses with larger samples and different measures of power distance to provide further evidence of our findings.

Study 3 was also limited in two other ways besides its correlational design. First, due to the anonymity of the data provided and the unique organizational structure, we could not rule out any possible nesting effects. Second, Study 3 relied on employees to report all variables in our model, which can create concerns of common method bias. Although we used time separation and ruled out some of these concerns in the supplemental analyses, future research could test our hypotheses using data from different sources (e.g., leaders and employees).

In addition, future research should expand our findings in several ways. Most critically, future research is needed to help explain how LH quantity and LH quality combine to influence follower emotional outcomes as well as help reconcile when both LH quantity and LH quality are more likely to enhance, as opposed to worsen, follower wellbeing outcomes. Our studies found independent effects of LH quantity and LH quality and that these variables do not interact to predict outcomes. Further, follower power distance moderated LH quantity's effects on surface acting, but it did not moderate LH quality's effects on surface acting or positive affect.

Although these initial results leave several questions unanswered regarding how LH quantity and LH quality combine to impact follower wellbeing outcomes, we believe they also provide a springboard for three promising theoretical directions to answer them.

First, future studies can investigate other boundary conditions, besides power distance, that may moderate the effects of LH quantity and LH quality. One promising direction is to

examine variables that may influence followers' beliefs about the appropriateness of humor use by their leaders in the workplace. Research finds that viewing humor expression as appropriate or inappropriate can significantly alter recipients' reactions, and that judgements of humor appropriateness can be independent from quality judgements (e.g., a person can find a joke amusing and inappropriate) (Bitterly, 2022). In relation to our studies, the more a follower finds a leader's use of humor at work to be inappropriate (appropriate), the more likely that LH quantity and LH quality will result in worse (better) emotional wellbeing outcomes. Key variables that likely determine followers' beliefs regarding the appropriateness of leader humor expression include followers' personal sense of humor (e.g., does the follower value humor and believe it to be a desirable trait?), norms about humor use at work (e.g., is humor common and valued among group members?; O'Neill & Rothbard, 2017), and the dominant type or style of humor that the leader uses. As an example of type of humor, research has identified selfenhancing and self-defeating humor (Martin, Puhlik-Doris, Larsen, Gray, & Weir, 2003). Selfenhancing humor (humor used to enhance one's image) could reinforce power boundaries among leaders and followers. In contrast, self-defeating humor (mocking oneself to seek acceptance) may blur such boundaries. Hence, the former (latter) may be seen as more appropriate and cause less (more) surface acting for followers with high power distance (with the opposite predictions for followers with low power distance), which future research can examine.

Second, future studies can identify other mechanisms, besides surface acting and positive affect, that may help explain LH quantity's and LH quality's combined effects on wellbeing outcomes. One key mechanism may be followers' appreciation of leader humor expression. In our pilot testing of Study 2, we learned that although some participants thought the leader's jokes were "bad" or "corny," they still appreciated the leader's attempt at humor because it helped

create a positive and comfortable environment. Similarly, research on failed humor attempts indicates that recipients of humor expression often show their appreciation for the expressor of humor even if they judged the humor as poor (e.g., they still laugh at the attempt; Bell, 2009; Hay, 2001). Specifying appreciation for leader humor attempts as a novel mechanism may help reconcile the overall effects of LH quantity and LH quality on wellbeing outcomes.

Third, future work should examine the interplay of LH quantity and LH quality on follower emotional outcomes at a within-person level, such as a daily or event level. For example, scholars could track leaders and followers throughout a day, capturing multiple instances of leader humor expression (including the type of humor and its quality) and subsequent reactions from followers. Aggregating such specific events and instances would help provide insight into important research questions such as how often do leaders express high- versus low-quality humor and what is the ratio of high- versus low-quality humor expression needed to achieve beneficial emotional outcomes without incurring the unintended costs?

Besides examining the interplay of LH quantity and LH quality, future studies can investigate other important directions related to our theorizing. First, studies can explore how specific types of leader humor expression may elicit more or less surface acting from followers. An example that comes from Study 2 is whether humor expression is accompanied by humor meta-communication. Humor meta-communication involves verbal messages (e.g., “I’ll tell you a joke” and “Wasn’t that funny?”) and non-verbal communication (e.g., laughing or smiling) that accompanies humor expression and signals or emphasizes to the recipient that the expresser intends to be humorous (Canestrari, 2010; Pickering et al., 2009). It is plausible that when humor meta-communication is present, followers may feel a stronger pressure to engage in surface acting to meet the expected display standard, which future research can examine.

Second, future research can explore display rules as the micro-mediator that we theorized between LH quantity and follower surface acting. Third, future studies can examine whether our predictions about power distance's moderating effects also apply to leader power. For example, perceptions of power and power distance are related (Adamovic, 2023), and our Study 2 manipulation of power distance may have also manipulated followers' perceptions of the leader's power. Future research can disentangle such possible effects of leader power and follower power distance. Fourth, another route for future research lies in exploring factors that reduce the harmful effects of LH quantity for followers with high power distance. One potential factor is leaders explicitly acknowledging the possible display rules caused by leader humor expression, thus relieving followers from the pressure to surface act.

CONCLUSION

Although high-quality leader humor can be useful to enhance employees' wellbeing, our theory and studies suggest that frequent leader humor expression, regardless of its quality, can cause employees to fake it with the bosses' jokes, reducing their wellbeing in the process. We also identify follower power distance values as a key moderating factor, such that leader humor expression is more (less) risky with followers who have high (low) power distance values. We hope our research motivates scholars to investigate how humor can be used at work to produce more beneficial and less harmful outcomes for employees.

REFERENCES

- Aiken, L. S., & West, S. G. 1991. *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA: Sage.
- Alves, J. C., Lovelace, K. J., Manz, C. C., Matsypura, D., Toyasaki, F., et al. 2006. A crosscultural perspective of self-leadership. *Journal of Managerial Psychology*, 21(4): 338– 359.
- Anderson, C., Hildreth, J. A. D., & Sharps, D. L. 2020. The Possession of High Status Strengthens the Status Motive. *Personality and Social Psychology Bulletin*, 46(12): 1712–1723.

- Barsade, S. G., & Gibson, D. E. 2007. Why Does Affect Matter in Organizations? *Academy of Management Perspectives*, 21(1): 36–59.
- Basch, J., & Fisher, C. D. 2000. Affective events–emotions matrix: A classification of work events and associated emotions. *Emotions in the workplace: Research, theory, and practice*: 36–48. Westport, CT, US: Quorum Books/Greenwood Publishing Group.
- Bennett, M. P., & Lengacher, C. A. 2006. Humor and laughter may influence health. I. History and background. *Evidence-Based Complementary and Alternative Medicine*, 3(1): 61–63.
- Bitterly, T. B. 2022. Humor and power. *Current Opinion in Psychology*, 43: 125–128.
- Bitterly, T. B., Brooks, A. W., & Schweitzer, M. E. 2017. Risky business: When humor increases and decreases status. *Journal of Personality and Social Psychology*, 112(3): 431.
- Bitterly, T. B., & Schweitzer, M. E. 2019. The impression management benefits of humorous self-disclosures: How humor influences perceptions of veracity. *Organizational Behavior and Human Decision Processes*, 151: 73–89.
- Bochner, S., & Hesketh, B. 1994. Power distance, individualism/collectivism, and job-related attitudes in a culturally diverse work group. *Journal of Cross-Cultural Psychology*, 25(2): 233–257.
- Bond, M. H., Wan, K.-C., Leung, K., & Giacalone, R. A. 1985. How are Responses to Verbal Insult Related to Cultural Collectivism and Power Distance? *Journal of Cross-Cultural Psychology*, 16(1): 111–127.
- Brislin, R. W. 1970. Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, 1(3): 185–216.
- Brotheridge, C. M., & Lee, R. T. 2003. Development and validation of the Emotional Labour Scale. *Journal of Occupational and Organizational Psychology*, 76(3): 365–379.
- Carlson, D., Ferguson, M., Hunter, E., & Whitten, D. 2012. Abusive supervision and work–family conflict: The path through emotional labor and burnout. *The Leadership Quarterly*, 23(5): 849–859.
- Carnevale, J. B., Huang, L., Yam, K. C., & Wang, L. 2022. Laughing with me or laughing at me? The differential effects of leader humor expressions on follower status and influence at work. *Journal of Organizational Behavior*, 43(7): 1153–1171.
- Chapman, A. J. 1983. Humor and laughter in social interaction and some implications for humor research. *Handbook of humor research*: 135–157. Springer.
- Cole, M. S., Carter, M. Z., & Zhang, Z. 2013. Leader–team congruence in power distance values and team effectiveness: The mediating role of procedural justice climate. *Journal of Applied Psychology*, 98(6): 962.
- Conger, J. A., & Kanungo, R. N. 1994. Charismatic leadership in organizations: Perceived behavioral attributes and their measurement. *Journal of Organizational Behavior*, 15(5): 439–452.
- Cooper, C. D. 2005. Just joking around? Employee humor expression as an ingratiation behavior. *Academy of Management Review*, 30(4): 765–776.
- Cooper, C. D., Kong, D. T., & Crossley, C. D. 2018. Leader Humor as an Interpersonal Resource: Integrating Three Theoretical Perspectives. *Academy of Management Journal*, 61(2): 769–796.

- Cortina, J. M. 1993. What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78: 98–104.
- Coser, R. L. 1960. Laughter among Colleagues†. *Psychiatry*, 23(1): 81–95.
- Côté, S. 2005. A Social Interaction Model Of The Effects Of Emotion Regulation On Work Strain. *Academy of Management Review*, 30(3): 509–530.
- Daniels, M. A., & Greguras, G. J. 2014. Exploring the Nature of Power Distance: Implications for Micro- and Macro-Level Theories, Processes, and Outcomes. *Journal of Management*, 40(5): 1202–1229.
- Decker, W. H. (1987). Managerial humor and subordinate satisfaction. *Social Behavior and Personality: an international journal*, 15(2), 225-232.
- Decker, W. H., & Rotondo, D. M. 2001. Relationships among gender, type of humor, and perceived leader effectiveness. *Journal of Managerial Issues*, 450–465.
- Diefendorff, J. M., & Gosserand, R. H. 2003. Understanding the emotional labor process: A control theory perspective. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 24(8): 945–959.
- Diefendorff, J., Morehart, J., & Gabriel, A. 2010. The influence of power and solidarity on emotional display rules at work. *Motivation and Emotion*, 34(2): 120–132.
- Dorfman, P. W., & Howell, J. P. 1988. Dimensions of national culture and effective leadership patterns: Hofstede revisited. *Advances in International Comparative Management*, 3(1): 127–150.
- Dulebohn, J. H., Bommer, W. H., Liden, R. C., Brouer, R. L., & Ferris, G. R. 2012. A MetaAnalysis of Antecedents and Consequences of Leader-Member Exchange: Integrating the Past With an Eye Toward the Future. *Journal of Management*, 38(6): 1715–1759.
- Edwards, J. R., & Lambert, L. S. 2007. Methods for integrating moderation and mediation: A general analytical framework using moderated path analysis. *Psychological Methods*, 12(1): 1.
- Elfenbein, H. A. 2007. Emotion in organizations: A review and theoretical integration. *Academy of Management Annals*, 1(1): 315–386.
- Erez, A., Misangyi, V. F., Johnson, D. E., LePine, M. A., & Halverson, K. C. 2008. Stirring the hearts of followers: Charismatic leadership as the transferal of affect. *Journal of Applied Psychology*, 93(3): 602.
- Evans, J. B., Slaughter, J. E., Ellis, A. P., & Rivin, J. M. (2019). Gender and the evaluation of humor at work. *Journal of Applied Psychology*, 104(8), 1077.
- Farmer, S. M., & Aguinis, H. 2005. Accounting for Subordinate Perceptions of Supervisor Power: An Identity-Dependence Model. *Journal of Applied Psychology*, 90(6): 1069–1083.
- Feldman, D. C. 1984. The development and enforcement of group norms. *Academy of Management Review*, 9(1): 47–53.
- Fisk, G. M., & Friesen, J. P. 2012. Perceptions of leader emotion regulation and LMX as predictors of followers' job satisfaction and organizational citizenship behaviors. *The Leadership Quarterly*, 23(1): 1–12.
- Glasø, L., & Einarsen, S. 2008. Emotion regulation in leader–follower relationships. *European Journal of Work and Organizational Psychology*, 17(4): 482–500. Glew, D. J., Mesmer-

- Magnus, J., & Viswesvaran, C. 2012. A meta-analysis of positive humor in the workplace. *Journal of Managerial Psychology*, 27(2): 155–190.
- Gloor, J. L., Cooper, C. D., Bowes-Sperry, L., & Chawla, N. 2021. Risqué business? Interpersonal anxiety and humor in the# MeToo era. *Journal of Applied Psychology*.
- Goncalo, J. A., & Duguid, M. M. 2012. Follow the crowd in a new direction: When conformity pressure facilitates group creativity (and when it does not). *Organizational Behavior and Human Decision Processes*, 118(1): 14–23.
- Goncalo, J. A., & Staw, B. M. 2006. Individualism–collectivism and group creativity. *Organizational Behavior and Human Decision Processes*, 100(1): 96–109.
- Goswami, A., Nair, P., Beehr, T., & Grossenbacher, M. 2016. The relationship of leaders' humor and employees' work engagement mediated by positive emotions: Moderating effect of leaders' transformational leadership style. *Leadership & Organization Development Journal*, 37(8): 1083–1099.
- Goswami, A., Nair, P. K., & Grossenbacher, M. A. 2015. Impact of aggressive humor on dysfunctional resistance. *Personality and Individual Differences*, 74: 265–269.
- Grandey, A. A. 2000. Emotional regulation in the workplace: A new way to conceptualize emotional labor. *Journal of Occupational Health Psychology*, 5(1): 95.
- Grandey, A. A. 2003. When “the show must go on”: Surface acting and deep acting as determinants of emotional exhaustion and peer-rated service delivery. *Academy of Management Journal*, 46(1): 86–96.
- Grandey, A. A., & Gabriel, A. S. 2015. Emotional Labor at a Crossroads: Where Do We Go from Here? *Annual Review of Organizational Psychology and Organizational Behavior*, 2(1): 323–349.
- Grandey, A. A., & Melloy, R. C. 2017. The state of the heart: Emotional labor as emotion regulation reviewed and revised. *Journal of Occupational Health Psychology*, 22(3): 407.
- Guenzi, P., Rangarajan, D., Chaker, N. N., & Sajtos, L. 2019. It is all in good humor? Examining the impact of salesperson evaluations of leader humor on salesperson job satisfaction and job stress. *Journal of Personal Selling & Sales Management*, 39(4): 352–369.
- Hecht, M. A., & LaFrance, M. 1998. License or Obligation to Smile: The Effect of Power and Sex on Amount and Type of Smiling. *Personality and Social Psychology Bulletin*, 24(12): 1332–1342.
- Hofstede, G. 1980. Cultural consequences. *Beverly Hills: Stage*. Hu, X., & Shi, J. 2015. Employees' surface acting in interactions with leaders and peers: Employees' Surface Acting. *Journal of Organizational Behavior*, 36(8): 1132–1152.
- Hughes, L. W., & Avey, J. B. 2009. Transforming with levity: Humor, leadership, and follower attitudes. *Leadership & Organization Development Journal*, 30(6): 540–562.
- Hülsheger, U. R., & Schewe, A. F. 2011. On the costs and benefits of emotional labor: A metaanalysis of three decades of research. *Journal of Occupational Health Psychology*, 16(3): 361–389.
- Huo, Y., Lam, W., & Chen, Z. 2012. Am I the Only One This Supervisor is Laughing at? Effects of Aggressive Humor on Employee Strain and Addictive Behaviors: Personnel Psychology. *Personnel Psychology*, 65(4): 859–885.

- Judge, T. A., Thoresen, C. J., Bono, J. E., & Patton, G. K. 2001. The job satisfaction–job performance relationship: A qualitative and quantitative review. *Psychological Bulletin*, 127(3): 376.
- Kammeyer-Mueller, J. D., Rubenstein, A. L., Long, D. M., Odio, M. A., Buckman, B. R., et al. 2013. A meta-analytic structural model of dispositional affectivity and emotional labor. *Personnel Psychology*, 66(1): 47–90.
- Karakowsky, L., Podolsky, M., & Elangovan, A. R. 2020. Signaling trustworthiness: The effect of leader humor on feedback-seeking behavior. *The Journal of Social Psychology*, 160(2): 170–189.
- Kirkman, B. L., Chen, G., Farh, J.-L., Chen, Z. X., & Lowe, K. B. 2009. Individual power distance orientation and follower reactions to transformational leaders: A cross-level, cross-cultural examination. *Academy of Management Journal*, 52(4): 744–764.
- Kong, D. T., Cooper, C. D., & Sosik, J. J. 2019. The state of research on leader humor. *Organizational Psychology Review*, 9(1): 3–40.
- Lehman, D. W., O'Connor, K., Kovács, B., & Newman, G. E. 2019. Authenticity. *Academy of Management Annals*, 13(1): 1–42.
- Levine, E. E. (2022). Community standards of deception: Deception is perceived to be ethical when it prevents unnecessary harm. *Journal of Experimental Psychology: General*, 151(2), 410.
- Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. 2002. To Parcel or Not to Parcel: Exploring the Question, Weighing the Merits. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2): 151–173.
- Little, T. D., Rhemtulla, M., Gibson, K., & Schoemann, A. M. 2013. Why the items versus parcels controversy needn't be one. *Psychological Methods*, 18: 285–300.
- Locke, E. A. 1976. The Nature and Consequences of Job Satisfaction. In M. D. Dunnette. *Handbook of Industrial and Organizational Psychology*, Chicago: Rand-McNally, 1297–1349.
- Magee, J. C., & Galinsky, A. D. 2008. Social Hierarchy: The Self-Reinforcing Nature of Power and Status. *Academy of Management Annals*, 2(1): 351–398.
- Martin, R. A., Puhlik-Doris, P., Larsen, G., Gray, J., & Weir, K. 2003. Individual differences in uses of humor and their relation to psychological well-being: Development of the Humor Styles Questionnaire. *Journal of Research in Personality*, 37(1): 48–75.
- McNeish, D., Stapleton, L. M., & Silverman, R. D. 2017. On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods*, 22(1): 114.
- Mesmer-Magnus, J., Glew, D. J., & Viswesvaran, C. 2012. A meta-analysis of positive humor in the workplace. *Journal of Managerial Psychology*, 27(2): 155–190.
- Morris, J. A., & Feldman, D. C. 1996. The dimensions, antecedents, and consequences of emotional labor. *Academy of Management Review*, 21(4): 986–1010.
- Morrison, E. W., & Milliken, F. J. 2000. Organizational Silence: A Barrier to Change and Development in a Pluralistic World. *Academy of Management Review*, 25(4): 706–725.
- Muthén, L. K., & Muthén, B. O. 2017. Mplus Version 8 user's guide. *Los Angeles, CA: Muthén & Muthén*.

- Ng, K.-Y., Koh, C., Ang, S., Kennedy, J. C., & Chan, K.-Y. 2011. Rating leniency and halo in multisource feedback ratings: Testing cultural assumptions of power distance and individualism-collectivism. *Journal of Applied Psychology*, 96(5): 1033.
- Ozcelik, H. 2013. An empirical analysis of surface acting in intra-organizational relationships. *Journal of Organizational Behavior*, 34(3): 291–309.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. 2003. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5): 879.
- Robert, C., Dunne, T. C., & Iun, J. 2016. The impact of leader humor on subordinate job satisfaction: The crucial role of leader–subordinate relationship quality. *Group & Organization Management*, 41(3): 375–406.
- Robert, C., & Wilbanks, J. E. 2012. The wheel model of humor: Humor events and affect in organizations. *Human Relations*, 65(9): 1071–1099.
- Robinson, D. T., & Smith-Lovin, L. 2001. Getting a laugh: Gender, status, and humor in task discussions. *Social Forces*, 80(1): 123–158.
- Romero, E. J., & Cruthirds, K. W. 2006. The use of humor in the workplace. *The Academy of Management Perspectives*, 20(2): 58–69.
- Rubenstein, A. L., Eberly, M. B., Lee, T. W., & Mitchell, T. R. 2018. Surveying the forest: A meta-analysis, moderator investigation, and future-oriented discussion of the antecedents of voluntary employee turnover. *Personnel Psychology*, 71(1): 23–65.
- Rupp, D. E., & Spencer, S. 2006. When customers lash out: The effects of customer interactional injustice on emotional labor and the mediating role of discrete emotions. *Journal of Applied Psychology*, 91(4): 971.
- Scandura, T. A., Graen, G. B., & Novak, M. A. 1986. When managers decide not to decide autocratically: An investigation of leader–member exchange and decision influence. *Journal of Applied Psychology*, 71(4): 579.
- Schaubroeck, J., Lam, S. S. K., & Cha, S. E. 2007. Embracing transformational leadership: Team values and the impact of leader behavior on team performance. *Journal of Applied Psychology*, 92(4): 1020–1030.
- Schermuly, C. C., & Meyer, B. 2016. Good relationships at work: The effects of Leader–Member Exchange and Team–Member Exchange on psychological empowerment, emotional exhaustion, and depression. *Journal of Organizational Behavior*, 37(5): 673–691.
- Seo, M.-G., Barrett, L. F., & Jin, S. 2008. *The structure of affect: History, theory, and implications for emotion research in organizations*.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. 2002. *Experimental and quasi-experimental designs for generalized causal inference*: xxi, 623. Boston, MA, US: Houghton, Mifflin and Company.
- Shumski Thomas, J., Olien, J. L., Allen, J. A., Rogelberg, S. G., & Kello, J. E. 2018. Faking it for the higher-ups: Status and surface acting in workplace meetings. *Group & Organization Management*, 43(1): 72–100.
- Smeltzer, L. R., & Leap, T. L. 1988. An analysis of individual reactions to potentially offensive jokes in work settings. *Human Relations*, 41(4): 295–304.

- Smith, P. B., Peterson, M. F., & Schwartz, S. H. 2002. Cultural Values, Sources of Guidance, and their Relevance to Managerial Behavior: A 47-Nation Study. *Journal of CrossCultural Psychology*, 33(2): 188–208.
- Sobral, F., & Islam, G. 2015. He Who Laughs Best, Leaves Last: The Influence of Humor on the Attitudes and Behavior of Interns. *Academy of Management Learning & Education*, 14(4): 500–518.
- Susa, A. M. 2002. *Humor type, organizational climate, and outcomes: The shortest distance between an organization's environment and the bottom line is laughter*. The University of Nebraska-Lincoln.
- Swider, B. W., & Zimmerman, R. D. 2010. Born to burnout: A meta-analytic path model of personality, job burnout, and work outcomes. *Journal of Vocational Behavior*, 76(3): 487–506.
- Taras, V., Steel, P., & Kirkman, B. L. 2016. Does country equate with culture? Beyond geography in the search for cultural boundaries. *Management International Review*, 56(4): 455–487.
- Taris, T. W., Horn, J. E. V., Schaufeli, W. B., & Schreurs, P. J. 2004. Inequity, burnout and psychological withdrawal among teachers: A dynamic exchange model. *Anxiety, Stress & Coping*, 17(1): 103–122.
- Tepper, B. J. (2000). Consequences of abusive supervision. *Academy of Management Journal*, 43, 178–190.
- Van Knippenberg, D. & Sitkin, S. B. 2013. A Critical Assessment of Charismatic—Transformational Leadership Research: Back to the Drawing Board? *The Academy of Management Annals*, 7(1): 1–60.
- Vecchio, R. P., Justin, J. E., & Pearce, C. L. 2009. The influence of leader humor on relationships between leader behavior and follower outcomes. *Journal of Managerial Issues*, 171–194.
- Wang, W., Mao, J., Wu, W., & Liu, J. 2012. Abusive supervision and workplace deviance: The mediating role of interactional justice and the moderating role of power distance. *Asia Pacific Journal of Human Resources*, 50(1): 43–60.
- Warren, C., Barsky, A., & McGraw, A. P. 2021. What makes things funny? An integrative review of the antecedents of laughter and amusement. *Personality and Social Psychology Review*, 25(1): 41–65.
- Wharton, A. S. 1993. The affective consequences of service work: Managing emotions on the job. *Work and Occupations*, 20(2): 205–232.
- Wijewardena, N., Härtel, C. E., & Samaratunge, R. 2017. Using humor and boosting emotions: An affect-based study of managerial humor, employees' emotions and psychological capital. *Human Relations*, 70(11): 1316–1341.
- Yam, K. C., Christian, M. S., Wei, W., Liao, Z., & Nai, J. 2018. The Mixed Blessing of Leader Sense of Humor: Examining Costs and Benefits. *Academy of Management Journal*, 61(1): 348–369.
- Zhu, X., Tian, G., & Liu, Z. 2022. The Smaller the Power Distance, the More Genuine the Emotion: Relationships between Power Distance, Emotional Labor, and Emotional Exhaustion among Chinese Teachers. *Sustainability*, 14(14): 8601

TABLE 1
Means, standard deviations, and correlations among variables (Study 1)

	Variables	Mean	SD	1	2	3	4	5	6
1	Leader humor quantity condition	0.52	0.50						
2	Follower power distance	2.96	0.79	-0.08					
3	Follower surface acting	1.95	0.88	0.25*	0.23*				
4	Follower emotional exhaustion	2.65	0.84	0.11	0.22*	0.48**			
5	Follower job satisfaction	3.29	0.74	-0.01	-0.17	-0.28**	-0.42**		
6	Follower positive affect	2.79	0.68	0.06	0.10	-0.002	-0.30**	0.48**	
7	Follower deep acting	2.52	0.88	0.14	0.01	0.27*	0.09	-0.09	0.13

Note: $N=88$. * $p < .05$, ** $p < .01$. LH quantity condition was coded 0 for control and 1 for LH quantity.

TABLE 2
Means, standard deviations, and ANOVAs of outcomes by condition (Study 1)

SURFACE ACTING		EMOTIONAL EXHAUSTION		JOB SATISFACTION	
Control	LH Quantity	Control	LH Quantity	Control	LH Quantity
M=1.72 (SD=0.83)	M=2.15 (SD=0.88)	M=2.55 (SD=0.88)	M=2.74 (SD=0.80)	M=3.30 (SD=0.72)	M=3.28 (SD=0.77)
$F(1, 86) = 5.50, p = .021, \eta_p^2 = .06$		$F(1, 86) = 1.10, p = .298, \eta_p^2 = .01$		$F(1, 86) = 0.01, p = .905, \eta_p^2 = .00$	

Note: $N=88$.

TABLE 3
Regression results (Study 1)

Variables	Surface acting				Emotional exhaustion				Job satisfaction			
	Model 1a	Model 1b	Model 2a	Model 2b	Model 3a	Model 3b	Model 4a	Model 4b	Model 5a	Model 5b	Model 6a	Model 6b
	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)
Intercept	1.94*** (0.18)	1.95*** (0.18)	1.96*** (0.18)	1.96*** (0.18)	3.00*** (0.29)	3.03*** (0.29)	3.12*** (0.27)	3.15*** (0.27)	4.52*** (0.32)	4.49*** (0.33)	4.45*** (0.33)	4.41*** (0.34)
LH quantity condition	0.27** (0.10)	0.27** (0.10)	0.27** (0.09)	0.27** (0.10)	0.15 (0.10)	0.13 (0.10)	0.04 (0.09)	0.01 (0.10)	-0.06 (0.09)	-0.03 (0.11)	0.01 (0.09)	0.05 (0.11)
Follower power distance	0.26** (0.10)	0.25** (0.10)	0.05 (0.14)	0.04 (0.14)	0.26** (0.10)	0.23* (0.10)	0.15 (0.09)	0.11 (0.09)	-0.23* (0.09)	-0.18 (0.10)	-0.17 (0.09)	-0.11 (0.10)
LH quantity \times Power distance			0.31* (0.13)	0.31* (0.13)								
Follower surface acting							0.43*** (0.08)	0.45*** (0.08)			-0.24* (0.10)	-0.27* (0.11)
<u>Control Variable</u>												
Follower positive affect	-0.04 (0.09)		-0.04 (0.09)		-0.33** (0.10)		-0.31*** (0.09)		0.51*** (0.09)		0.49*** (0.09)	
R ²	0.126	0.124	0.179	0.178	0.171	0.063	0.333	0.237	0.281	0.031	0.333	0.094

Note: N=88. Standardized estimates provided. Standard errors reported in parentheses. LH quantity condition was coded 0 for control and 1 for LH quantity. All predictor variables (except for the LH quantity condition) centered at sample mean values. * $p < .05$, ** $p < .01$, *** $p < .001$.

TABLE 4
Indirect effects analyses for Studies 1-3

Indirect Effects	STUDY 1		STUDY 2		STUDY 3 (Employee Survey DVs)		STUDY 3 (Company Records DVs)	
	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
Leader humor quantity → surface acting → emotional exhaustion	0.19	[.06, .40]	0.21	[.10, .38]	0.05	[.01, .14]	0.10	[.01, .27]
At High Follower PD	0.36	[.14, .68]	0.36	[.20, .58]	0.08	[.01, .18]	0.14	[.01, .37]
At Low Follower PD	0.03	[-.18, .23]	0.07	[-.01, .19]	0.02	[-.02, .09]	0.03	[-.04, .19]
Difference	0.34	[.04, .73]	0.30	[.15, .52]	0.06	[.004, .17]	0.11	[.01, .35]
Leader humor quantity → surface acting → job satisfaction	-0.10	[-.25, -.02]	-0.12	[-.28, -.02]	-0.01	[-.06, .03]	0.02	[-.06, .13]
At High Follower PD	-0.18	[-.41, -.04]	-0.20	[-.43, -.04]	-0.01	[-.08, .04]	0.03	[-.10, .17]
At Low Follower PD	-0.01	[-.15, .08]	-0.04	[-.13, .002]	-0.003	[-.04, .01]	0.01	[-.02, .10]
Difference	-0.17	[-.44, -.02]	-0.16	[-.40, -.03]	-0.01	[-.08, .03]	0.03	[-.07, .16]

Note: Study 1 $N = 88$; Study 2 $N = 198$; Study 3 $N = 126$. Bias-corrected bootstrapping = 10,000 draws. DVs = dependent variables. PD = power distance. Indirect effects for emotional exhaustion and job satisfaction were estimated simultaneously. In calculating all moderated indirect effects, the moderating effect of power distance was specified on the first-stage relationship between leader humor quantity and surface acting. Control variables outlined for each study were included when calculating these reported indirect effects.

TABLE 5
Means, standard deviations, and correlations among variables (Study 2)

Variables	Mean	SD	1	2	3	4	5	6	7
-----------	------	----	---	---	---	---	---	---	---

1	Leader humor quantity condition	0.51	0.50							
2	Power distance condition	0.48	0.50	-0.01						
3	Follower surface acting	2.17	1.33	0.18*	0.25**					
4	Follower emotional exhaustion	1.88	1.07	-0.01	0.21**	0.43**				
5	Follower work satisfaction	5.50	1.31	0.07	-0.21**	-0.26**	-0.63**			
6	Follower positive affect	2.94	0.97	0.16*	-0.12	-0.14	-0.45**	0.62**		
7	Leader humor quality	4.02	1.69	0.46**	-0.23**	-0.17*	-0.37**	0.33**	0.49**	
8	Follower deep acting	4.25	1.38	0.08	-0.13	-0.02	-0.18*	0.30**	0.39**	0.21**

Note: $N=198$. Leader humor quantity and power distance conditions were coded 0 for low and 1 for high.

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

TABLE 6
Means and standard deviations of outcomes by condition (Study 2)

	SURFACE ACTING		EMOTIONAL EXHAUSTION		WORK SATISFACTION	
	Low Power Distance	High Power Distance	Low Power Distance	High Power Distance	Low Power Distance	High Power Distance
Low LH Quantity	1.89 (1.03)	1.99 (1.31)	1.69 (0.72)	2.09 (1.40)	5.63 (1.02)	5.19 (1.75)
High LH Quantity	1.81 (1.08)	3.05 (1.52)	1.62 (0.75)	2.13 (1.19)	5.91 (0.85)	5.24 (1.37)

Note: $N=198$. Standard deviations reported in parentheses.

TABLE 7
Regression results (Study 2)

Variables	Surface acting				Emotional exhaustion				Work satisfaction			
	Model 1a	Model 1b	Model 2a	Model 2b	Model 3a	Model 3b	Model 4a	Model 4b	Model 5a	Model 5b	Model 6a	Model 6b
	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)
Intercept	1.15*** (0.12)	1.21*** (0.13)	1.35*** (0.10)	1.42*** (0.11)	1.49*** (0.15)	1.57*** (0.17)	1.65*** (0.16)	1.75*** (0.16)	4.39*** (0.29)	4.36*** (0.31)	4.32*** (0.30)	4.26*** (0.31)
LH quantity	0.30*** (0.07)	0.18** (0.07)	0.09 (0.06)	-0.03 (0.06)	0.16* (0.07)	-0.01 (0.07)	0.06 (0.07)	-0.08 (0.07)	-0.04 (0.06)	0.07 (0.07)	0.004 (0.06)	0.11 (0.07)
Follower power distance	0.20** (0.06)	0.25*** (0.06)	-0.01 (0.07)	0.04 (0.09)	0.12 (0.06)	0.21** (0.07)	0.05 (0.06)	0.11 (0.07)	-0.13** (0.05)	-0.21** (0.07)	-0.11* (0.05)	-0.15* (0.07)
LH quantity \times Power distance			0.36*** (0.10)	0.37** (0.11)								
Follower surface acting							0.34*** (0.08)	0.42*** (0.08)			-0.15* (0.07)	-0.24** (0.08)
<u>Control Variables</u>												
LH quality	-0.24** (0.08)		-0.24** (0.07)		-0.25** (0.09)		-0.17* (0.08)		0.03 (0.07)		-0.01 (0.07)	
Follower positive affect	-0.04 (0.08)		-0.05 (0.08)		-0.34*** (0.07)		-0.32*** (0.07)		0.59*** (0.06)		0.59*** (0.06)	
R ²	0.149	0.096	0.192	0.141	0.264	0.046	0.360	0.205	0.402	0.049	0.421	0.099

Note: N=198. Cluster-robust standard errors used for regression analysis to account for group nesting. Standardized estimates provided. Standard errors reported in parentheses. LH quantity and power distance conditions were coded 0 for low and 1 for high. All predictor variables (except for the LH quantity and power distance conditions) centered at sample mean values.

* $p < .05$, ** $p < .01$, *** $p < .001$.

TABLE 8

Means, standard deviations, and correlations among variables (Study 3)

	Variables	Mean	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1.	LH quantity	2.41	0.82														
2.	Follower power distance	2.55	0.49	0.23*													
3.	Follower surface acting	1.92	0.87	0.34*	0.34*												
4.	Emotional exhaustion (employee survey)	2.01	0.76	0.24*	0.27*	0.35*											
5.	Job satisfaction (employee survey)	4.05	0.65	-0.11	-0.19*	-0.20*	-0.58*										
6.	Emotional exhaustion (company records)	3.24	1.58	0.12	0.20*	0.26*	0.58*	-0.51*									
7.	Job satisfaction (company records)	7.94	1.33	0.11	-0.13	-0.02	-0.44*	0.54*	-0.71*								
8.	Follower positive affect	3.49	0.78	-0.16	-0.28*	-0.25*	-0.62*	0.64*	-0.52*	0.55*							
9.	LMX	3.99	0.66	0.17	-0.19*	-0.15	-0.28*	0.43*	-0.30*	0.25*	0.34*						
10.	LH quality	3.62	0.71	0.35*	0.11	-0.09	-0.19*	0.24*	-0.17	0.23*	0.21*	.54*					
11.	Follower age	32.83	5.06	-0.13	-0.15	-0.10	-0.02	-0.06	-0.06	-0.08	0.02	-0.05	-0.09				
12.	Follower gender	0.43	0.50	-0.05	-0.01	-0.23*	-.002	0.01	0.02	-0.07	0.02	0.17	0.11	-.10			
13.	Leader gender	0.25	0.43	0.12	-0.13	-0.05	0.15	-0.05	0.22*	-0.21*	-0.002	0.18*	-.11	-.03	0.10		
14.	Follower deep acting	2.55	0.78	0.32*	0.27*	0.43*	0.23*	0.03	0.04	0.10	-0.04	0.19*	-0.03	-0.12	-0.06	0.18*	
15.	Charismatic leadership	4.12	0.71	0.08	-0.17	-0.15	-0.20*	0.40*	-0.28*	0.19*	0.31*	0.73*	0.48*	-0.004	0.15	0.15	0.01

Note: N=126. Gender was coded 0 for male and 1 for female. *

$p < 0.05$

TABLE 9
Regression results (Study 3)

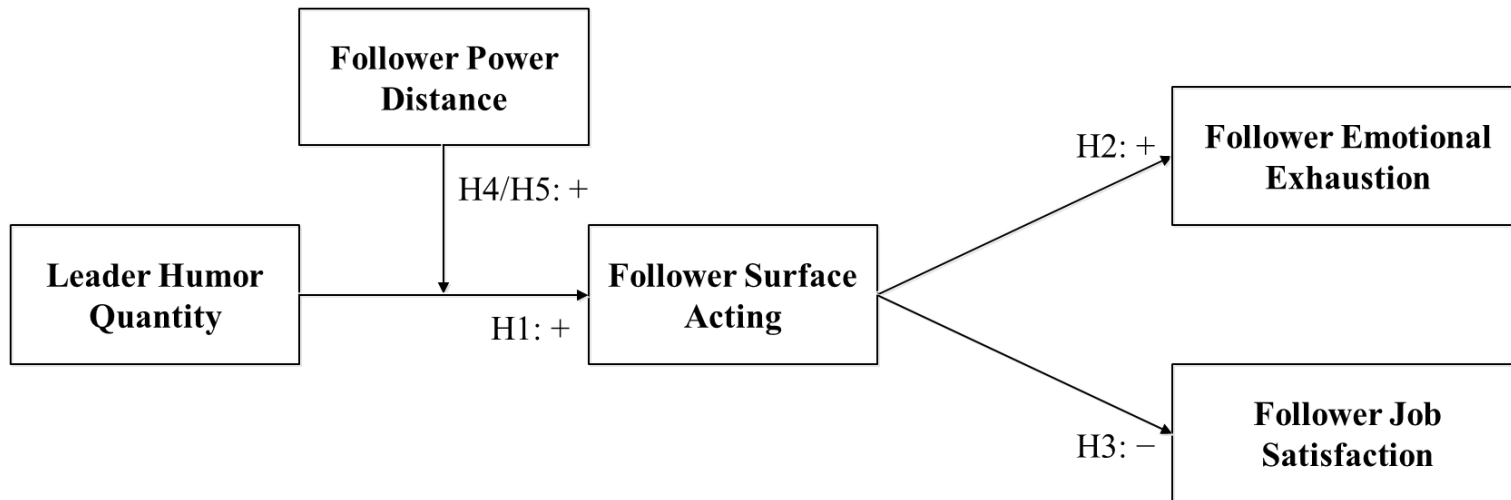
Variables	Surface acting				Emotional exhaustion (company records)				Job satisfaction (company records)			
	Model 1a	Model 1b	Model 2a	Model 2b	Model 3a	Model 3b	Model 4a	Model 4b	Model 5a	Model 5b	Model 6a	Model 6b
	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)
Intercept	2.42*** (0.15)	2.21*** (0.12)	2.37*** (0.15)	2.17*** (0.12)	1.88*** (0.15)	2.06*** (0.12)	1.85*** (0.15)	2.06*** (0.12)	6.21*** (0.36)	6.01*** (0.35)	6.21*** (0.36)	6.01*** (0.35)
LH quantity	0.32** (0.11)	0.27** (0.10)	0.29** (0.10)	0.24* (0.10)	0.02 (0.08)	0.08 (0.09)	-0.03 (0.08)	0.02 (0.09)	0.22** (0.08)	0.15 (0.08)	0.21** (0.08)	0.16 (0.09)
Follower power distance	0.26** (0.08)	0.28*** (0.07)	0.31*** (0.08)	0.32*** (0.07)	0.05 (0.07)	0.18* (0.08)	0.01 (0.08)	0.12 (0.09)	-0.05 (0.07)	-0.16 (0.08)	-0.06 (0.08)	-0.15 (0.09)
LH quantity \times Power distance			0.16* (0.07)	0.17** (0.07)								
Follower surface acting							0.16* (0.08)	0.21* (0.09)			0.04 (0.08)	-0.02 (0.10)
<u>Control Variables</u>												
LH quality	-0.22* (0.11)		-0.23* (0.11)		0.06 (0.09)		0.09 (0.09)		-0.03 (0.09)		-0.02 (0.09)	
LMX	0.04 (0.11)		0.07 (0.11)		-0.23** (0.09)		-0.23** (0.09)		0.09 (0.09)		0.09 (0.09)	
Follower positive affect	-0.08 (0.08)		-0.06 (0.08)		-0.44*** (0.08)		-0.43*** (0.08)		0.55*** (0.08)		0.55*** (0.08)	
Follower age	-0.06 (0.06)		-0.06 (0.06)		-0.04 (0.07)		-0.03 (0.07)		-0.08 (0.06)		-0.08 (0.06)	
Follower gender	-0.19* (0.08)		-0.20** (0.08)		0.04 (0.07)		0.07 (0.08)		-0.06 (0.07)		-0.05 (0.07)	
Leader gender	-0.07 (0.09)		-0.04 (0.09)		0.26*** (0.07)		0.27*** (0.07)		-0.26*** (0.07)		-0.26*** (0.07)	
R ²	0.286	0.188	0.309	0.216	0.359	0.045	0.377	0.081	0.410	0.037	0.411	0.037

Note: N=126. Standardized estimates provided. Standard errors reported in parentheses. All predictor variables except for gender centered at sample mean values.

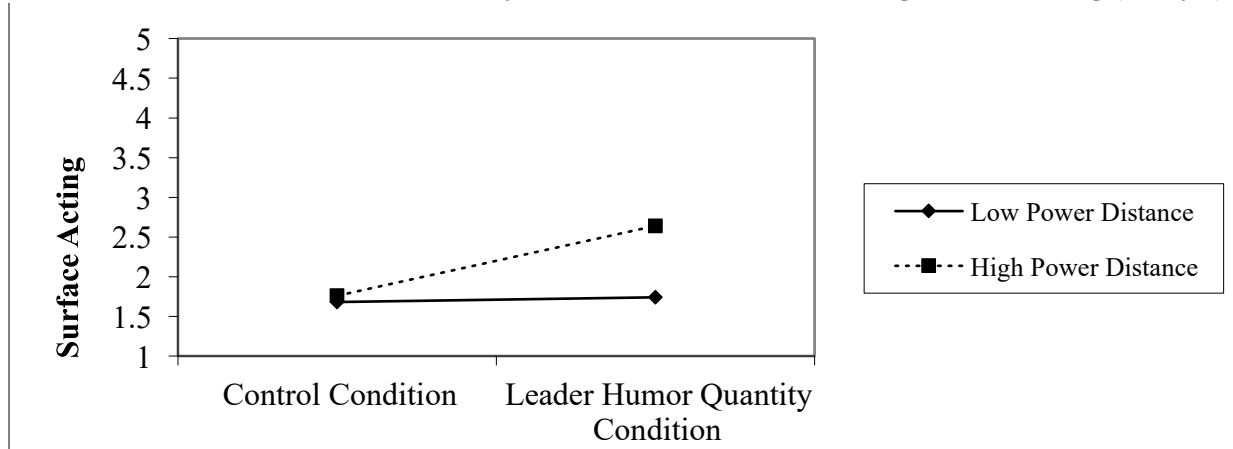
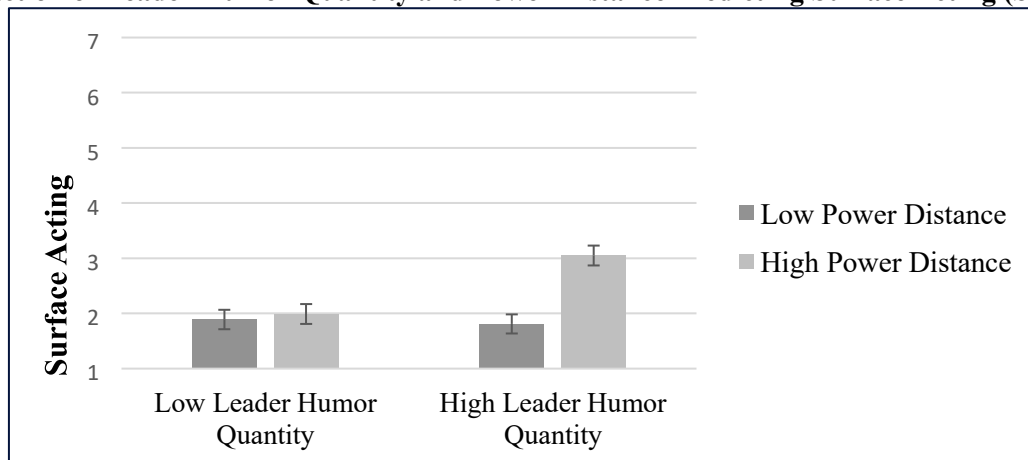
Gender was coded 0 for male and 1 for female.

* $p < .05$, ** $p < .01$, *** $p < .001$.

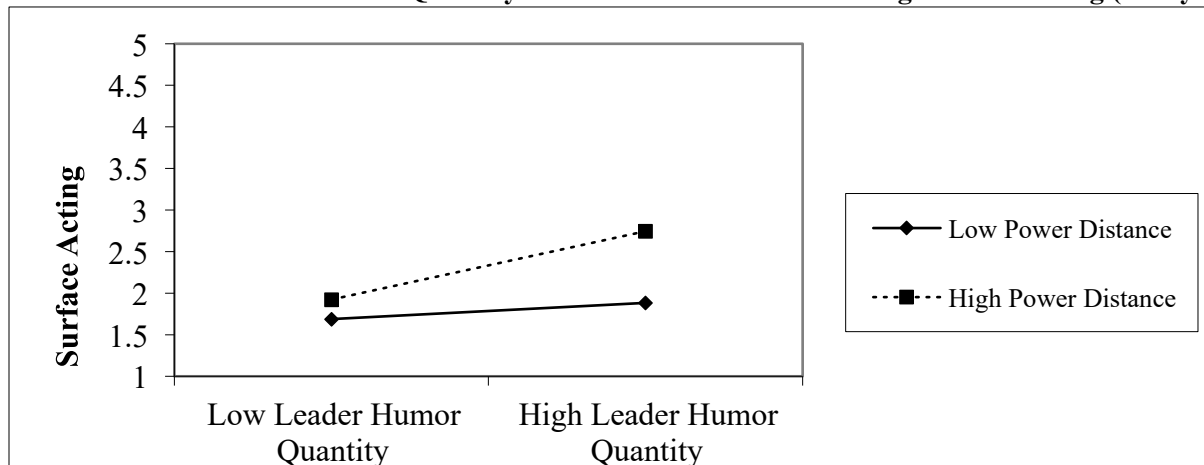
FIGURE 1 Theoretical model



Note: H2 and H3 denote indirect effects of leader humor quantity via surface acting. H4 denotes direct moderation effect on surface acting while H5 denotes moderated indirect effects on outcomes via surface acting.

FIGURE 2**Interaction of Leader Humor Quantity and Power Distance Predicting Surface Acting (Study 1)****FIGURE 3****Interaction of Leader Humor Quantity and Power Distance Predicting Surface Acting (Study 2)**

Note: lines on bar graphs represent standard errors.

FIGURE 4**Interaction of Leader Humor Quantity and Power Distance Predicting Surface Acting (Study 3)**

APPENDIX A Study 2 Experimental Manipulation of Leader Humor Quantity Embedded in the Focus Group Leader's Script

Low LH Quantity Condition	High LH Quantity Condition
<p>For instance, would you be willing to pay \$4 for this Balloon [point to balloon]? The price may seem high. But I'll remind you, the price of balloons is truly dependent on the economy.</p>	<p>For instance, would you be willing to pay \$4 for this Balloon [point to balloon]? The price may seem high. But I'll remind you, the price of balloons is truly dependent on inflation. [Haha] Inflation, get it? The price is dependent on inflation.</p>
<p>Moving on, in the bookstore, there is an ongoing debate about which item is the most popular, but it typically ends up being one of our men's accessories.</p>	<p>Moving on, in the bookstore, there is an ongoing debate about which item is the most popular, but it typically ends up being a tie. [Haha] It ends up being a tie.</p>
<p>The next item we'll discuss is a branded jacket. This new jacket will be reversible.</p>	<p>The next item we'll discuss is a branded jacket. This new jacket will be reversible... as it turns out. [Haha]. As it turns out, the reversible jacket.</p>
<p>I've always thought it is a shame that we don't have reversible Blazers; but our president is worried they would sell out very quickly.</p>	<p>I've always thought it is a shame that we don't have reversible Blazers; but our president is worried they would sell so fast they'd catch on fire. Get it? They'd be blazing off the shelves. [Haha]</p>

Biographical Sketches

Xiaoran Hu (X.Hu16@lse.ac.uk) is an assistant professor of management at the London School of Economics and Political Science. He received his PhD from London Business School. His research interests include teams and leadership effectiveness, cultural values, emotions, creativity and innovation.

Michael R. Parke (mrparke@wharton.upenn.edu) is an assistant professor of management at The Wharton School, University of Pennsylvania. He received his PhD in organizational behavior from the Robert H. Smith School of Business, University of Maryland. His research examines proactive behaviors, such as voice and creativity, and emotion management at work.

Randall S. Peterson (rpeterson@london.edu) is Professor at London Business School. He received his Ph.D. from University of California, Berkeley. His current research focuses on the effects of CEO and leader personality, top team and boardroom dynamics, leading diverse teams, and managing conflict in teams.

Grace Simon (gmsimon@wharton.upenn.edu) is a doctoral candidate in management at the Wharton School, University of Pennsylvania. Her research focuses on the topics of prosocial behaviors, emotions, and interpersonal support at work.