© *Academy of Management Journal*
2015, Vol. 58, No. 6, 1804–1825.
http://dx.doi.org/10.5465/amj.2013.1135

# EMOTIONAL LABOR DYNAMICS: A MOMENTARY APPROACH

**ALLISON S. GABRIEL**
**University of Arizona**

**JAMES M. DIEFENDORFF**
**University of Akron**

**Emotional labor has been described as a dynamic self-regulatory process that unfolds over the course of customer interactions, with employees continuously monitoring and adjusting their felt and expressed emotions via two emotion regulation strategies: surface acting and deep acting. Despite dynamic theory on the topic, empirical tests have largely ignored within-episode variability in emotional labor, relying on assessments of emotional labor focused on the person, day, or interaction level of analysis. The current study elaborated on theory pertaining to within-episode emotional labor dynamics, utilizing a call center simulation to examine how shifts in customer incivility impacted on continuous measures (captured every 200 milliseconds) of participants' felt emotions, surface acting, deep acting, and vocal tone during a single interaction. Results provided evidence that customer behavior causally influences within-episode changes in emotions, emotion regulation, and vocal tone, and that these key emotional labor variables significantly relate to each other at the momentary level of analysis. Further, by modeling lagged effects, we were able to gain insight into the causal direction in the relationships among these continuously measured variables. Moreover, we showed for the first time that surface acting and deep acting are used simultaneously to manage emotional labor demands.**

Emotional labor—that is, the process of employees managing their emotions to conform to organizational expectations during customer interactions (Hochschild, 1983)—is a key topic in organizational research (Elfenbein, 2007). The core premise of emotional labor is that employees must adhere to display rules that specify which emotions should be expressed on the job (Hochschild, 1983; Rafaeli & Sutton, 1987). In many service occupations (including the context of this study), employees are encouraged to display positive emotions and to suppress negative emotions—what has been collectively referred to as "integrative display rules" (Wharton & Erickson, 1993). If employees feel emotions that differ from display rules, it has been theorized that they will actively regulate their emotions to ensure appropriate displays (Grandey, 2000; Hochschild, 1983). Emotional labor research has primarily focused on two regulation strategies: surface acting, which involves faking required emotions and suppressing felt emotions; and deep acting, which involves exerting effort to feel and express required emotions (Grandey, Diefendorff, & Rupp, 2013; Hochschild, 1983). Although many variables have been implicated in emotional labor theory, most models include display rules, felt emotions, emotion regulation, and emotional expressions as core constructs (Diefendorff & Gosserand, 2003; Grandey, 2000).

Over the past 30 years, much has been learned about the relations among emotional labor variables (see meta-analyses by Hülsheger & Schewe, 2011;

This paper is based on work completed by Allison Gabriel as part of her doctoral dissertation, supervised by James Diefendorff, which received the Society for Industrial and Organizational Psychology's 2014 S. Rains Wallace Dissertation Award for best dissertation in industrial/organizational psychology. This paper was accepted for publication when Allison Gabriel was an assistant professor in the Department of Management at Virginia Commonwealth University. We thank Jason Dahling, Alicia Grandey, and Doug Pugh for their helpful feedback on earlier versions of this paper, and Iris Mauss for her guidance on the methods. We also thank Steven Ash, Joelle Elicker, Rosalie Hall, and Paul Levy for their feedback during Allison Gabriel's dissertation process, as well as Matthew Garvin, Kevin Jackson, Adam Krutko, Christopher Petsko, Ariel Roberts, Steven Salapski, Leigh Ann Seibert, Matthew Sloan, and Marc Tolbert for their assistance in data collection, and Michael Rossetti for his assistance with piloting study materials and patience. Finally, we thank Brent Scott and three anonymous reviewers for their constructive feedback and guidance during the review process.

Kammeyer-Mueller et al., 2013; Mesmer-Magnus, DeChurch, & Wax, 2012), with most research focused on broader, trait-level conceptualizations of the constructs (that is, "in general" levels of emotion regulation, dispositional affect, and emotional performance: Brotheridge & Grandey, 2002; Brotheridge & Lee, 2002; Grandey, 2003). Recently, research has considered the relationships among emotional labor constructs over briefer periods, such as entire work shifts (that is, constructs measured at the end of each day: Judge, Woolf, & Hurst, 2009; Scott & Barnes, 2011; Scott, Barnes, & Wagner, 2012; Wagner, Barnes, & Scott, 2014) or short time periods within work shifts (that is, captured at random or fixed intervals during a shift: Beal, Trougakos, Weiss, & Dalal, 2013; Bono, Foldes, Vinson, & Muros, 2007; Totterdell & Holman, 2003). However, careful inspection of much of the theory about how emotional labor variables relate to one another is at an even more finely grained level of analysis than these person-level and event-level studies can capture, describing dynamic processes that unfold within a single customer interaction (Beal & Trougakos, 2013; Côté, 2005; Diefendorff & Gosserand, 2003).

As an example, Rafaeli and Sutton (1987: 28) described emotional labor as follows:

> [T]he initial emotions sent by a focal employee (an "act") stimulate the target person to respond with implicit or explicit feedback about the continuation of the displayed emotion (an "interact"). The sender of emotion reacts to such feedback by readjustments including abandoning, revising, or maintaining the displayed emotion (completing a "double interact").

Building on these ideas, Diefendorff and Gosserand (2003) used the structure of control theory models of behavioral self-regulation (Carver & Scheier, 1998) to conceptualize emotional labor as involving the constant monitoring of discrepancies between emotions and display rules and attempts to reduce observed discrepancies by using emotion regulation strategies. This theory is explicitly dynamic, outlining how moment-to-moment changes in emotional labor variables occur through discrepancy detection and reduction processes. In addition to regulating emotional expressions, this theory suggests that employees also continuously monitor the superordinate goal of satisfying customers, with discrepancies from this goal being observed when customers display negative emotions. Together, these ideas highlight the possibility of *within-episode emotional labor dynamics* and point toward the need to understand the transient, moment-to-moment nature of such processes during employee–customer interactions (Beal & Trougakos, 2013).

The purpose of our paper is to elaborate on the theoretical processes involved in within-episode emotional labor dynamics between customers and employees, and to provide an empirical test of these ideas in a call center simulation in which we manipulated customer (that is, confederate) behaviors and captured continuous assessments of felt emotions, emotion regulation, and emotional vocal tone. Prior emotional labor research often describes within-episode and moment-to-moment processes (see, e.g., Judge et al., 2009; Scott & Barnes, 2011; Scott et al., 2012), but the methodology employed in past studies does not allow for a direct test of these ideas. In contrast, our methodology allowed us to capture fully each of the emotional labor variables over the course of an interpersonal interaction, giving us the opportunity to test the presumed underlying emotional labor dynamics during a customer service interaction. Specifically, we were able to test the causal role of customer behavior in shaping within-episode emotional labor dynamics, and to examine the interrelationships and directionality of influence among continuously assessed felt emotions, emotion regulation (that is, surface acting, deep acting), and vocal expressions of emotions. We also examined a neglected aspect of emotional labor processes—the relationship between surface and deep acting—and whether these strategies should be construed as alternatives that are used independently or can be used simultaneously as individuals regulate different aspects of their emotional experience.

## EMOTIONAL LABOR DYNAMICS

Grandey (2000) described emotional labor as the process of regulating one's emotions to produce organizationally desired emotional displays (see also Hochschild, 1983). In recent years, research has adopted a within-person perspective on emotional labor, testing relationships among variables such as felt emotions, emotion regulation, and emotional expressions over short timespans, such as work shifts or single customer interactions (see, e.g., Beal et al., 2013; Judge et al., 2009; Scott & Barnes, 2011; Scott et al., 2012; Totterdell & Holman, 2003; Wagner et al., 2014). Yet these within-person studies have primarily provided information on how work shifts or customer interactions differ from each other and not how (or why) emotional labor processes change *within* episodes. On the one hand, it may be that there is little substantive variability in emotional

labor within performance episodes (that is, that felt emotions or emotion regulation are fairly constant during an interaction), suggesting that a more micro-level, moment-to-moment analysis is not necessary. On the other hand, several theories of emotion processes (see, e.g., Frijda, 1993; Krumhuber, Kappas, & Manstead, 2013) suggest that experienced and expressed emotions change quickly, and that they may exhibit substantial variability over short time frames (see, e.g., Mauss, Levenson, McCarter, Wilhelm, & Gross, 2005; Mauss et al., 2011). In addition, theoretical descriptions of emotional labor often explicitly discuss within-episode dynamics (Beal & Trougakos, 2013; Côté, 2005; Diefendorff & Gosserand, 2003; Diefendorff & Richard, 2008; Rafaeli & Sutton, 1987, 1990). As such, we suspect that valuable insight may be gained by adopting a within-episode analysis of emotional labor processes.

Beal and Trougakos (2013) opined that emotional labor research has ignored temporal issues, including the possibility that emotion regulation may vary during single interactions as employees experience discrepancies between emotional goals (display rules) and felt emotions. These ideas are consistent with the *self-regulatory view of emotional labor* (Diefendorff & Gosserand, 2003; Diefendorff & Richard, 2008), which argues that emotional labor processes unfold during customer interactions as employees monitor discrepancies between emotions and emotion expectations, and seek to reduce such discrepancies when they occur. From this perspective, display rules represent relatively low-level goals that act in service of higher-level goals aimed at satisfying customers and achieving high-quality emotional performance. According to Diefendorff and Gosserand (2003; see also Diefendorff & Richard, 2008), a low-level discrepancy between felt or expressed emotions and display rules, or a higher-level discrepancy between customer affective displays and the goal of having happy customers, triggers employees to engage in behaviors aimed at reducing the discrepancy. Consistent with general models of emotional labor (Hochschild, 1983), this control theory view suggests that employees may seek to reduce discrepancies by using the emotion regulation strategies of deep and surface acting to show more positive emotions (and to make the customer happier). As noted earlier, deep acting involves changing one's felt emotions to align with emotional display rules, whereas surface acting involves suppressing felt emotions and faking organizationally desired emotions (Hochschild, 1983); in most customer service contexts (including that of the current study), this would involve hiding negative emotions and faking positive ones. Although surface and deep acting are almost always invoked in emotional labor descriptions, little theory or research is aimed at understanding how they relate to each other, including whether they may be used in tandem by employees to produce effective displays and satisfied customers (Beal & Trougakos, 2013). We address this theoretical gap and provide empirical insight into this issue. When considering the dynamic nature of service exchanges (Côté, 2005) and self-regulatory processes (Carver & Scheier, 1998), we therefore expect to observe meaningful variability in felt emotions, emotion regulation, and emotional vocal tone during the course of an exchange with a customer:

*Hypothesis 1. (a) Felt emotions, (b) surface acting, (c) deep acting, and (d) vocal tone exhibit variability within a single emotional labor episode.*

### Impact of Customer Behavior

Assuming that we would observe within-episode emotional labor variability, we sought to understand *why* such variability might exist. A key objective of emotional labor is to produce effective emotional displays that enhance the customer experience (Diefendorff & Gosserand, 2003; Grandey, 2000; Hochschild, 1983). Thus one variable monitored by employees during emotional labor episodes is how the customer behaves, including how pleasant or friendly the customer is to the employee (Diefendorff & Richard, 2008). In control theory terms, the goal would be a satisfied customer, with this goal being constantly monitored during an interaction; when a discrepancy from the goal is sensed, the employee is motivated to reduce the discrepancy and make the customer happy (Diefendorff & Richard, 2008). This idea is consistent with the process described by Rafaeli and Sutton (1987), in which employees and customers mutually influence each other over the course of an interaction. It is also consistent with Côté's (2005) social interaction model of emotion regulation, which outlined how customer reactions to employee emotion regulation attempts can determine whether the employee experiences that regulation as harmful or beneficial. In essence, effortful emotion regulation may produce employee benefits when this regulation produces a happy customer, who in turn treats the employee well. Thus the social benefits of regulation can outweigh the costs of exerting oneself (see, e.g., Heaphy & Dutton, 2008). These ideas suggest

that customer behavior plays a prominent role in shaping employee behavior and well-being over the course of an emotional labor episode.

Consistent with the self-regulatory view of emotional labor (Diefendorff & Gosserand, 2003), we theorized that uncivil customer behavior produces a discrepancy from the objective of satisfying the customer. Employees may then respond to this discrepancy by enhancing the positivity of their own emotional displays in an attempt to appease the uncivil customer. Additionally, employees may feel negative emotions in response to mistreatment (Dallimore, Sparks, & Butcher, 2007), which will produce a discrepancy between feelings and display rules, or what has been labeled emotion–rule dissonance (Grandey et al., 2013). Such a discrepancy signals to employees that they must regulate their feelings and/or expressions so as to show the required emotions, which they may accomplish by attempting to alter their felt emotions to make them more positive (deep acting) or by manipulating how positive they sound (surface acting).

To test systematically the causal role of customers in shaping within-episode emotional labor processes, we manipulated customer (confederate) behavior at two points during a call center simulation. At the onset of the call (Phase I), customers were affectively neutral, which was intended to establish a baseline for the interaction. We suspected that this neutral customer demeanor would signal to participants that things were going reasonably well, resulting in little need to regulate emotions. However, at a specific point in the call (the start of Phase II), we introduced a negative change in customer demeanor, which was expected to produce an increase in employee negative feelings and an increase in emotion regulation. Past research has linked customer incivility to employee negative emotions (Rupp & Spencer, 2006; Wegge, Vogt, & Wecking, 2007; Yang & Diefendorff, 2009) and emotion–rule dissonance (Mesmer-Magnus et al., 2012). Rupp and Spencer (2006) found that exposure to manipulated customer incivility corresponded to increased self-ratings of emotional labor (that is, the effort exerted in regulating one's emotions) and observers' perceptions that participants were engaging in emotion regulation.

In addition to impacting on internal processes, we expected that customer demeanor would negatively impact employee vocal tone (that is, that employees would sound more negative). However, we anticipated that the effect on vocal tone would be smaller than the effect on feelings and regulation precisely because the increased regulation is aimed at maintaining a positive demeanor despite customer mistreatment. Although efforts to regulate vocal tone—the primary emotional behavior in call center interactions (Van Jaarsveld & Poster, 2013)—should largely prevent negative vocal emotional expressions, we did not expect that vocal tone when being mistreated would be as positive as vocal tone when the customer is affectively neutral (and employee feelings are more positive and there is less need to regulate). Moreover, research has found that felt emotions relate to expressions even when the individual is attempting to conceal feelings (Ekman, Friesen, & O'Sullivan, 1988), suggesting that the voice is a somewhat "leaky channel" (Ekman & Friesen, 1969; Elfenbein & Ambady, 2002) and that customer incivility will translate into more negative vocal tone. Taken together, we hypothesize that:

> *Hypothesis 2. In response to a within-episode increase in customer incivility, (a) felt emotions become more negative, (b) surface acting increases, (c) deep acting increases, and (d) vocal tone becomes more negative.*

Research manipulating customer incivility or measuring it at the end of a customer interaction implicitly assumes that incivility is stable during the interaction, such that some customers are uncivil and others are not (see, e.g., Goldberg & Grandey, 2007; McCance, Nye, Wang, Jones, & Chiu, 2013; Rupp & Spencer, 2006). However, we theorized that customer behavior is also dynamic: While some customers may persist in being uncivil (or become even more uncivil), other customers may shift from being rude to being pleasant, which would presumably impact on the moment-to-moment changes in emotional labor during an interaction. Andersson and Pearson (1999; see also Groth & Grandey, 2012) theorized about "incivility spirals," in which the customer and employee can become increasingly hostile to one another in a tit-for-tat exchange, suggesting a linear trend toward greater negativity by both parties. However, it is possible that a difficult service interaction can be repaired, resulting in a once-uncivil customer becoming pleasant, yielding a curvilinear trend in customer incivility, and (possibly) a curvilinear trend in employee affect, regulation, and vocal tone. These ideas resonate with the assertion of Côté (2005; see also Rafaeli & Sutton, 1987) that how the customer behaves during an interaction can produce within-episode changes in the feelings and behavior of employees.

To fully examine these ideas, we introduced a between-subjects manipulation in the final phase of the

call (the start of Phase III) in which the customer who was uncivil in Phase II either became more uncivil (signaling service failure and a continued negative trajectory) or became civil (signaling service recovery and a change in the trajectory). Consistent with our self-regulatory view of emotional labor, we theorized that further increases in customer incivility would trigger additional changes in emotional labor processes that extend the same trajectory proposed in Hypothesis 2 (that is, increased felt negative emotions, emotion regulation, and negative vocal tone). However, a positive change in the customer should signal a discrepancy reduction, producing improved employee felt and expressed emotions, and a reduced need to regulate feelings and displays. We predicted the following for service failure and recovery in Phase III:

*Hypothesis 3. In response to within-episode customer changes from incivility to further incivility, (a) felt emotions become more negative, (b) surface acting increases, (c) deep acting increases, and (d) vocal tone becomes more negative.*

*Hypothesis 4. In response to within-episode customer changes from incivility to civility, (a) felt emotions become more positive, (b) surface acting decreases, (c) deep acting decreases, and (d) vocal tone becomes more positive.*

### Dynamic Interplay of Felt Emotions, Emotion Regulation, and Vocal Tone

In addition to the role of external customer influences, the self-regulatory approach to emotional labor provides insight into how felt emotions, displayed emotions, and emotion regulation may relate to each other over time. Diefendorff and Gosserand (2003) suggested that a central goal of emotional labor is to monitor internal emotional states and, when needed, regulate them by means of surface and/or deep acting to ensure a steady outward affective demeanor (that is, vocal tone). Prior theory has described causal links among these variables using language that suggests within-episode dynamic processes (Judge et al., 2009; Scott & Barnes, 2011). Although our description of emotional labor is not at odds with these ideas, our data can provide the first (to our knowledge) test of these ideas at the momentary, within-episode level of analysis. In addition to providing a constructive replication of the relationships among emotional labor variables at a within-episode level (Lykken, 1968), our unique data structure enabled us (a) to examine the stability of each variable over time (see, e.g., Ebner-Priemer & Trull, 2011), (b) to test the presumed causal progression among the emotional labor variables (emotions → regulation → vocal tone) within a performance episode, and (c) to explore the possibility of reciprocally causal relations among the variables. We also considered how the emotion regulation strategies relate to each other and whether they are mutually exclusive (that is, negatively related) or can be used together to manage difficult situations (that is, positively related).

Consistent with our previous discussion, we theorized that changes in customer behavior from Phase I to Phase II and from Phase II to Phase III would produce changes in employee felt emotions. However, it is also likely that there may be moment-to-moment fluctuations in felt emotion that are idiosyncratic to the individual or specific interaction. Fitting with Diefendorff and Gosserand (2003), changes in felt emotions should impact on the discrepancy between feelings and display rules, impacting on the amount of emotion regulation performed by employees. Both feelings and regulation are presumed to contribute to the emotional tone of participants' voices. In general, felt emotions (that is, hedonic tone, ranging from very negative on the low end, neutral at the midpoint, and very positive on the high end) and emotion regulation (ranging from none on the low end to a very large extent on the high end) should be inversely related: When employees feel positively, they have little need to regulate, because there is unlikely to be a sensed discrepancy (see, e.g., Carver & Scheier, 1998). Thus we anticipated that felt emotions at a point in time should negatively relate to emotion regulation and positively relate to vocal tone:

*Hypothesis 5. Within an emotional labor episode, momentary felt emotions are negatively related to momentary (a) surface acting and (b) deep acting; as felt emotions become more negative, the use of surface acting and deep acting increases.*

Additionally, fitting with our discussion of the voice being a leaky emotional channel (Elfenbein & Ambady, 2002; Rosenthal & DePaulo, 1979; Zuckerman, Amidon, Bishop, & Pomerantz, 1982), we expected that what individuals feel may leak through and impact on the tone of their voice. We did not expect participants to express strong negative emotions (such as anger) even if they felt such emotions, because doing so would violate display rules (Grandey, 2000). As previously noted, it is likely that moment-to-moment changes in felt emotions would

be larger than changes in vocal tone. Nonetheless, we expected felt emotions to covary with emotions expressed in the voice, such that felt negative emotions would be associated with less positive vocal tone, whereas felt positive emotions would be associated with a more positive vocal tone:

*Hypothesis 6. Within an emotional labor episode, momentary felt emotions are positively related to momentary vocal tone; as felt emotions become more negative, vocal tone also becomes more negative.*

According to most descriptions of emotional labor, the purpose of emotion regulation is to produce effective emotional displays. Consistent with meta-analytic work (Hülsheger & Schewe, 2011) and the idea that deep acting aligns felt emotions with expectations to produce appropriate displays (Grandey, 2000; Hochschild, 1983), we theorized that deep acting would positively relate to vocal tone. For surface acting, the relation with vocal tone is less clear. One might expect surface acting to improve vocal tone, because the strategy is aimed at reducing display rule discrepancies (Diefendorff & Gosserand, 2003). Indeed, Beal, Trougakos, Weiss, and Green (2006) found that surface acting at the event level positively predicted supervisor affective delivery ratings. However, Grandey (2000) theorized that surface acting is less authentic, with empirical work linking surface acting to less positive vocal tone (Tsai & Huang, 2002) and affective delivery (Grandey, 2003). Further, suppressing one's emotions during a conversation has been shown to harm interpersonal functioning (Butler, Egloff, Wilhelm, Smith, Erickson, & Gross, 2003). Even so, it may be that surface acting will not relate to vocal tone at the momentary level after controlling for the effects of felt emotions, because the presumed detriment of surface acting is that it does not fully conceal a person's felt negativity. However, we presume that the inauthenticity associated with surface acting may prevail to negatively impact on vocal tone, even after controlling for the effects of felt emotions. As such, we expected that high surface acting would lead to worse vocal tone ratings than low surface acting:

*Hypothesis 7. Within an emotional labor episode, momentary surface acting is negatively related to momentary vocal tone.*

*Hypothesis 8. Within an emotional labor episode, momentary deep acting is positively related to momentary vocal tone.*

The final within-episode relationship that we consider is between surface acting and deep acting. Surprisingly, past research is largely silent on the interrelationship among these strategies, although some expectations might be derived from Gross's (1998) process model. Gross (1998) developed a model of emotion regulation that takes the emotion generation process and identifies when different emotion regulation strategies are likely to occur. In his model, an emotional stimulus is encountered and individuals can regulate their emotions by utilizing antecedent-focused strategies (that is, deep acting: Grandey, 2000) before the emotion is fully formed, or response-focused strategies (that is, surface acting: Grandey, 2000) after an emotion occurs. On the surface, this model might suggest that individuals use deep acting early in the process *or* use surface acting late in the process. Indeed, such a view is consistent with many descriptions of these strategies as being mutually exclusive alternatives for regulating one's emotions (Austin, Dore, & O'Donovan, 2008; Kruml & Geddes, 2000; Zapf, 2002). These ideas suggest that, at the momentary level of analysis, the relationship between surface acting and deep acting should be negative, because one strategy is used at a time in isolation.

Despite this theorizing, meta-analytic results indicate that the strategies are positively, although weakly, correlated ($\rho$ = .22; Hülsheger & Schewe, 2011), suggesting that people may be more likely to use both than to use one at the expense of the other. Indeed, recent work using latent profile analysis identified a subset of emotional labor actors that report using both strategies at a high level (Gabriel, Daniels, Diefendorff, & Greguras, 2015). Perhaps more telling is that within-person research has also found a weak positive correlation between the two strategies (for example $r$ = .20 in Judge et al., 2009; $r$ = .07 in Scott & Barnes, 2011), suggesting again that the strategies may co-occur. Of course, the positive correlations in past studies could be the result of some people, days, or situations requiring more emotion regulation than others and not of individuals simultaneously using the strategies. For instance, a positive day-level relationship between surface and deep acting means that some days require more regulation than others, but does not preclude the possibility that individuals are using one strategy at a time and vacillating between them within a given day. However, we propose that employees may use surface acting and deep acting simultaneously to address different aspects of the discrepancy between felt emotions and display

rules, with surface acting primarily targeting outward expressions and deep acting targeting internal feelings. As suggested by Beal and Trougakos (2013), individuals may reduce this discrepancy by hiding undesirable feelings (surface acting) while also trying to change these negative feelings so that they are more positive (deep acting). For instance, an employee faced with customer incivility may grit his or her teeth and smile all while trying to reinterpret the situation in a more positive way. Although the idea that both strategies may be used concurrently may seem contrary to models of the emotion regulation process (Gross, 1998), such models are largely silent on the issues of timing and duration of regulation strategies. Further, empirical tests typically do not track strategy use over time and across changing emotional stimuli. Thus we hypothesize that, when faced with a dynamic and difficult customer service encounter, individuals will utilize the strategies together to address the overarching problem of making the customer happy and handling their own changing emotions, yielding a positive moment-to-moment relationship between surface acting and deep acting:

*Hypothesis 9. Within an emotional labor episode, momentary surface acting is positively related to momentary deep acting.*

### Exploring Causality among Emotional Labor Variables

Although these hypotheses could be considered in isolation, they suggest a causal sequence, with felt emotions triggering surface acting and deep acting, and all three variables impacting on vocal tone (that is, emotional performance: Grandey, 2000). A significant limitation of past research examining relations among these (and similar) variables is that the data are not modeled longitudinally, preventing inferences about causality. Even studies with longitudinal data typically do not attempt to examine how the emotional labor variables relate to each other over time, looking instead at relationships among variables within particular time periods (see, e.g., Scott et al., 2012). Our continuous measurement of the emotional labor variables permits us to model lagged values of independent and dependent variables simultaneously when predicting a dependent variable, a condition that is needed to isolate causal direction of effects (Granger, 1969). Thus, in addition to testing concurrent relations among variables, we modeled a full set of lagged relationships, separating the variables into three time periods that correspond to the presumed causal progression of felt emotions → emotion regulation → vocal tone. Such an approach enabled us to examine the influence of earlier levels of independent variables on changes in the dependent variables, which helps to isolate the causal direction of effects (Cole & Maxwell, 2003).

Our data structure also enabled us to examine the possibility of reciprocally causal relations among our variables (see, e.g., Hülsheger, Lang, & Maier, 2010). The self-regulatory view of emotional labor (Diefendorff & Richard, 2008) suggests that employees may compare not only their feelings, but also their expressions, with display rules to determine if a discrepancy exists. Thus a negative change in vocal tone might produce a discrepancy from the display rule and result in a subsequent increase in emotion regulation. In contrast, an improvement in vocal tone might signal to employees that the discrepancy has been reduced, resulting in less subsequent surface and deep acting (Diefendorff & Gosserand, 2003; Rafaeli & Sutton, 1987). As such, in addition to emotion regulation impacting on subsequent vocal tone, vocal tone may influence emotion regulation as employees detect that more or less effort is needed to adhere to the display rule. Moreover, although models suggest that felt emotions precede expressed emotions (Grandey, 2000; Gross, 1998), research on the facial feedback hypothesis and related theories (see, e.g., Strack, Martin, & Stepper, 1988) suggests that showing an emotion can impact on how one feels; we explore this causal direction in the data.

Additionally, emotional labor models (Diefendorff & Gosserand, 2003; Grandey, 2000) often include affect (and the comparison of emotions against display rules: Grandey et al., 2013) as preceding emotion regulation. Yet daily diary research has positioned surface and deep acting as antecedents of felt affect (Judge et al., 2009; Scott & Barnes, 2011; Wagner et al., 2014). The theoretical rationale is that surface acting (that is, being fake) acts as a stressor, making individuals feel worse, whereas deep acting with its focus on changing feelings, may improve affect. Thus exploring causality between felt emotions and emotion regulation warrants more attention.

Finally, we consider whether surface and deep acting exhibit causal precedence on each other. As mentioned earlier, Grandey (2000; see also Gross, 1998) argued that deep acting is an antecedent-focused strategy, whereas surface acting is a response-focused strategy, suggesting that the use of deep acting may result in individuals not needing to use surface

acting. However, there is some evidence that surface acting, with its focus on suppression, may occur quickly and be the first response to unexpected situations, whereas deep acting, which may require more deliberate control (for example cognitively reappraising a situation) may occur more slowly (Beal & Trougakos, 2013). As Beal and Trougakos (2013) mentioned, in difficult situations, employees may first rely on an easier mode of regulation (such as displaying a fake smile), but as time elapses they may try to use a more effortful strategy (that is, try to create positive feelings to yield a more genuine customer experience). Given our continuous data structure, we were able to tease apart the causal ordering of these strategies.

## METHODS

### Participants, Design, and Procedure

We recruited 84 undergraduate students to participate in exchange for course credit and a gift card. We dropped eight participants because poststudy checks revealed that they did not accurately perceive the manipulated civility of the customer across the three phases of the calls (for example seven participants indicated that the caller was uncivil the entire call; one participant indicated that the call started uncivilly and ended politely, when the actual progression was the opposite). However, testing the hypotheses with or without these individuals did not alter the substantive conclusions reported in this paper. The resulting sample of 76 participants was mostly female (65.8%) and Caucasian (69.7%), with an average age of 20.94 ($SD$ = 4.99).

Prior to the experimental session, participants completed an online survey containing demographic variables. During the experiment, participants played the role of a call center employee in a 90-minute simulation. Participants were told that the study was in partnership with a local call center and that the purpose of the study was to pilot new training materials aimed at providing customer service over the phone. To do this, they were told that they would take actual calls from students at another university needing help to create presentations in Microsoft PowerPoint; this was based partially on McCance and colleagues (2013). Participants were told that the students whom they were assisting could win US$100 if they were the first in their session to finish their PowerPoint, helping to justify the task's importance (and the possibility that callers may act emotionally). Similarly to previous research (Goldberg & Grandey, 2007; McCance et al., 2013) and service contexts (Grandey et al., 2013), we established a positive display rule ("We want you to be *happy*, *positive*, and *respectful* to all customers calling into the hotline"). Given that display rules may vary slightly for individuals (see, e.g., Diefendorff & Richard, 2008), we enforced the display rule in the scripted greeting, follow-up, and closing statements that participants had to follow ("My name is ___ and I am happy to serve you", "We are always happy to help you through any problem") and through signage within the laboratory space (such as posters encouraging "service with a smile"). Participants were told that their calls would be monitored for service quality and that caller evaluations of their performance would determine if the participants were eligible to win a gift card.

Before interacting with a customer, participants went through a practice call with the researcher. The practice call lasted 82 seconds on average and was intended: (a) to allow participants to practice the greeting, follow-up, and closing statements; (b) to give participants a chance to practice providing step-by-step PowerPoint instructions; and (c) to let participants practice providing continuous ratings of felt emotions and emotion regulation after the call. Postpractice, participants began the simulation and took one phone call; they did not know how many calls they would receive. Although participants were given a script for the greeting, follow-up, and closing statements, the remainder of their dialogue was unscripted and dependent on the questions asked by the caller. Participants were unaware of what they would be asked by the callers and were unaware of how they would be treated. Importantly, confederate callers followed the same script for all participants. Immediately postcall, participants listened to a recording of the call three times. Each time, participants provided continuous ratings of a different variable: felt emotions, surface acting, and deep acting. Although ratings were made postcall (instead of during the actual call), playing back the conversation helped to alleviate memory biases in the ratings (Ruef & Levenson, 2007). Additionally, Mauss and colleagues (2005) found that cued-recall continuous ratings after a stimulus presentation correlated highly with online continuous ratings during a stimulus presentation ($r$ = .80 for amusement, $r$ = .73 for sadness), supporting the idea that ratings from cued recall closely correspond to ratings provided during lived experiences.

Following past precedent (see, e.g., Goldberg & Grandey, 2007), our seven confederate callers were

male, to eliminate a gender effect.[1] Based on the within- and between-subject manipulations, the calls were divided into three phases. The within-subject manipulation for the first two phases was the same for all participants, such that all calls started with customers being civil in Phase I (neutral affective tone, no complaints) and uncivil in Phase II (negative affective tone, complaints). For the between-subjects manipulation, participants were randomly assigned to the service recovery ($n = 38$) or the service failure ($n = 38$) condition. For the service recovery condition, the caller shifted to being civil (positive affective tone, conversational) in Phase III. In the service failure condition, the caller shifted to being even more uncivil (heightened negative affective tone, complaints, insults) in Phase III. To distinguish the three phases of each call, given that each call unfolded over a slightly different period of time, a research assistant blind to study condition and hypotheses listened to recordings of the calls in which only the confederate's voice could be heard, and marked the two points in which the call increased or decreased in civility (that is, marked the experimental phases).

## Measures

Lamenting the fact that static scales cannot capture fluctuations in emotional states, Ruef and Levenson (2007: 47) made the case for collecting data "on a moment-to-moment basis, either *on-line* as the emotion is first experienced or *retrospectively* as the temporal dimension of the original episode is 'replayed' while real-time momentary self-report measures [are collected]." Building on these ideas, we collected continuous ratings via a program in E-Prime 2.0 Professional that enabled us to play back recordings of the calls and have participants provide continuous ratings of how they felt, how much they surface acted, and how much they deep acted. The program recorded participants' mouse movements along a 20-point scale every 200 milliseconds. We sampled participant responses every 200 milliseconds to provide "near-continuous data" (Mauss et al., 2005: 180). This sample rate was chosen to strike a balance between capturing the maximum possible by the software package (1,000 ratings per second, which would be likely to be many more observations than is theoretically meaningful) and still oversampling the ratings of constructs so as to not miss any changes. Participants first provided ratings of felt emotions and then ratings of either surface acting followed by deep acting or vice versa; the order of the ratings for the regulation strategies was counterbalanced. A similar approach using continuous ratings with retrospective recordings was employed in marital research in which husbands and wives were asked to talk with their spouses about particular issues, then each person watched a video of the conversation and continuously rated how he or she felt during the interaction (Gottman & Levenson, 1985; Levenson & Gottman, 1983).

***Continuous ratings of participant emotionality.*** Participants rated how they felt during the call on a 20-point scale (1 = *very negative*; 10 = *neutral*; 20 = *very positive*), which is similar to other continuous ratings of emotion (see, e.g., Fredrickson & Kahneman, 1993). Felt emotions ratings began with the cursor at "10" (neutral). Participants were instructed to move the mouse as much or as little as they liked depending on how they felt during the call.

***Continuous ratings of surface acting.*** Participants rated the extent to which they hid their felt emotions from the customer ("How much did you hide your emotions?"). The wording of the item focused participants on the extent to which they engaged in emotional suppression, which is a key component of surface acting in the emotional labor literature (Brotheridge & Lee, 2002; Diefendorff, Croyle, & Gosserand, 2005; Grandey, 2003) and also fits with emotional suppression in the emotion regulation literature (Gross, 1998), which is closely connected to emotional labor (Grandey, 2000). The 20-point scale had six anchors, ranging from no emotion regulation to regulating to the largest extent possible (1 = *none*; 4.5 = *very small extent*; 8.5 = *small extent*; 12.5 = *moderate extent*; 16.5 = *large extent*; 20 = *very large extent*). Surface acting ratings began with the cursor at "1" (that is, none).

***Continuous ratings of deep acting.*** Participants provided ratings of the extent to which they changed their emotions to feel the emotions that they were to express with the caller ("How much did you try to change your emotions?"); the wording matches descriptions of deep acting in the literature (Diefendorff et al., 2005; Grandey, 2003; Kruml & Geddes, 2000). In the verbal and written instructions, participants were told that deep acting reflects exerting effort to try to feel positively about the situation or the interaction with a caller in order to

[1] We tested six dummy codes as Level-2 control variables in our analyses to see if confederate affected ratings of felt emotions, surface acting, deep acting, and vocal tone. Results remained unchanged.

express positive emotions. The same scale anchors used for surface acting were used for deep acting.

***Continuous ratings of employee vocal tone.*** Two trained research assistants who were blind to study condition and hypotheses separately rated participants' vocal tone using the rating scale that participants used for felt emotions. The confederate's voice was removed from the audio recordings, preventing it from impacting on the ratings of participant vocal tone (Rupp & Spencer, 2006). Raters were instructed to adjust their ratings when a change in the affective quality of a participant's voice occurred. When there was silence, raters were instructed to not adjust their ratings. The average of the two ratings was used to operationalize vocal tone. To support this decision, we looked at the average within-person correlation between the two third-party ratings and found that it was 0.60 ($p < .001$), suggesting high convergence between the two raters.

## Analytic Approach

Following past precedent (Gottman & Levenson, 1985; Mauss et al., 2005), the 200-millisecond continuous ratings were aggregated to 1-second averages. Given that the average call lasted 4 minutes and 24 seconds, this resulted in an average of 264.75 data points per participant. Service failure calls lasted an average of 255.37 seconds and service recovery calls lasted an average of 274.13 seconds; the difference in call length was not significant ($t_{(74)} = 1.72$, n.s.), and controlling for call length in analyses did not affect results. Because our data were nested (measurements nested within participants), we conducted analyses using multilevel modeling (MLM) in MPlus 7.11 (Muthén & Muthén, 1998–2013). We first tested whether there was sufficient within-person variability in the emotional labor variables. Within-person variability accounted for substantial total variance in each construct (felt emotions, 56.63%; surface acting, 62.47%; deep acting, 55.73%; vocal tone, 58.16%) supporting MLM.

To test Hypotheses 2–4, which predicted changes in emotions, emotion regulation, and vocal tone across phases of the call, we used codes to model growth factors corresponding to the three phases of the call, with follow-up bivariate codes used to compare particular phases. For Hypotheses 5–9, we utilized within-person path analysis to test relationships among the emotional labor variables using current values of all variables (T), as well as two lag values (T-10 and T-20 seconds). Given our focus on within-person processes only, all effects were modeled as fixed effects (Kreft & DeLeeuw, 1998). This produced a data structure similar to a cross-lagged panel design with three occasions (see, e.g., Cacioppo, Hawkley, & Thisted, 2010), and enabled us to examine the stability of the variables and how the variables related to each other concurrently and between time periods, to tease apart causality. The lag amount of 10 seconds was chosen after assessing the stability of the constructs across a range of lag amounts, as well as testing the causal relationships among the variables over the various lag amounts. To examine stability, we calculated within-person autocorrelations of each variable at a point in time (T), with lags of the same variable ranging from T-1 to T-20 seconds (see Table 1 for a sampling of these values). Each measure exhibited high stability, with 2-second lags having autocorrelations between 0.97 and 0.98 (accounting for over

**TABLE 1**
**Within-Person Auto-Correlations for Various Time Lags**

| Lag quantities | Felt emotions | Surface acting | Deep acting | Vocal tone |
|---|---|---|---|---|
| T-2 seconds | 0.97 | 0.98 | 0.97 | 0.98 |
| T-4 seconds | 0.92 | 0.94 | 0.93 | 0.94 |
| T-6 seconds | 0.88 | 0.91 | 0.89 | 0.91 |
| T-8 seconds | 0.84 | 0.88 | 0.85 | 0.87 |
| T-10 seconds | 0.80 | 0.85 | 0.82 | 0.84 |
| T-12 seconds | 0.76 | 0.82 | 0.78 | 0.81 |
| T-14 seconds | 0.72 | 0.79 | 0.75 | 0.79 |
| T-16 seconds | 0.68 | 0.76 | 0.72 | 0.76 |
| T-18 seconds | 0.65 | 0.73 | 0.68 | 0.73 |
| T-20 seconds | 0.61 | 0.69 | 0.65 | 0.70 |

*Note*: Focal variables are assessed at time T and correlated with the lagged assessment of themselves at T-2–T-20; variables were group-mean-centered prior to analyses; all correlations are significant at $p < .001$. For brevity, we present only the within-person correlations for lags every 2 seconds (full analyses for all lags are available upon request from the first author).

94% of the variance in T measurements) and 20-second lags having autocorrelations between 0.61 and 0.70 (accounting for 37–49% of the variance in T measurements). Further, most of the anticipated causal relationships were detected across the various lag lengths.

In an attempt to strike a balance between having very little variance left in the variables after modeling autocorrelations for short lags (for example 1-second lags, in which the autocorrelations would leave less than 10% of their variance in our dependent variables to be predicted) and having very long lags that may stretch the time frame of causal effects (for example a 20-second lag may be slower than the presumed psychological process of felt emotions leading to emotion regulation and emotion regulation leading to vocal tone), we modeled 10-second lags in the model that we report here. In sum, our final path model included: (a) the autocorrelations of each variable across time periods (that is, the lagged value of a variable predicting itself); (b) the concurrent relationships among variables within a time period (to understand momentary relationships among the residuals); (c) the theorized causal paths across time periods (for example felt emotions at T-20 → emotion regulation at T-10 → vocal tone at T); (d) the causal paths reflecting reverse causality; and (e) exogenous growth curve codes representing the within-person and between-person customer incivility manipulations.

## RESULTS

Means, standard deviations, and within-person correlations of our continuously rated variables are reported in Table 2. In support of Hypothesis 1, there was a large amount of within-episode variability, highlighting that, within the customer service encounter, individuals' felt emotions, emotion regulation, and vocal tone varied. To illustrate these findings, Figure 1 plots the average levels of the continuous rating variables across the entire call length for the service failure (Figure 1a) and service recovery (Figure 1b) conditions, respectively. These plots show the ebb and flow of felt emotions, emotion regulation, and vocal tone during a single interaction, suggesting that using average or end-of-call ratings would have missed substantial variability during the episode. Interestingly, the variability for vocal tone was much smaller than that of felt emotions and emotion regulation, fitting with the control theory idea that individuals are trying to maintain a steady outward demeanor despite large changes in internal feelings and emotion regulation.

### Impact of Customer Behavior (Hypotheses 2–4)

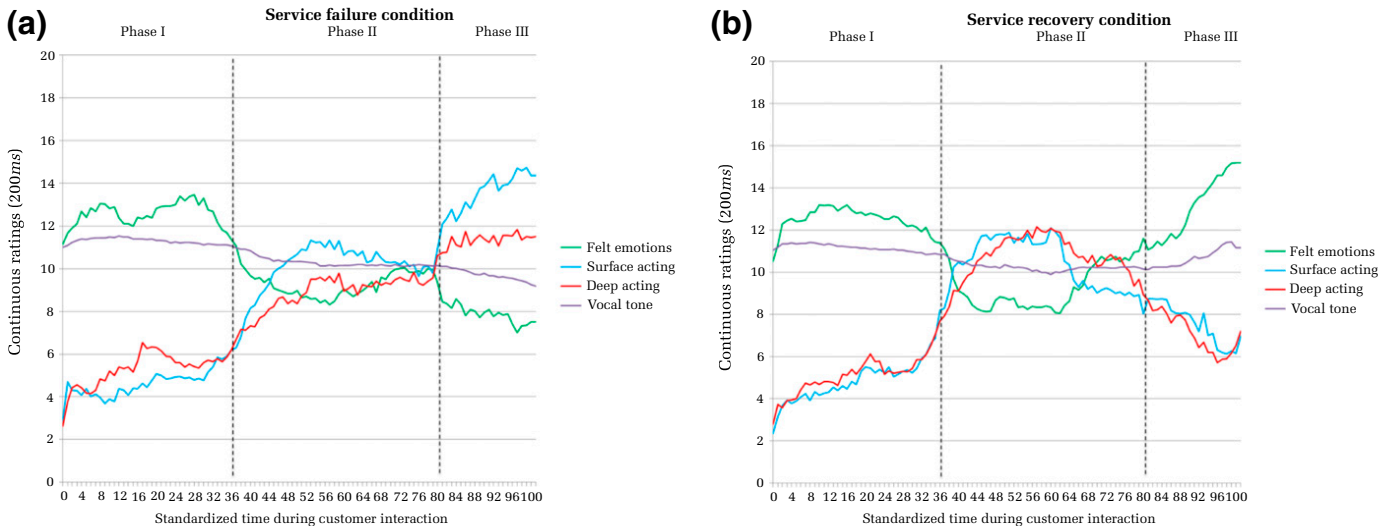
Hypotheses 2–4 focused on the causal role that customers play in shaping emotional labor dynamics. We anticipated that a shift from customer neutrality in Phase I to customer incivility in Phase II would cause participants to (a) feel more negative emotions, (b) increase surface acting, (c) increase deep acting, and (d) express less positive emotions vocally (Hypothesis 2). We expected that these effects would continue in the same direction during Phase III (a linear trend) for the service failure condition (Hypothesis 3), but change direction (a curvilinear trend) for the service recovery condition (Hypothesis 4). Results for these hypotheses can be found in Table 3.

**TABLE 2**
**Within-Person Correlations among Concurrent and Lagged Variables**

| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Felt emotions – T | – | | | | | | | | | | | |
| 2 | Felt emotions – T-10 | 0.80 | – | | | | | | | | | | |
| 3 | Felt emotions – T-20 | 0.61 | 0.78 | – | | | | | | | | | |
| 4 | Surface acting – T | –0.61 | –0.57 | –0.49 | – | | | | | | | | |
| 5 | Surface acting – T-10 | –0.53 | –0.61 | –0.56 | 0.85 | – | | | | | | | |
| 6 | Surface acting – T-20 | –0.42 | –0.52 | –0.60 | 0.69 | 0.84 | – | | | | | | |
| 7 | Deep acting – T | –0.59 | –0.56 | –0.49 | 0.65 | 0.60 | 0.52 | – | | | | | |
| 8 | Deep acting – T-10 | –0.51 | –0.58 | –0.54 | 0.59 | 0.65 | 0.61 | 0.82 | – | | | | |
| 9 | Deep acting – T-20 | –0.40 | –0.49 | –0.56 | 0.51 | 0.59 | 0.66 | 0.65 | 0.81 | – | | | |
| 10 | Vocal tone – T | 0.42 | 0.44 | 0.42 | –0.49 | –0.49 | –0.47 | –0.41 | –0.43 | –0.42 | – | | |
| 11 | Vocal tone – T-10 | 0.34 | 0.40 | 0.42 | –0.45 | –0.49 | –0.50 | –0.36 | –0.41 | –0.43 | 0.84 | – | |
| 12 | Vocal tone – T-20 | 0.27 | 0.33 | 0.39 | –0.40 | –0.45 | –0.48 | –0.30 | –0.36 | –0.41 | 0.70 | 0.85 | – |

*Note:* Correlations are all within-person and group-mean-centered (rated by two third-party raters); T, T-10 seconds, and T-20 seconds represent the different time lags used in our path analytic model; all correlations are significant at $p < .001$.

**FIGURE 1**
**Within-Person Trajectories for the (a) Service Failure and (b) Service Recovery Conditions**




*Note*: Because the calls and the amount of time in each phase were different per participant, we rescaled the time variable such that each participants' call would be on the same 100-point time metric, with Phase I representing about 36% of the call, Phase II representing about 44% of the call, and Phase III representing about 20% of the call. (These percentages mirror the average duration spent in each phase across participants.)

As a first step, we examined models that included a linear growth code (Phase I = −1, Phase II = 0, Phase III = 1) for all participants. We then introduced condition codes that distinguished Phase III for the service failure (1, 1, 1) and service recovery (1, 1, −1) conditions. When multiplying the growth code by these condition codes, it produced the same linear trend for the service failure condition (–1, 0, 1), but added a curvilinear trend for the service recovery condition (–1, 0, −1). Using the change in −2 log likelihood (see Table 3), we observed that the model adding the curvilinear trend for the service recovery condition fit the data better than the linear-only model. After establishing that the combined linear/curvilinear model best fit the data, we ran follow-up analyses to test our hypotheses, which compared two phases at a time.[2]

In support of Hypothesis 2, we found that, within and across experimental conditions, participant felt emotions became more negative, surface and deep acting increased, and vocal tone became more negative when moving from Phase I to Phase II (see Table 3). For instance, across participants, felt emotions started out positive (Phase I: $M = 12.49$) and became less positive when customers became uncivil (Phase II: $M = 9.28$). We also found support for Hypothesis 3, which pertained to the service failure condition: Felt emotions and vocal tone were more negative when moving from Phase II (felt emotions: $M = 9.28$; vocal tone: $M = 10.29$) to Phase III (felt emotions: $M = 7.87$; vocal tone: $M = 9.73$), and both surface and deep acting increased when moving from Phase II (surface acting: $M = 10.13$; deep acting: $M = 8.91$) to Phase III (surface acting: $M = 13.60$; deep acting: $M = 11.36$). Thus, with customer incivility increasing further, the ratings of all four emotional labor variables exhibited linear trends across the three phases.

We also found support for Hypothesis 4, which pertained to the service recovery condition. In response to the switch from customer incivility to civility, felt emotions and vocal tone became more positive when moving from Phase II (felt emotions: $M = 9.28$; vocal tone: $M = 10.25$) to Phase III (felt emotions: $M = 13.04$; vocal tone: $M = 10.61$), and levels of surface and deep acting decreased when moving from Phase II (surface acting: $M = 10.42$; deep acting: $M = 10.75$) to Phase III (surface acting: $M = 7.59$; deep acting: $M = 7.11$). Across the three phases, participants in the service recovery condition exhibited curvilinear trends in all four variables. In sum, these results are consistent with theory proposed by Diefendorff and Gosserand (2003) and by Côté (2005) in that customer behavior impacted on changes in emotional labor variables, perhaps signaling the emergence and increase of

[2] These analyses are analogous to conducting a repeated-measures analysis of variance (ANOVA) running condition (service recovery and service failure) as a between-subject factor and phase (I, II, and III) as a within-subject factor, with felt emotions, surface acting, deep acting, and vocal tone as dependent variables.

**TABLE 3**
**Growth Curve Models for Hypotheses 2–4**

| | Null | Linear | Linear service failure/ Curvilinear service recovery | | Means for phases | | |
|---|---|---|---|---|---|---|---|
| | | | | | Phase I (A) | Phase II (B) | Phase III (C) |
| *Felt emotions* | | | | | | | |
| $-2$*LL (*df*) | 106,358.35 (2) | 105,162.70 (3) | 101,346.31 (5) | All participants | $12.49^{B,C}$ | $9.28^{A,C}$ | $10.82^{A,B}$ |
| $\Delta -2*LL_{null}$ (*df*) | | 1,195.65 (1)*** | 5,012.04 (3)*** | Service failure | $12.53^{B,C}$ | $9.28^{A,C}$ | $7.87^{A,B}$ |
| $\Delta -2*LL_{linear}$ (*df*) | | | 3,816.39 (2)*** | Service recovery | $12.45^{B}$ | $9.28^{A,C}$ | $13.04^{B}$ |
| *Surface acting* | | | | | | | |
| $-2$*LL (*df*) | 119,889.15 (2) | 115,413.91 (3) | 111,987.74 (5) | All participants | $4.82^{B,C}$ | $10.28^{A}$ | $10.17^{A}$ |
| $\Delta -2*LL_{null}$ (*df*) | | 4,475.24 (1)*** | 7,901.41 (3)*** | Service failure | $4.68^{B,C}$ | $10.13^{A,C}$ | $13.60^{A,B}$ |
| $\Delta -2*LL_{linear}$ (*df*) | | | 3,426.17 (2)*** | Service recovery | $4.95^{B,C}$ | $10.42^{A,C}$ | $7.59^{A,B}$ |
| *Deep acting* | | | | | | | |
| $-2$*LL (*df*) | 116,240.13 (2) | 113,429.39 (3) | 110,381.79 (5) | All participants | $5.27^{B,C}$ | $9.85^{A}$ | $8.93^{A}$ |
| $\Delta -2*LL_{null}$ (*df*) | | 2,810.74 (1)*** | 5,858.34 (3)*** | Service failure | $5.33^{B,C}$ | $8.91^{A,C}$ | $11.36^{A,B}$ |
| $\Delta -2*LL_{linear}$ (*df*) | | | 3,047.60 (2)*** | Service recovery | $5.20^{B,C}$ | $10.75^{A,C}$ | $7.11^{A,B}$ |
| *Vocal tone* | | | | | | | |
| $-2$*LL (*df*) | 45,811.74 (2) | 39,258.11 (3) | 35,221.51 (5) | All participants | $11.22^{B,C}$ | $10.27^{A}$ | $10.23^{A}$ |
| $\Delta -2*LL_{null}$ (*df*) | | 6,553.63 (1)*** | 10,590.23 (3)*** | Service failure | $11.29^{B,C}$ | $10.29^{A,C}$ | $9.73^{A,B}$ |
| $\Delta -2*LL_{linear}$ (*df*) | | | 4,036.60 (2)*** | Service recovery | $11.15^{B,C}$ | $10.25^{A,C}$ | $10.61^{A,B}$ |

*Note:* LL = log likelihood. Linear effects were coded (Phase I = −1; Phase II = 0; Phase III = 1). A combined linear/curvilinear effect was created by multiplying the linear effect code (–1, 0, 1) by a condition code that distinguished Phase III for the service failure (1, 1, 1) and service recovery (1, 1, −1) conditions. This produced a linear trend for the service failure condition (–1, 0, 1) and a curvilinear trend for the service recovery condition (–1, 0, −1). Values with superscripts indicate that the mean for a particular phase within condition was significantly different from that of another phase.

*** $p < .001$

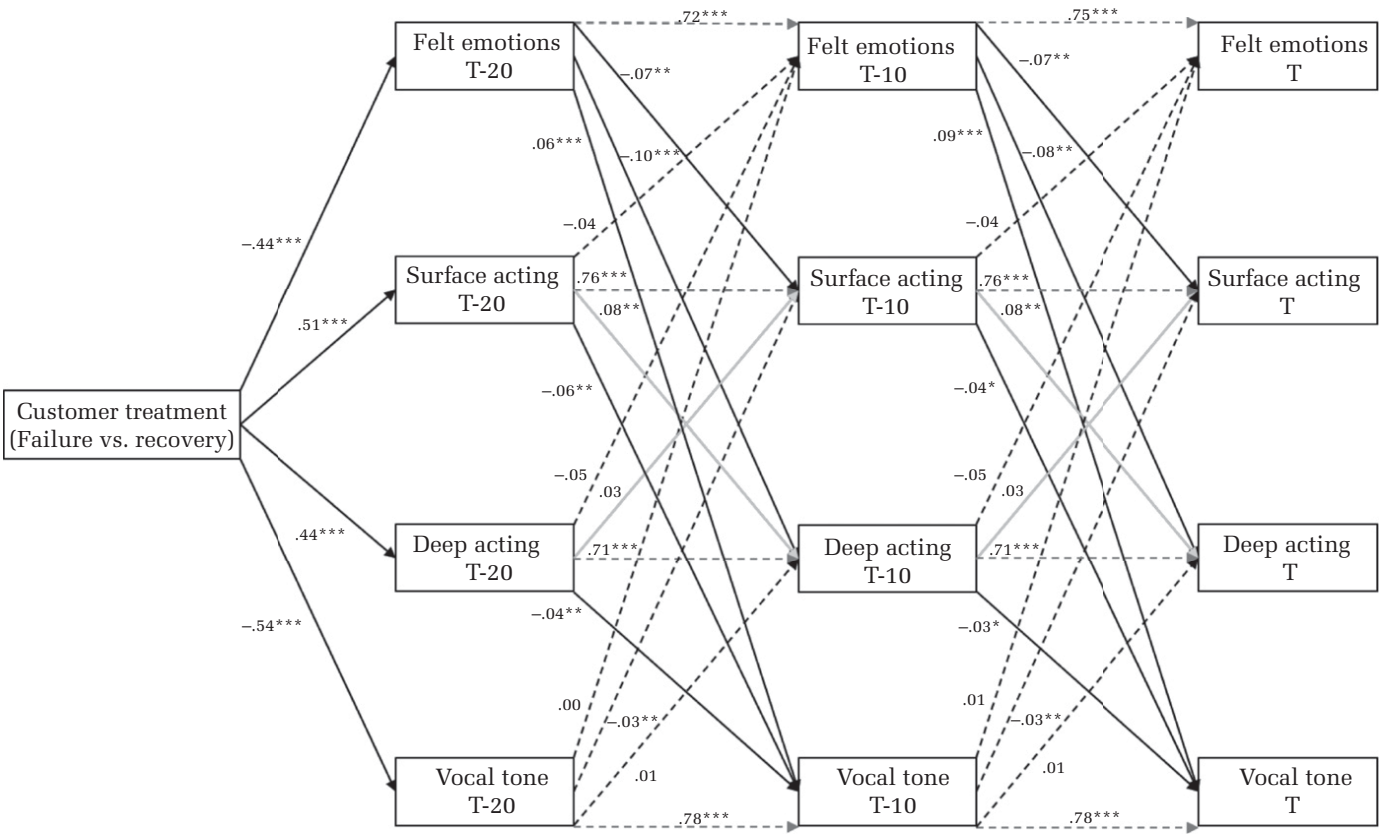
a discrepancy for the service failure condition, and the emergence and reduction of a discrepancy for the service recovery condition.

## Dynamics of Felt Emotions, Emotion Regulation, and Vocal Tone (Hypotheses 5–9)

After establishing that customer behavior affects emotional labor dynamics, we focused on the interrelationships among the continuously measured variables both in the moment (that is, concurrent) and across our lagged occasions. We first modeled the customer influence growth factor corresponding to the mixed linear and curvilinear effects as an exogenous predictor of each variable in the model (see Figure 2). Having established that the continuously measured variables exhibit high levels of stability at lags of 10 seconds (see Table 2), we modeled autocorrelations when testing all substantive relationships. We tested the relationships among the four variables within measurement occasions, as well as across lagged occasions, in an attempt to provide a strong test of the theoretical propositions assumed by our self-regulatory model of emotional labor. Testing lagged relationships in the presence of autocorrelations enabled us to better infer the causal order among variables across the three time periods (T, T-10, T-20). In addition, we modeled lagged effects that would provide insight into the potential for reverse causality among variables, as well as the causal direction in the relationship between surface acting and deep acting. The full results are presented in Figure 2.

Hypothesis 5 predicted a negative relationship between felt emotions and emotion regulation, such that, as felt emotions became more negative, (a) surface acting and (b) deep acting would increase (that is, a negative relationship). The concurrent within-person relationships in the full model were significant and negative (for example, at T: felt emotions with surface acting, $\gamma = -.37$, $p < .001$; felt emotions with deep acting, $\gamma = -.31$, $p < .001$), supporting the idea that momentary felt emotions were more negative (less positive) when momentary surface acting or deep acting were high compared to when they were low. In testing lagged effects, these negative within-person relationships held for surface acting (for example, at T, emotions at T-10: $\gamma = -.07$, $p < .01$) and deep acting (for example, at T, emotions at

**FIGURE 2**
**Path Analytic Model**




*Note*: Customer treatment represents the service failure (linear: −1, 0, 1) and recovery (curvilinear: −1, 0, −1) growth factors. Solid black lines represent causal paths. Dashed gray lines represent the stability model. Dashed black lines represent the reverse causal paths. To provide the most stringent test, all possible paths were modeled among the variables; these paths are in light gray. Concurrent relations were also modeled, but are not shown. The concurrent relations were:

felt emotions/surface acting: at T, $\gamma = -.37$; at T-10, $\gamma = -.37$; at T-20, $\gamma = -.50$;
felt emotions/deep acting: at T, $\gamma = -.31$; at T-10, $\gamma = -.31$; at T-20, $\gamma = -.47$;
felt emotions/vocal tone: at T, $\gamma = .15$; at T-10, $\gamma = .15$; at T-20, $\gamma = .23$;
surface acting/deep acting: at T, $\gamma = .37$; at T-10, $\gamma = .37$; at T-20, $\gamma = .57$;
surface acting/vocal tone: at T, $\gamma = -.13$; at T-10, $\gamma = -.13$; at T-20, $\gamma = -.31$;
deep acting/vocal tone: at T, $\gamma = -.10$; at T-10, $\gamma = -.11$; at T-20, $\gamma = -.25$.
All concurrent relations were significant at $p < .001$ except for deep acting and vocal tone at T ($p < .01$).

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

T-10: $\gamma = -.08$, $p < .01$), suggesting that felt emotions trigged within-person changes in emotion regulation. We also considered the reverse causal order, testing whether surface and deep acting predicted within-person changes in felt emotions. Across time lags, these effects were nonsignificant (see Figure 2). Thus momentary felt emotions were primarily an antecedent, as opposed to an outcome, of momentary regulation over the course of a customer service episode.

Hypothesis 6 predicted that momentary felt emotions would have a positive within-person relationship with momentary vocal tone. This hypothesis was supported at concurrent assessments (for example, at T: $\gamma = .15$, $p < .001$) and when modeling lagged felt emotions as a predictor of within-person changes in vocal tone (for example, vocal tone at T, felt emotions at T-10: $\gamma = .09$, $p < .001$). We also tested whether momentary vocal tone exhibited causal precedence over felt emotions, but found no support for this causal

direction (see Figure 2). As such, changes in vocal tone were shaped by felt emotions within a customer interaction, but changes in felt emotions were not the result of the affective tone in participants' voices.

For Hypotheses 7 and 8, we examined the within-person relations of vocal tone with surface acting and deep acting, predicting that surface acting would be negatively related to vocal tone (Hypothesis 7) and that deep acting would be positively related to vocal tone (Hypothesis 8). Concurrently, the within-person relationships of both strategies with vocal tone were negative, such that more momentary emotion regulation was associated with worse momentary vocal tone (for example, at T: surface acting with vocal tone, $r = -.13$, $p < .001$; deep acting with vocal tone, $r = -.10$, $p < .01$). Modeling lagged emotion regulation as a within-person predictor of changes in vocal tone revealed a similar pattern, with momentary surface acting being negatively related to subsequent vocal tone (for example, vocal tone at T, surface acting at T-10: $\gamma = -.04$, $p < .05$), as was momentary deep acting (for example, vocal tone at T, deep acting at T-10: $\gamma = -.03$, $p < .05$). These results support Hypothesis 7, but do not support Hypothesis 8. Thus, although past work has generally found that high deep acting is associated with better affective delivery (Hülsheger & Schewe, 2011), at the momentary level of analysis greater deep acting was associated with worse vocal tone (controlling for surface acting and felt emotions). We also considered whether vocal tone might exhibit causal precedence over both forms of emotion regulation. Interestingly, vocal tone did exhibit causal precedence over surface acting (for example, surface acting at T, vocal tone at T-10: $\gamma = -.03$, $p < .01$), but not deep acting, suggesting that when participants sounded more negative, they were more likely to subsequently report that they had been trying to hide their negative feelings. Thus surface acting and vocal tone appeared to influence each other reciprocally, whereas deep acting preceded only vocal tone (in a negative manner).

Supporting Hypothesis 9, we found a positive within-person relation between momentary surface and deep acting (for example, at T: $\gamma = .37$, $p < .001$; see Figure 2), supporting the idea that the strategies are utilized simultaneously. Supplemental analyses revealed that, at any moment, participants reported using both strategies to some extent during about 72.7% of the call, providing evidence that surface and deep acting can occur at the same point in time. Further, we examined whether there was causal precedence between the two strategies. In our cross-lagged panel analysis, deep acting was consistently and positively predicted by lagged surface acting (for example, deep acting at T, surface acting at T-10: $\gamma = .08$, $p < .01$); surface acting was not predicted by lagged deep acting (see Figure 2). Thus surface acting appears to precipitate deep acting, but not vice versa. These results suggest that surface acting may be an immediate response in difficult encounters, triggering downstream deep acting (see, e.g., Beal & Trougakos, 2013).

In sum, our hypotheses suggest a specific causal sequence similar to that theorized by Grandey (2000), in which felt emotions impact on affective performance (that is, vocal tone) through emotion regulation. Further, we estimated the indirect effect of felt emotions (T-20) to vocal tone (T), which was significant (estimate = .112, $p < .001$). Contained within this estimate are four specific indirect effects operating through the following variables assessed at the lag of T-10: felt emotions (estimate = .043, $p < .001$); vocal tone (estimate = .062, $p < .001$); surface acting (estimate = .004, $p = .076$); and deep acting (estimate = .003, $p = .072$). Although the indirect effects of emotions at T-20 on vocal tone at T through surface acting and deep acting at T-10 are not conventionally significant ($p < .05$), these effects are in the presence of the much stronger relationship of felt emotions with subsequent vocal tone (with either felt emotions or vocal tone at T-10 as the mediator) and each other, and became significant when each was modeled by itself.

## DISCUSSION

By utilizing continuous rating methodology in a call center simulation, we addressed the intricacies surrounding the self-regulatory view of emotional labor (Diefendorff & Gosserand, 2003; Diefendorff & Richard, 2008). Our results illustrated that felt emotions, emotion regulation, and vocal tone vary substantially within a single interaction and are impacted on by customer behavior. By modeling lagged values of all continuously assessed variables (similarly to a cross-lagged panel design), we were able to examine causal relations among the emotional labor variables, allowing us to gain deeper insight into the ways in which these variables relate to each other within individuals over the course of a dynamic customer service interaction.

### Theoretical Implications

The current findings suggest that more careful consideration of temporal issues is needed in

emotional labor theory and research. For example, Grandey's (2000) model of emotional labor is one of the most prominent, but has been interpreted as representing a static, causal sequence in which external factors (that is, display rules, affective events) impact on internal emotion regulation, which influences employee well-being and performance. Indeed, the model does not explicitly address the within-person or within-episode variability that is known to characterize emotions (Mauss et al., 2005). The current study is the first to provide empirical evidence of moment-to-moment, *within-episode* changes in emotional labor variables, capturing the ebb and flow of felt emotions, emotion regulation, and vocal tone during a customer service interaction. As such, our results contribute to theory and research on emotional labor in several ways.

First, our results suggest that employees may actively search their social environments for information on whether the customer is satisfied and that within-interaction variability affects emotional labor processes. This idea was highlighted by the fact that manipulated customer civility had substantial effects on changes in felt emotions, emotion regulation, and vocal tone, and that these continuously measured variables then influenced each other in ways that were largely consistent with theory. Indeed, it seems that individuals may have an overarching standard for "keeping the customer happy" and that they are very sensitive to discrepancies from this standard (Diefendorff & Gosserand, 2003; Diefendorff & Richard, 2008). Further, participants reported feeling worse when customers were not treating them well, perhaps because this signaled that they were not providing quality service, as well as the fact that they were being treated in an uncivil manner. This change in participant affect resulted in increased emotion regulation via surface *and* deep acting, as well as worse vocal tone during the call. Thus, even though the effects of customer behavior have been described in emotional labor models (Côté, 2005; Diefendorff & Gosserand, 2003) and highlighted empirically (Goldberg & Grandey, 2007; McCance et al., 2013; Rupp & Spencer, 2006), this is the first evidence that within-episode changes in customer behavior can causally affect the trajectories of employee emotional labor processes (see, e.g., Rafaeli & Sutton, 1987). Interestingly, our results suggest that, even after being mistreated, positive changes in customer demeanor can return employee affective states and regulation to pre-incivility levels, suggesting that a rebound effect can occur when a dissatisfied customer becomes pleased. Thus training employees on methods of service recovery may not only benefit customers directly, but also benefit employees indirectly by altering how customers treat them (Côté, 2005).

Second, we found that our continuously measured variables exhibited a causal order largely consistent with self-regulatory theory (see, e.g., Diefendorff & Gosserand, 2003; Diefendorff & Richard, 2008)—that is, felt emotions predicted changes in emotion regulation, which in turn predicted changes in the quality of participant vocal tone. Changes in felt emotions also predicted changes in vocal tone, and there were small indirect effects for felt emotions through emotion regulation in predicting vocal tone. This pattern of causal effects was observed over a 10-second lag period despite each of the variables exhibiting a high level of stability. Further, in only one instance was reverse causality consistently observed: Vocal tone exhibited a causal relationship with changes in surface acting (and surface acting predicted changes in vocal tone), suggesting that how participants sounded resulted in changes in the extent to which they suppressed felt emotions. This idea suggests that suppression may be the strategy of choice when individuals detect that they do not sound as positive as they should (see, e.g., Beal & Trougakos, 2013). In addition, the sign of all of the proposed effects was in the expected direction, except for deep acting to vocal tone. Counter to expectations, deep acting *negatively* predicted vocal tone, suggesting that, on a moment-to-moment basis, deep acting may not be as effective as has been previously found at the event and person levels (see, e.g., Beal et al., 2006; Hülsheger & Schewe, 2011).

Our third contribution is that we explored in detail the interrelatedness of surface and deep acting. The two strategies were highly positively correlated at a given moment, fitting with recent theorizing about the two regulation techniques (Beal & Trougakos, 2013), as well as meta-analytic (Hülsheger & Schewe, 2011) and within-person (Judge et al., 2009) results showing positive relationships. As such, our findings suggest that surface acting and deep acting should not be considered in an "either/or" manner, but as complementary strategies that individuals may simultaneously pursue to manage difficult situations (see, e.g., Gabriel et al., 2015)—that is, when faced with an uncivil customer and negative emotions, individuals simultaneously suppress their negative feelings and actively try to change those negative feelings to be more positive. Interestingly, we found that surface acting predicts changes in deep acting, but not the reverse, suggesting that changes in surface acting may precipitate changes in deep acting. This finding raises questions about which regulation strategy comes first in the emotion

management process and suggests that, when faced with an unpredictably rude customer, it may be that suppressing negative feelings occurs first, followed by more deliberate attempts to change those feelings. These findings might suggest ways in which to extend emotion regulation theories (Gross, 1998) to more dynamic emotional contexts that unfold over time.

Fourth, although we observed within-person variability in all continuously measured variables, vocal tone exhibited much less variability (that is, the observable external behavior in call center work) than the internal constructs of felt emotions and emotion regulation. This finding suggests that, from the perspectives of two independent listeners, vocal tone was fairly steady and consistent during the interactions (ranging about 2 points on the 20-point scale), despite large swings in felt emotions (fluctuating some 5–7 points), emotion regulation (fluctuating about 10 points), and manipulations of customer incivility. These results are consistent with the self-regulatory idea (Diefendorff & Gosserand, 2003) that the purpose of emotional labor is to keep emotional behavior (that is, vocal tone) steady despite turmoil in the environment (that is, incivility) and wide swings in the internal states (that is, felt emotions). Thus, even though feelings changed and a good deal of effort was needed to regulate emotions, participants maintained a fairly steady outward display that was consistent with the overarching emotional display rule.

Finally, a broader implication is that our findings suggest that emotion regulation theory (Gross, 1998) may benefit from considering more fully the variability and time course of emotion regulation strategies. Past experimental work (Butler et al., 2003; Richards & Gross, 2000) often manipulates which strategy is used by instructing individuals to regulate their emotions using one strategy or the other (for example suppress, reappraise) during an event (such as watching a film clip), and then measures the effects of this manipulation on emotional states, expressions, and physiology. This approach implicitly assumes that the manipulated regulation strategy is used consistently over the event. Based on our findings, there is reason to believe that the level of regulation may systematically vary over emotional episodes, and that this variability may be tied to within-episode changes in the environment and the person. A fruitful avenue for future work might be to examine the effects of "micro-events" (for example particular words, images, stimuli) within a time period on affect and regulation trajectories. Further, our results provide a first look at the stability of these regulation strategies over time periods ranging from 1 second to 20 seconds. More work is needed to better understand when and why regulation strategies exhibit change or stability, but these initial data provide a starting point for exploring these ideas. Another avenue to consider is the role of individual differences in predicting stability, change, and trajectories of emotions, emotion regulation, and vocally expressive behavior.

### Practical Implications

Our results shed additional light on the problems associated with incivility in call centers specifically, and service occupations in general. Consistent with past work (Goldberg & Grandey, 2007; Rupp & Spencer, 2006), exposure to incivility, even briefly, can increase felt negative emotions and the need to regulate, both of which are harmful for employees and the emotional performance provided. However, the present results extend past findings by showing that brief exposure to positive treatment from customers after incivility (that is, service recovery) produced a rebound in positive affect and a decreased need to regulate emotions. In negative social exchanges, these results suggest that striving for a positive resolution may be beneficial not only for customers, but also for employee well-being. Research on apologies and forgiveness in interpersonal relationships suggests that being forgiven can have ameliorative effects (Takaku, 2001). Perhaps an analogous process occurs in dynamic service interactions in which efforts to attain service recovery are successful.

In the absence of producing a positive conclusion to a service interaction, there may be other avenues for employees to experience a restoration of their baseline positive affect. For instance, providing respite breaks between interactions (see, e.g., Trougakos, Beal, Green, & Weiss, 2008) may help employees to recover resources lost. If respite breaks are not possible, another option would be to create periods of time during which employees can capitalize on social support by processing difficult encounters with coworkers (see, e.g., McCance et al., 2013). Additionally, Beal and Trougakos (2013) highlighted that strong negative emotional experiences, such as those stemming from incivility exposure, could promote negative coping responses such as rumination, which can create a vicious cycle of negative affect. As such, managers may want to train employees on how to detach from negative experiences as they occur or immediately afterward.

In civil exchanges, employees may be encouraged to reflect on the positive aspects of the exchange to foster psychological resources (see, e.g., Bono, Glomb, Shen, Kim, & Koch, 2013) as a buffer against future incivility. Encouraging individuals to focus on good things can lead to increased positive emotions (Seligman, Steen, Park, & Peterson, 2005), and having employees focus on successful interactions could increase positive emotional states.

### Limitations and Future Directions

As with any study, there are limitations. First, we exposed participants to only one call, which is incongruent with call center work. However, we were focused on providing a micro-level account of basic processes. As such, only one call was needed to capture this variability. Moreover, given that participants listened back to the call three times to provide their continuous ratings, we sought to reduce the fatigue experienced by participants from the procedure. Thus our results may actually slightly underestimate the negative effects on workers. Nonetheless, future work should consider multiple calls. By doing so, one could detect whether (a) emotion regulation and felt emotions from one call carry over to the next call, and (b) trajectories of emotional labor change across calls. As evidenced in Figure 1a, during Phase III in the service failure condition, ratings of surface acting rose slightly higher than deep acting. In as much as this trend is salient, participants may rely on surface acting more in subsequent calls.

Additionally, although our continuous rating approach was advantageous for capturing momentary assessments, there is the possibility for slight retrospection bias given that participants listened back to the call to provide their ratings. As mentioned earlier, past research has tested whether continuous ratings differ if provided "online" (that is, simultaneously with stimulus presentation) or immediately after stimulus presentation, finding no differences (Mauss et al., 2005). Moreover, given that participants were required to engage actively in our call center simulation, we deemed online continuous ratings to be difficult, if not entirely impossible. Although we do not see an alternative to conducting continuous ratings in such a simulation, we acknowledge that this bias may have slightly affected our results. For instance, as can be seen in Figure 1, participants did pre-emptively begin to shift their ratings in advance of the different phase manipulations. This could be attributed to participants recalling being treated well or treated poorly by the caller with whom they interacted slightly ahead of the point at which the instance actually occurred in the recording. We do not suspect that such movement in ratings affected our results, but nonetheless recognize this as a cautionary note to our findings.

Related to our analytic choices, we used a 10-second lag to try to tease apart the causal direction of our effects. Although we chose this lag in an attempt to balance the consideration of the stability of effects (that is, autocorrelations) with the potential time frame within which one variable is likely to impact on the other, we acknowledge that this choice was somewhat arbitrary. However, our supplemental analyses revealed similar patterns of causal effects across a variety of lag amounts, helping to alleviate concerns that these particular results are a product of our specific lag choice. Nonetheless, we encourage future research to examine lag amounts in modeling causal effects more fully to determine whether and in what ways different lag choices might impact on the observed effects.

In regards to vocal tone, another way in which to assess employee emotional performance in call center exchanges would be through assessing emotional authenticity, which is an equally important aspect of affective delivery in service exchanges (see, e.g., Grandey, Fisk, Mattila, Jansen, & Sideman, 2005; Hennig-Thurau, Groth, Paul, & Gremler, 2006). As such, future research could replicate our findings with emotional authenticity rated by third-party observers, as opposed to only vocal tone. In doing so, researchers could identify whether felt emotions, surface acting, and deep acting affect the authenticity of momentary displays. Other continuous ratings that would be beneficial for future work to consider would be parsing surface acting into faking and suppression, given that both aspects are common in emotional labor measures (Diefendorff et al., 2005), as well as continuous ratings of employee display rule perceptions to test whether such perceptions vary throughout a customer interaction (see, e.g., Diefendorff & Richard, 2008). Moreover, another extension for future research would be to consider dispositional traits that affect the different trajectories of emotional labor processes, such as customer orientation, dispositional affectivity, or emotion demands–abilities fit (for a recent consideration of emotional labor antecedents, see Gabriel et al., 2015).

One may question the realism and the extent to which participants were committed to playing the role of call center worker. In an attempt to address

this concern proactively, we sought to enhance the realism of the task by means of the instructions given to participants, the layout and signage of the room, and the use of an established cover story. We also sought to enhance participant motivation for working on the task by letting participants know that their performance could impact on the customer's chances of winning $100 and their own chances of winning a gift card. However, we acknowledge that there may be concerns with realism and recommend that future continuous rating research using actual call center workers be conducted. For instance, a viable approach would be to collect recordings of actual customer calls, then have real employees provide ratings postshift of their emotional experience and emotion regulation. Such ratings could then be paired with actual customer and supervisor affective evaluations.

A natural extension of our findings is to consider whether there is a point at which customer incivility sparks the abandonment of display rules and the goal of satisfying customers. In the current study, our participants were treated rudely, but we did not let the calls escalate to extreme amounts (for example there was no cursing), in order to simulate typical types of incivility that customer service employees experience (see, e.g., Sliter, Jex, Wolford, & McInnerney, 2010). However, it may be the case that customer incivility can escalate—potentially changing from incivility to aggression—to the point at which employees can no longer maintain "service with a smile," or other required emotional displays (such as a bill collector's ability to display anger when the customer being called is crying or distressed). These ideas were suggested by Beal and Trougakos (2013), and we encourage future research to consider whether such a threshold exists at which employees begin to deviate from emotional display rules.

## CONCLUSION

The current study moved emotional labor into new territory, studying whether, and in what ways, emotional labor dynamically varied in a single interaction. By providing a first test of these processes, we recommend that future research consider the within-episode variability of felt emotions, emotion regulation, and vocal tone, as well as micro-event triggers that impact on this variability. We encourage emotional labor scholars to place more focus on moment-to-moment shifts in emotional labor, aligning our dynamic theories with dynamic empirical results.

## REFERENCES


Andersson, L. M., & Pearson, C. M. 1999. Tit for tat? The spiraling effect of incivility in the workplace. ***Academy of Management Review***, 24: 452–471.

Austin, E., Dore, T., & O'Donovan, K. 2008. Associations of personality and emotional intelligence with display rule perceptions and emotional labor. ***Personality and Individual Differences***, 44: 679–688.

Beal, D., & Trougakos, J. 2013. Episodic intrapersonal emotion regulation: Or, dealing with life as it happens. In A. A. Grandey, J. M. Diefendorff, & D. Rupp (Eds.), Emotional labor in the 21st century: Diverse perspectives on emotion regulation at work: 31–55. New York: Psychology Press/Routledge.

Beal, D. J., Trougakos, J. P., Weiss, H. M., & Dalal, R. S. 2013. Affect spin and the emotion regulation processes at work. ***Journal of Applied Psychology***, 98: 593–605.

Beal, D. J., Trougakos, J. P., Weiss, H. M., & Green, S. G. 2006. Episodic processes in emotional labor: Perceptions of affective delivery and regulation strategies. ***Journal of Applied Psychology***, 91: 1053–1065.

Bono, J., Glomb, T., Shen, W., Kim, E., & Koch, A. 2013. Building positive resources: Effects of positive events and positive reflection on work-stress and health. ***Academy of Management Journal***, 56: 1601–1627.

Bono, J. E., Foldes, H. J., Vinson, G., & Muros, J. P. 2007. Workplace emotions: the role of supervision and leadership. ***Journal of Applied Psychology***, 92: 1357–1367.

Brotheridge, C., & Grandey, A. 2002. Emotional labor and burnout: Comparing two perspectives of "people work." ***Journal of Vocational Behavior***, 60: 17–39.

Brotheridge, C., & Lee, R. 2002. Testing a conservation resources model of the dynamics of emotional labor. ***Journal of Occupational Health Psychology***, 7: 57–67.

Butler, E., Egloff, B., Wilhelm, F., Smith, N., Erickson, E., & Gross, J. 2003. The social consequences of expressive suppression. ***Emotion***, 3: 48–67.

Cacioppo, J. T., Hawkley, L. C., & Thisted, R. A. 2010. Perceived social isolation makes me sad: 5-year cross-lagged analyses of loneliness and depressive symptomatology in the Chicago Health, Aging, and Social Relations Study. ***Psychology and Aging***, 25: 453–463.

Carver, C. S., & Scheier, M. F. 1998. ***On the self-regulation of behavior***. New York: Cambridge University Press.

Cole, D. A., & Maxwell, S. E. 2003. Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. ***Journal of Abnormal Psychology***, 112: 558–577.

Côté, S. 2005. A social interaction model of the effects of emotion regulation and work strain. ***Academy of Management Review***, 30: 509–530.

Dallimore, K. S., Sparks, B. A., & Butcher, K. 2007. The influence of angry customer outbursts on service providers' facial displays and affective states. ***Journal of Service Research***, 10: 78–82.

Diefendorff, J. M., & Gosserand, R. 2003. Understanding the emotional labor process: A control theory perspective. ***Journal of Organizational Behavior***, 24: 945–959.

Diefendorff, J. M., & Richard, E. M. 2008. Not all emotional display rules are created equal: Distinguishing between prescriptive and contextual display rules. In. N. M. Ashkanasy & C. L. Cooper (Eds.), Research companion to emotion in organizations: 316–334. Northampton, CT: Edward Elgar.

Diefendorff, J. M., Croyle, M. H., & Gosserand, R. H. 2005. The dimensionality and antecedents of emotional labor strategies. ***Journal of Vocational Behavior***, 66: 339–357.

Ebner-Priemer, U. W., & Trull, T. J. 2011. Investigating temporal instability in psychological variables: Understanding the real world as time dependent. In M. R. Mehl & T. S. Conner (Eds.), ***Handbook of research methods for studying daily life***: 423–458. New York: Guilford Press.

Ekman, P., & Friesen, W. V. 1969. Nonverbal leakage and clues to deception. ***Psychiatry***, 32: 88–106.

Ekman, P., Friesen, W. V., & O'Sullivan, M. 1988. Smiles when lying. ***Journal of Personality and Social Psychology***, 54: 414–420.

Elfenbein, H. A. 2007. Emotion in organizations: A review and theoretical integration. ***The Academy of Management Annals***, 1: 315–386.

Elfenbein, H. A., & Ambady, N. 2002. Predicting workplace outcomes from the ability to eavesdrop on feelings. ***Journal of Applied Psychology***, 87: 963–971.

Fredrickson, B. L., & Kahneman, D. 1993. Duration neglect in retrospective evaluations of affective episodes. ***Journal of Personality and Social Psychology***, 65: 45–55.

Frijda, N. H. 1993. The place of appraisal in emotion. ***Cognition and Emotion***, 7: 357–387.

Gabriel, A. S., Daniels, M. A., Diefendorff, J. M., & Greguras, G. J. 2015. Emotional labor actors: A latent profile analysis of emotional labor strategies. ***Journal of Applied Psychology***, 100: 863–879.

Goldberg, L. S., & Grandey, A. A. 2007. Display rules versus display autonomy: Emotion regulation, emotional exhaustion, and task performance in a call center simulation. ***Journal of Occupational Health Psychology***, 12: 301–318.

Gottman, J. M., & Levenson, R. W. 1985. A valid procedure for obtaining self-report of affect in marital interaction. ***Journal of Consulting and Clinical Psychology***, 53: 151–160.

Grandey, A. A. 2000. Emotion regulation in the workplace: A new way to conceptualize emotional labor. ***Journal of Occupational Health Psychology***, 5: 95–110.

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: 86–96.

Grandey, A. A., Diefendorff, J. M., & Rupp, D. 2013. Emotional labor: Overview of definitions, theories, and evidence. In A. A. Grandey, J. M. Diefendorff, & D. Rupp (Eds.), Emotional labor in the 21st century: Diverse perspectives on emotion regulation at work: 3–27. New York: Psychology Press/Routledge.

Grandey, A. A., Fisk, G. M., Mattila, A. S., Jansen, K. J., & Sideman, L. A. 2005. Is "service with a smile" enough? Authenticity of positive displays during service encounters. ***Organizational Behavior and Human Decision Processes***, 96: 38–55.

Granger, C. W. J. 1969. Investigating causal relations by econometric models and cross-spectral methods. ***Econometrica***, 37: 424–438.

Gross, J. J. 1998. The emerging field of emotion regulation: An integrative review. ***Review of General Psychology***, 2: 271–299.

Groth, M., & Grandey, A. A. 2012. From bad to worse: Negative exchange spirals in employee–customer service interactions. ***Organizational Psychology Review***, 2: 208–233.

Heaphy, E. D., & Dutton, J. E. 2008. Positive social interactions and the human body at work: Linking organizations and physiology. ***Academy of Management Review***, 33: 137–162.

Hennig-Thurau, T., Groth, M., Paul, M., & Gremler, D. D. 2006. Are all smiles created equal? How emotional contagion and emotional labor affect service relationships. ***Journal of Marketing***, 70: 58–73.

Hochschild, A. R. 1983. ***The managed heart: The commercialization of feeling***. Berkeley, CA: University of California Press.

Hülsheger, U. R., & Schewe, A. 2011. On the costs and benefits of emotional labor: A meta-analysis of three decades of research. ***Journal of Occupational Health Psychology***, 16: 361–389.

Hülsheger, U. R., Lang, J. W., & Maier, G. W. 2010. Emotional labor, strain, and performance: Testing reciprocal relationships in a longitudinal panel study. ***Journal of Occupational Health Psychology***, 15: 505–521.

Judge, T., Woolf, E., & Hurst, C. 2009. Is emotional labor more difficult for some than for others? A multilevel, experience-sampling study. ***Personnel Psychology***, 62: 57–88.

Kammeyer-Mueller, J. D., Rubenstein, A. L., Long, D. M., Odio, M. A., Buckman, B. R., Zhang, Y., & Halvorsen-Ganepola, M. D. K. 2013. A meta-analytic structural model of dispositional affectivity and emotional labor. ***Personnel Psychology***, 66: 47–90.

Kreft, I. G. G., & DeLeeuw, J. 1998. ***Introducing multilevel modeling***. Newbury Park, CA: Sage.

Krumhuber, E. G., Kappas, A., & Manstead, A. S. 2013. Effects of dynamic aspects of facial expressions: A review. ***Emotion Review***, 5: 41–46.

Kruml, S., & Geddes, D. 2000. Exploring the dimensions of emotional labor: The heart of Hochschild's work. ***Management Communication Quarterly***, 14: 8–49.

Levenson, R. W., & Gottman, J. M. 1983. Marital interaction: Physiological linkage and affective exchange. ***Journal of Personality and Social Psychology***, 45: 587–597.

Lykken, D. T. 1968. Statistical significance in psychological research. ***Psychological Bulletin***, 70: 151–159.

Mauss, I. B., Levenson, R., McCarter, L., Wilhelm, F., & Gross, J. 2005. The tie that binds? Coherence among emotion experience, behavior, and physiology. ***Emotion***, 5: 175–190.

Mauss, I. B., Shallcross, A. J., Troy, A. S., John, O. P., Ferrer, E., Wilhelm, F. H., & Gross, J. J. 2011. Don't hide your happiness! Positive emotion dissociation, social connectedness, and psychological functioning. ***Journal of Personality and Social Psychology***, 100: 738–748.

McCance, A. S., Nye, C. D., Wang, L., Jones, K., & Chiu, C. Y. 2013. Alleviating the burden of emotional labor: The role of social sharing. ***Journal of Management***, 39: 392–415.

Mesmer-Magnus, J. R., DeChurch, L. A., & Wax, A. 2012. Moving emotional labor beyond surface and deep acting: A disconcordance–congruence perspective. ***Organizational Psychology Review***, 2: 6–53.

Muthén, L.K., & Muthén, B.O. 1998–2013. Mplus user's guide. Los Angeles, CA: Muthén & Muthén.

Rafaeli, A., & Sutton, R. I. 1987. Expression of emotion as part of the work role. ***Academy of Management Review***, 12: 23–37.

Rafaeli, A., & Sutton, R. I. 1990. Busy stores and demanding customers: How do they affect the display of positive emotion? ***Academy of Management Journal***, 33: 623–637.

Richards, J. M., & Gross, J. J. 2000. Emotion regulation and memory: The cognitive costs of keeping one's cool. ***Journal of Personality and Social Psychology***, 79: 410–424.

Rosenthal, R., & DePaulo, B. M. 1979. Sex differences in eavesdropping on nonverbal cues. ***Journal of Personality and Social Psychology***, 37: 273–285.

Ruef, A. M., & Levenson, R. W. 2007. Continuous measurement of emotion: The affect rating dial. In J. A. Coan & J. J. B. Allen (Eds.), ***Handbook of emotion elicitation and assessment***: 286–297. New York: Oxford University Press.

Rupp, D., & Spencer, S. 2006. When customers lash out: The effects of customer interactional justice on emotional labor and the mediating role of discrete emotions. ***Journal of Applied Psychology***, 91: 971–978.

Scott, B., & Barnes, C. 2011. A multilevel field investigation of emotional labor, affect, work withdrawal, and gender. ***Academy of Management Journal***, 54: 116–136.

Scott, B., Barnes, C., & Wagner, D. 2012. Chameleonic or consistent? A multilevel investigation of emotional labor variability and self-monitoring. ***Academy of Management Journal***, 55: 905–926.

Seligman, M. E. P., Steen, T. A., Park, N., & Peterson, C. 2005. Positive psychology progress: Empirical validation of interventions. ***The American Psychologist***, 60: 410–421.

Sliter, M., Jex, S., Wolford, K., & McInnerney, J. 2010. How rude! Emotional labor as a mediator between customer incivility and employee outcomes. ***Journal of Occupational Health Psychology***, 15: 468–481.

Strack, F., Martin, L. L., & Stepper, S. 1988. Inhibiting and facilitating conditions of the human smile: A nonobtrusive test of the facial feedback hypothesis. ***Journal of Personality and Social Psychology***, 54: 768–777.

Takaku, S. 2001. The effects of apology and perspective taking on interpersonal forgiveness: A dissonance-attribution model of interpersonal forgiveness. ***The Journal of Social Psychology***, 141: 494–508.

Totterdell, P., & Holman, D. 2003. Emotion regulation in customer service roles: Testing a model of emotional labor. ***Journal of Occupational Health Psychology***, 8: 55–73.

Trougakos, J. P., Beal, D. J., Green, S. G., & Weiss, H. M. 2008. Making the break count: An episodic examination of recovery activities, emotional experiences, and positive affective displays. ***Academy of Management Journal***, 51: 131–146.

Tsai, W.-C., & Huang, Y. M. 2002. Mechanisms linking affective delivery and customer behavioral intentions. ***Journal of Applied Psychology***, 87: 1001–1008.

Van Jaarsveld, D., & Poster, W. R. 2013. Call centers: Emotional labor over the phone. In A. A. Grandey, J. M. Diefendorff, & D. Rupp (Eds.), Emotional labor in the 21st century: Diverse perspectives on emotion regulation at work: 153–173. New York: Psychology Press/Routledge.

Wagner, D. T., Barnes, C. M., & Scott, B. A. 2014. Driving it home: How workplace emotional labor harms employee home life. ***Personnel Psychology***, 67: 487–516.

Wegge, J., Vogt, J., & Wecking, C. 2007. Customer induced stress in call centre work: A comparison of

audio and videoconference. ***Journal of Occupational and Organizational Psychology***, 80: 693–712.

Wharton, A. S., & Erickson, R. J. 1993. Managing emotions on the job and at home: Understanding the consequences of multiple emotional roles. ***Academy of Management Review***, 18: 457–486.

Yang, J., & Diefendorff, J. M. 2009. The relations of daily counterproductive workplace behavior with emotions, situational antecedents, and personality moderators: A diary study in Hong Kong. ***Personnel Psychology***, 62: 259–295.

Zapf, D. 2002. Emotion work and psychological well-being: A review of the literature and some conceptual considerations. ***Human Resource Management Review***, 12: 237–268.

Zuckerman, M., Amidon, M. D., Bishop, S. E., & Pomerantz, S. D. 1982. Face and tone of voice in the communication of deception. ***Journal of Personality and Social Psychology***, 43: 347–357.

**Allison S. Gabriel** (asgabriel@email.arizona.edu) received her PhD in industrial-organizational psychology from the University of Akron. She is an assistant professor in the Department of Management and Organizations at the University of Arizona. Her research focuses on emotions at work, motivation, job demands, and employee well-being.

**James M. Diefendorff** (jdiefen@uakron.edu) received his PhD in industrial-organizational psychology from the University of Akron. He is a professor in the Department of Psychology at the University of Akron. His research focuses on emotional labor and work motivation.