Episodic Processes in Emotional Labor: Perceptions of Affective Delivery and Regulation Strategies

Daniel J. Beal
Rice University

John P. Trougakos, Howard M. Weiss, and Stephen G. Green
Purdue University

This study examined emotional labor processes from a within-person, episodic framework. The authors hypothesized that the influence of negative emotions on affective delivery would be lessened by regulation strategies for supervisor perceptions but not self-perceptions. In addition, difficulty maintaining display rules was hypothesized to mediate the relation between negative emotions and self-perceptions of affective delivery. Finally, the influence of surface acting strategies on these processes as well as correlations with individual differences was investigated. Hypotheses were tested using ecological momentary assessment of a sample of cheerleading instructors. Results suggest that surface actors can regulate emotions effectively on an episode-to-episode basis but find the episode more difficult. In addition, surface actors exhibit more general tendencies to devalue themselves and experience fewer positive emotions.

Keywords: emotion regulation, emotional labor, affective delivery, episodic processes, negative emotion

Emotional labor is a term used in the organizational literature to describe a family of constructs involving the regulation of emotions in work settings (Hochschild, 1983; Morris & Feldman, 1996). Whether the term refers to instances of emotion regulation, tasks that require high levels of emotion regulation, or jobs for which emotion regulation is necessary and frequent, the underlying theme is that the appropriate display of emotion is an integral part of successful performance (Grandey, 2000; Pugh, 2001; Totterdell & Holman, 2003). Certainly, jobs emphasizing emotional labor can have other important and more traditional performance requirements, but, generally, the *emotional labor* term is reserved for the managing of outward expressive displays that are desired by the organization. Frequently, the outcome of interest in jobs that involve emotional labor is referred to as *affective delivery,* or the extent to which employees maintain these expressive display requirements while at work (Grandey, 2003; Tsai & Huang, 2002). It is important to note that whether an employee maintains display rules depends on whose perception is considered; therefore, ratings of affective delivery may differ depending on who is doing the perceiving.

In the last decade or so, the importance of affective delivery for success in many jobs has increased in accordance with the shift to a service-oriented economy (Morris & Feldman, 1996; Schneider & Bowen, 1995). Research on affective delivery has found it linked to the length of time customers spend in stores, customers' positive mood, ratings of service quality, and the willingness to return to the store and recommend it to friends (Pugh, 2001; Tsai & Huang, 2002). Given the obvious importance of affective delivery for the success of organizations and the workers they employ, a detailed understanding of how employees successfully regulate their emotional expressions at work seems necessary. Unfortunately, given that affective delivery is a new concept in industrial–organizational psychology, little is known about the antecedents of emotion regulation at work. In an effort to inform this burgeoning literature, organizational researchers have connected information from more basic research on emotion regulation, particularly regulation strategies, with the literature on emotional labor and affective delivery (Grandey, 2000; Morris & Feldman, 1996; Zapf, 2002). The current study elaborates further on these connections, describing different psychological processes for different constituents involved in evaluating emotion regulation at work.

Strategies for Emotion Regulation at Work

For most people, consistent regulation of emotion often can be a very difficult task (Baumeister, Bratslavsky, Muraven, & Tice, 1998; Gross, 2002). Regulation becomes increasingly difficult when feelings are discordant with the emotions that people would prefer to experience, a state often referred to as *emotional dissonance* (Morris & Feldman, 1996). When emotional dissonance occurs at work, regulation of the emotional state often is a necessity for maintaining organizationally desired display rules. Not all jobs require the display of positive emotions (e.g., bill collectors, security personnel), but workplace emotional dissonance typically occurs when employees are experiencing negative emotions and

Daniel J. Beal, Department of Psychology, Rice University; John P. Trougakos and Stephen G. Green, Krannert School of Management, Purdue University; Howard M. Weiss, Department of Psychological Sciences, Purdue University.

Correspondence concerning this article should be addressed to Daniel J. Beal, Department of Psychology, Rice University, Mail Stop 205, 6100 Main Street, Houston, TX 77005. E-mail: dbeal@rice.edu

display rules require the expression of positive emotions. As such, the current study focuses on the experience of negative emotions in service encounters.

A key ingredient of successful emotion regulation is the particular strategy used to help present a controlled exterior. Following a dramaturgical model, organizational researchers have made a useful distinction between two strategies (Grandey, 2000, 2003; Hochschild, 1983). *Surface acting* refers to efforts to change the external expression of the emotion. People "choke down" unwanted feelings and present the expression that matches their desired display. In contrast, *deep acting* refers to efforts to change the actual experience of the emotion. This form of regulation is usually achieved either through cognitive reappraisal of the events surrounding the emotional experience or through what Gross (1998b) has called *attentional deployment*: thinking about something else to change the emotional experience to that which is desired. Recent research has found that the extent to which people engage in one strategy is largely independent of the extent to which they engage in the other. Further differentiation of the two strategies comes from evidence that they differ in their prediction of various work outcomes (Brotheridge & Grandey, 2002; Brotheridge & Lee, 2003; Grandey, 2003).

Although research on surface and deep acting emphasizes general tendencies to engage in both strategies, it is important to remember that emotion regulation occurs on an episode-to-episode basis. That is, because emotions are experienced episodically, with punctuated beginnings and limited life spans (Frijda, 1993), emotion regulation, too, must occur on an episode-to-episode basis. Therefore, if we are to better understand the process of emotional labor, we should examine it within a similarly time-bound, episodic framework. Such an examination is especially important in work contexts, where daily experiences often unfold as a series of performance episodes (Beal, Weiss, Barros, & MacDermid, 2005).

The current research therefore presents an episodic study of the emotional labor process. Specifically, we focus on the difficulty and ability to maintain high levels of affective delivery in the face of contrasting emotional states. In addition, we examine the influence of surface and deep acting strategies on emotion regulation in an episodic fashion as well as the implications for more stable aspects of the self. An important contribution of the research is the demonstration that the nature of emotional labor for the person engaging in it (e.g., the employee) is often completely different from how it is perceived by others (e.g., the employee's supervisor). As a preview, we suggest that self-ratings of affective delivery are based on a consideration of one's affective state during the episode and the difficulty that affective state creates in attempting to maintain positive displays. Supervisor ratings, however, cannot consider these issues because supervisors are not privy to the internal states of employees; instead, these ratings are based on the outward expression of positive affect. So, each source uses different information with respect to ratings of affective delivery. In the next sections, we offer further evidence for these separate processes.

Supervisor Perceptions

Maintaining display rules while experiencing negative emotions involves eliminating interfering aspects of the negative emotion either through suppression or by altering the experience of the emotion. From the supervisor's perspective, however, all that is known is how well the employee conforms to the organization's display rules. Thus, as long as those display rules are upheld, whether the employee is regulating the emotional display or changing the emotional experience is irrelevant. Of course, it is possible that one strategy is more effective than another, and this effectiveness will be detected and evaluated by supervisors. The literature on this topic, however, is mixed.

Although there is some evidence that people can be trained to detect lying (e.g., Mann, Vrij, & Bull, 2004), there is ample evidence that in most situations, people cannot detect authentic from inauthentic stories (DePaulo, 1992). At a more microlevel, Ekman's (e.g., Frank, Ekman, & Friesen, 1993) work has shown that an analysis of individual smiles can reveal differences between authentic and inauthentic smiles. Still, determinations of truthfulness and careful studies of individual smiles are not the same as noticing that one regulation strategy is more effective than another. Indeed, recent laboratory work supports the idea that surface and deep acting are equally effective across an entire episode of emotion regulation. For instance, Gross (1998a) found that the suppression of emotion was just as effective as reappraisal (a form of deep acting) in terms of concealing behavioral expressions of emotions.

Considering that most work contexts involve efforts at regulation on an episode-to-episode basis and that supervisors are not likely to make fine-grained judgments of facial expressions, we hypothesize that, for supervisors, the important element of emotion regulation is not whether it occurs through surface or deep acting. Because both surface and deep acting can be effective at any given moment in time, the factor that should determine supervisor ratings of affective delivery is the extent to which either type of emotion regulation strategy is used. So, supervisory ratings of affective delivery will be influenced by the total effort given to regulating negative emotions, regardless of the type of strategy.

> *Hypothesis 1A:* The extent to which someone engages in any form of regulation (i.e., total of surface and deep acting) will moderate the relation between the number of negative emotions experienced during an episode and supervisor-rated affective delivery for that episode, such that the negative relation will become weaker as total regulation increases.

> *Hypothesis 1B:* Surface and deep acting will exhibit the same moderating effects: As surface or deep acting increases, the relation between the number of negative emotions experienced during an episode and supervisor-rated affective delivery for that episode will become less negative.

Self-Perceptions

Outside observers may not have access to information concerning employees' internal states. The employees themselves, however, are all too aware of how they are feeling during a service encounter. Because of this awareness, the experience of negative emotion will make affective delivery seem more difficult compared with times when people are not experiencing negative emotion. In addition, the perceptions of difficulty at the hands of a

negative emotional experience will likely lead to lower self-ratings of affective delivery. These main effects of negative emotion should be present irrespective of the particular strategies used by employees. That is, regardless of the particular strategy on which people rely, experiencing negative emotions should still create perceptions of difficulty and, subsequently, lower ratings of affective delivery.

Hypothesis 2A: The number of negative emotions experienced during an episode will be positively related to self-rated difficulty in maintaining display rules for that episode.

Hypothesis 2B: The number of negative emotions experienced during an episode will be negatively related to self-rated affective delivery for that episode.

The two relations hypothesized above do not reflect isolated self-perceptions. People often use self-relevant information as the basis for other judgments of self, particularly if the two domains are very similar (Bem, 1972; Farh & Dobbins, 1989; Wells & Sweeney, 1986). Indeed, the difficulty involved in maintaining display rules seems to be the most relevant piece of information used in making self-ratings of how well positive displays were maintained during the episode (Lane & Herriot, 1990). Again, this process should occur regardless of whether the particular acting strategy is perceived as being more or less difficult. Therefore, we predict the following chain of self-perceptions: negative emotional state → perceptions of difficulty maintaining display rules → lower self-perceptions of affective delivery. Formally, we hypothesize the following:

Hypothesis 3: Difficulty maintaining display rules will mediate the relation between the number of negative emotions experienced during the episode and self-perceptions of affective delivery for that episode.

Although this mediational chain should operate regardless of the particular regulation strategy used, we realize that some forms of affect regulation are more difficult than others. First, people who engage in surface acting should perceive greater difficulty in affective delivery because surface acting is a more stressful strategy that requires substantially greater levels of self-regulation when compared with deep acting. Support for this assertion comes from research showing that suppression of emotional expression increases physiological activity, negatively influences later efforts at self-regulation, and generally is perceived as more difficult than other forms of emotion regulation (Baumeister et al., 1998; Butler et al., 2003; Gross, 2002). In terms of service encounters, as surface actors experience negative emotions, they should find it increasingly difficult to maintain display rules compared with those who do not engage as much in surface acting. As a result of this perceived difficulty, the hypothesized mediational chain predicts that high surface actors will feel that they are less able to provide high levels of affective delivery (Bandura, 1997).

Hypothesis 4A: Surface acting will moderate the relation between the number of negative emotions experienced during the episode and difficulty maintaining display rules for that

episode, such that as surface acting increases, the positive relation will become stronger.

Hypothesis 4B: Surface acting will moderate the relation between the number of negative emotions experienced during the episode and self-rated affective delivery for that episode, such that as surface acting increases, the negative relation will become stronger.

The predictions for deep actors are less clear. We do believe it likely that deep actors will perceive less difficulty in maintaining display rules compared with high surface actors. That is, when they initially experience a negative emotion, deep actors will engage in reappraisal or positive refocusing in an effort to change the experience altogether. Assuming that these efforts are sometimes successful, deep actors should not perceive as much difficulty in maintaining display rules because their actual emotional state will no longer be dissonant with what is required. Although we would expect these differences when comparing high deep actors with high surface actors, it is unclear how high deep actors will compare with those who are low in deep or surface acting. It is possible that those who engage in neither surface nor deep acting perceive no difficulty associated with negative emotional experiences and that their self-ratings of affective delivery are consequently unaffected by their emotional state. It is also possible, however, that individuals who do not regulate their emotions are completely at the mercy of these affective states and, as such, their ratings are greatly influenced by their emotional experiences. So, although we expected high deep actors and high surface actors to differ with respect to their self-ratings of difficulty and affective delivery, we took an exploratory approach in the analysis because of our uncertainty regarding those who are low in deep and surface acting.

Consequences of Emotion Regulation Strategies

Assuming that, on an episodic basis, surface and deep acting are equally likely to result in the maintenance of organizational display rules, what then is the cost for those who use the more stressful and difficult surface acting style? If, as others have suggested (e.g., Butler et al., 2003; Richards & Gross, 1999), surface acting is a more difficult strategy, then chronically engaging in this strategy should have some effect on more global aspects of one's life. Various authors have noted relations between surface acting and a variety of negative self-relevant life and work outcomes (e.g., Brotheridge & Grandey, 2002; Erickson & Wharton, 1997; Grandey, 2000; Zapf, 2002). Hochschild (1983) suggested that hiding feelings was associated with burnout and diminished self-esteem. Pugliesi (1999), although she did not examine surface and deep acting per se, did provide evidence that attempts to control the expression of emotion at work were negatively related to job satisfaction and positively related to psychological distress. Indeed, evidence suggests a pattern of negativity toward one's job and life for surface actors. In the current study, we extend this line of reasoning to several new outcomes of interest.

Positive and negative affectivity, for example, are frequently linked to other evaluative constructs such as job satisfaction and psychological distress (e.g., McWilliams, Cox, & Enns, 2003;

Thoresen, Kaplan, Barsky, Warren, & de Chermont, 2003). It may be the case, then, that surface actors experience more negative states and fewer positive states in general. In addition, the pattern of negativity that seems to be associated with surface acting undoubtedly will spread to include overall self-evaluations. This hypothesis is based on recent evidence that constructs such as neuroticism and generalized self-efficacy are part of an underlying self-evaluative construct (Judge, Erez, Bono, & Thoreson, 2003). Thus, the tendency to be in a negative state is closely tied to evaluating oneself negatively. We therefore propose the following hypotheses:

Hypothesis 5A: Surface acting will be positively related to trait negative affect and negatively related to trait positive affect.

Hypothesis 5B: Surface acting will be negatively related to core self-evaluations.

Finally, as mentioned above, the personal implications for chronic use of deep acting strategies is more ambiguous, particularly with respect to more global outcomes. Relatively few findings have been obtained for deep acting, and the relations that have been noted are somewhat inconsistent and relatively small. For example, Brotheridge and Grandey (2002) found that deep acting was positively related to feelings of personal accomplishment at work; however, Grandey (2003) found that deep acting was negatively related to job satisfaction. Given the mixed results in the literature on deep acting, we again adopted an exploratory approach and made no formal hypotheses.

Method

Overview and Sample

We examined our hypotheses using experience-sampling techniques with a sample of cheerleading instructors during summer cheerleading camps. These camps provided a work structure that was segmented into discrete episodes of customer interaction, with short breaks in between each episode. Every summer across the United States, well over 100,000 cheerleaders attend these camps. Instructors employed by a large cheerleading organization that administers many of these camps agreed to take part in the study. Typically, these camp instructors are current or former college cheerleaders who spend a substantial portion of their summer moving from camp to camp. These instructors gather to teach cheerleading skills to cheerleading teams who are high school age and younger. Each camp spans 4 days and consists of nine 2–3-hr instructional sessions with the customers, who consist of both camp attendees and team coaches.

Being an instructor at these camps is an emotionally laborious job. The display rule requirements of the job are emphasized by the organization at all times. According to company literature, instructor responsibilities include maintaining a positive attitude to help foster an environment of excitement, enthusiasm, and fun; motivating customers to learn and perform at their best; and listening to and dealing with customer (camper) concerns and complaints. These requirements suggest that cheerleading instructors are an excellent example of a high-involvement emotional labor job. As verification of these requirements, we asked our sample to complete a brief questionnaire assessing the extent to which they must hide negative emotions and show positive emotions as a part of their job (ranging from 1 [*not at all*] to 5 [*all the time*]). The means for our sample were 2.97 and 3.61, respectively, which are higher than the means for any

other job group assessed in a previous study with this scale (Brotheridge & Grandey, 2002).

In addition to the display rule requirements, instructors are expected to maintain the ideal company image (which company literature describes as "clean cut, athletic, All American, and confident") and must follow explicit rules regarding the regulation of other behaviors, as outlined by company conduct guidelines. Failure to do so may result in employees facing disciplinary action, including warnings, automatic probation, and automatic dismissal. Additionally, employees must strive to be "ideal role models"; this requires always being professional, maintaining good posture in front of campers, and maintaining a positive attitude and purpose. These examples represent just a handful of an extensive number of company policies and help to illustrate the emphasis and importance the company puts on adhering to display rules.

Participants

Twenty-three male and 46 female camp instructors from a large cheerleading organization participated in the study. Data were collected at five cheerleading training camps in the midwestern and southern United States. The average age of the instructors was 19.8 years old. Instructors had worked for the company for an average of about 2 years and had been involved in cheerleading at some level for approximately 7.5 years. Each instructor was observed for the first eight sessions of one cheerleading camp. The final (ninth) session was not used, as it consisted mostly of closing ceremonies and presentation of awards.

Supervisors

All company supervisors were former instructors. Supervisors in this study (2 men, 6 women) had an average of 7.25 years of experience (*SD* = 3.1 years) with the company and had been supervisors for an average of 5.25 years (*SD* = 3.1 years). Prior to the summer camp season, supervisors receive 3 days of head instructor training, during which they engage in extensive instruction as to how employees are to be evaluated. Supervisors then participate in a week of regional general staff training, during which they help to facilitate the training of instructors as well as continuing their own training.

The main activities for supervisors during each session involved addressing the concerns of coaches and moving from group to group ensuring that all was going well and noting the performance and affective delivery of the instructors. The average number of instructors for whom supervisors were accountable was 8.6 (*SD* = 5.2). Accordingly, the supervisory context was not constant but usually allowed for frequent monitoring of each instructor. Typically, if instructors do not perform up to par, there is a series of disciplinary actions that supervisors can pursue, including discussing the issue with the instructor, pulling the instructor off camps in the immediate future, scheduling the instructor for very few camps, or asking for the instructor to sit out a summer to reevaluate his or her priorities and needs. Thus, the affective delivery of each instructor is taken seriously, and supervisors are authorized to take actions that could result in a variety of negative consequences if affective delivery is not deemed satisfactory.

Procedure

The participants (instructors) were each given a personal digital assistant (PDA) to record their data after each camp session. The PDAs were used to gather assessments of employees' emotional states and perceptions of affective delivery throughout the course of the camps. Group training on the use of the PDA was conducted the day before each camp was to start. Prior to the training, each instructor completed a questionnaire packet designed to gather demographic information as well as individual-differences measures.

Table 1
Means, Standard Deviations, and Zero-Order Correlations for Individual-Differences Constructs

Variable	*M*	*SD*	1	2	3	4	5
1. Core self-evaluations	3.66	0.49	—				
2. Trait positive affect	5.15	0.79	.45*	—			
3. Trait negative affect	2.60	0.62	−.48*	−.11	—		
4. Surface acting	2.75	1.02	−.41*	−.31*	.19	—	
5. Deep acting	3.45	0.91	−.17	.02	.08	.31*	—

* $p < .05$.

Participants met with us immediately after each session[1] to complete the measures on their PDAs. Each assessment took approximately 3 min to complete. The camps were organized as follows: Morning sessions began at 8:30 a.m. and went to approximately 12:00 p.m. Afternoon sessions began at 2:00 p.m. and continued until 4:30 p.m., and evening sessions lasted from 6:00 p.m. to 9:00 p.m. The 1st day of camp began with an afternoon session (i.e., no morning session on the 1st day), and the 2nd and 3rd days of camp included morning, afternoon, and evening sessions, for a total of eight sessions over 3 days. At the end of the eighth session, the PDAs were collected from the instructors, a final paper questionnaire was administered, and participants were thanked for their time and effort. In addition to the instructors' data, we obtained supervisor ratings of the instructors' affective delivery at the end of each of the eight sessions.

Materials

With the exception of Deep and Surface Acting Scales, individual-differences measures were administered during the orientation session the day before each camp began. Means, standard deviations, and zero-order correlations for all individual-differences measures appear in Table 1.

Core self-evaluations. Participants' core self-evaluations were measured with the 12-item Core Self-Evaluations Scale developed by Judge et al. (2003). This scale measures the underlying self-evaluative factor that is present across the four more specific traits of self-esteem, generalized self-efficacy, neuroticism, and locus of control. A sample item from the measure is "Sometimes when I fail I feel worthless" (reverse scored). The measure is rated on a 5-point scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*) and had high internal consistency ($\alpha = .81$).

Trait emotions. Trait emotions were assessed with the personality structure of affect measure designed by Diener, Smith, and Fujita (1995). This measure asks respondents to rate the extent to which they generally feel each of 24 emotions on a 7-point scale ranging from 1 (*never*) to 7 (*always*). The measure can yield scores for prototypical dimensions of emotional experience or two broader scores for positive affect and negative affect. Scores are derived by combining all of the positive emotion words or all of the negative emotion words, respectively (see Diener et al., 1995, for details). Given that our hypotheses for trait emotion involved only positive and negative affect, we focused our examination on these scales (αs = .80 and .81, respectively).

Surface and deep acting. After the eighth session, on completing the PDA measures for the final time, participants were given a modified measure of surface and deep acting based on the Emotional Labor Scale (Brotheridge & Grandey, 2002; Brotheridge & Lee, 2003). These three-item measures tap how participants regulate emotions by hiding feelings, faking feelings, and modifying feelings to fulfill their work display rules. Modifications were made to refer specifically to the regulatory style used during the camp (i.e., the stem was changed to "On average during this camp"). These measures used a 5-point scale ranging from 1 (*not at all*) to 5 (*all the time*). An exploratory factor analysis with oblique rotation resulted in a two-factor solution for the six items (the two factors were

correlated at .31). The internal consistencies for the Surface and Deep Acting Scales were adequate but somewhat lower than those found in other samples ($\alpha = .71$ and $\alpha = .66$, respectively).

The next measures discussed were administered at the end of each of the eight camp sessions (episodic). Means, standard deviations, and zero-order correlations for all episodic measures appear in Table 2. For ease of presentation, all measures have been aggregated to the individual level.

Negative emotions. PDAs prompted instructors to select all of the emotions they had experienced during the session. The emotions listed were based on an empirically derived structure of prototypical emotional experience (Diener et al., 1995; Shaver, Schwartz, Kirson, & O'Connor, 1987) and consisted of anger, anxiety, happiness, love, sadness, shame, and a "none of the above emotions" option. Because affective delivery in this job calls for expressing high levels of cheerfulness and enthusiasm, we examined situations in which instructors would be likely to experience emotional dissonance with these display rules; therefore, our formative measure (see Edwards & Bagozzi, 2000, for differences between formative and reflective measures) consisted of the mean of the anger, anxiety, sadness, and shame options.

Affective delivery. After each session, we used two items to determine the extent to which instructors believed they had maintained display rules. Instructors were asked to rate their agreement with the items "I feel that I was very spirited/enthusiastic during the entire session" and "I was able to keep a positive/upbeat attitude throughout the entire session." Both items used a scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). We calculated an internal consistency of .92 for this measure from the empirical Bayes estimates of the two items for each instructor.

Difficulty maintaining display rules. We used two items to examine the difficulty participants experienced in maintaining display rules during the just-completed session. The first item asked, "How difficult was it to be very spirited/enthusiastic throughout the entire session" (1 = *very difficult*, 5 = *very easy*). The second item was "How difficult was it to keep a positive/upbeat attitude throughout the entire session," rated on the same scale. These items were reverse scored so that higher scores indicate greater difficulty. We calculated an internal consistency of .96 for this measure from the empirical Bayes estimates of the two items for each instructor.

Supervisor ratings of affective delivery. On completion of each session, supervisors completed a brief paper-and-pencil questionnaire for each instructor's affective delivery during the session. There were 8 supervisors, and the number of instructors they were responsible for rating ranged from 5 to 16 people at any given camp. To eliminate variation in the ratings from each particular supervisor, we regressed the ratings on a series of seven

---

[1] The measures examined in this article were taken only at the end of each session. We did, however, also administer other measures at the beginning of each session. Because these measures are irrelevant to the current article, we do not discuss them further.

Table 2
Means, Standard Deviations, and Zero-Order Correlations for Event-Based Measures,
Aggregated to the Individual Level

Variable	M	SD	1	2	3	4
1. Negative emotions (self-ratings)	0.17	0.13	—			
2. Affective delivery (self-ratings)	4.08	0.49	−.31*	—		
3. Difficulty (self-ratings)	3.78	0.65	.37*	.69*	—	
4. Affective delivery (residualized supervisor ratings)	−0.02	0.44	−.11	.05	.06	—

* $p < .05$.

dummy codes (i.e., representing the 8 supervisors) and obtained residual scores. All analyses reported below are based on these residualized scores.

At the end of each session, supervisors answered two items with language parallel to that for the instructors (e.g., "Thinking about how this instructor behaved toward the campers: Rate the instructor's spirit level and enthusiasm during this session"). These items were rated on a 5-point Likert scale (1 = *very low*, 5 = *very high* [with a "not applicable" category in case supervisors did not observe this instructor during the last session]). An internal consistency of .95 was calculated for this measure from the empirical Bayes estimates of the residualized items for each instructor.

Analytic Strategy

Across the 3-day period, there were a total of eight sessions and therefore eight measurement periods (two sessions on Day 1, three sessions on Day 2, and three sessions on Day 3), creating a total of 552 measurement occasions across all participants. Consistent with prior recommendations for this type of data, we used multilevel modeling techniques to test most of our hypotheses (Beal & Weiss, 2003; Bolger, Davis, & Rafaeli, 2003). Level 1 consisted of the eight episodes within each person, and Level 2 examined the slopes and intercepts for each instructor. Hypotheses 5A and 5B involved Level 2 data for both predictor and criterion variables; therefore, these hypotheses were tested with Pearson correlations.

Many researchers have noted that regression models that occur over time may contain forms of time-based dependency, such as trends, cycles, and autoregressive patterns (Beal & Weiss, 2003; West & Hepworth, 1991). These dependencies create systematic variance in the data that can be modeled but is not necessarily a part of the hypothesized relation. For example, irrespective of the predictor variables of interest, affective delivery at time T is probably influenced to some extent by affective delivery at time $T − 1$. It is also possible that criterion variables such as affective delivery or difficulty maintaining display rules contain trends or cycles over the course of the data collection period. We address issues of unwanted trends, cycles, and serial dependency by first entering the following control variables in all of our Level 1 models:[2] To model day-to-day trends, we included a variable that increased monotonically each day. To model cycles within a day, we followed the procedure recommended by West and Hepworth (1991) and included sine and cosine functions with a period of a single day. Finally, to capture variability due to serial dependency, we included the criterion score from the previous time if the previous time had occurred the same day as the current criterion measure.[3] With these control variables included in each regression model, our Level 1 results can be interpreted as the average within-person relation between a particular predictor and criterion over and above the many effects time may have on the criterion.

To analyze the data, we used the HLM5 program (Raudenbush, Bryk, Cheong, & Congdon, 2001). In the absence of a theory relating one's average level of negative affect to differences in the relations between each affective experience and the criterion variables, we concurred with other authors that grand mean centering of our predictors would be most appropriate (Kenny, Korchmaros, & Bolger, 2003; Kreft, de Leeuw, & Aiken, 1995).[4] As is customary with multilevel analysis, we used restricted maximum likelihood estimation procedures. On some of the more complex analyses, the maximum likelihood estimation failed to converge. In these situations, we removed random effects sequentially until the model could be estimated. Our logic was to remove first the random effects of the temporal control variables (e.g., sine and cosine functions, daily trend coefficient) before proceeding to the random effects of the primary predictor variables (e.g., negative emotions, difficulty maintaining display rules). Fortunately, we could estimate all of the models without having to remove random effects from the primary predictor variables.

Results

Hypotheses 1A and 1B

Hypothesis 1A predicted that total amount of regulation would moderate the relation between negative emotions and supervisor ratings of affective delivery. To test this hypothesis, we averaged scores on the Surface Acting and Deep Acting Scales to create a measure of total regulation. On the high end of this continuum are people who highly endorse surface acting strategies as well as deep acting strategies. On the low end of this continuum are people who endorse neither surface acting nor deep acting strategies. This variable significantly predicted the slopes of the negative emotion–affective delivery relation ($\gamma_{11} = .55$), $t(67) = 3.83$, $p < .05$.

One can see the nature of this interaction by noting the slopes for low regulators (minus 1 standard deviation from the mean;

---

[2] It seemed possible that the number of camps an instructor had worked so far that summer, representing a more immediate job experience variable, might serve to modify the relations in which we were interested. We therefore initially included it in all of the hypothesis tests. We found that including it or not including it had no noticeable influence on the results reported. Therefore, we present the models without controlling for this variable.

[3] The criterion score from the previous measurement occasion held a significant positive relation to the current criterion score in almost every model. Although the coefficients for trend and cycles were not always significant, we have included them in all of the reported models to lend consistency to our set of analyses. Notably, the inclusion or exclusion of the trend and cycle variables did not substantially alter the pattern of results.

[4] We compared the results of our grand-mean-centered models with a set of models that used person-mean centering. Once again, although there were slight differences in the exact slope and intercept values, the pattern of effects was identical.

$\gamma_{11} = -.53$) and high regulators (plus 1 standard deviation from the mean; $\gamma_{11} = .23$). For low regulators, experiencing negative emotions resulted in increasingly lower ratings from supervisors. For high regulators, supervisor-rated affective delivery actually increased as negative emotion increased. Although the positive slope for high regulators is somewhat surprising, these results offer clear support for Hypothesis 1A.

After examining the effects of total regulation, we looked at the separate effects of surface and deep acting (Hypothesis 1B). As explained above, we predicted that because the cheerleading supervisors were not aware of the emotion regulation strategies used by the cheerleading instructors, the total amount of either surface or deep acting would predict the nature of the relation between negative emotion and ratings of affective delivery. Consistent with these predictions, amount of both surface acting ($\gamma_{11} = .35$), $t(67) = 3.00$, $p < .05$, and deep acting ($\gamma_{11} = .40$), $t(67) = 3.00$, $p < .05$, predicted the slope of the relation between negative emotion and ratings of affective delivery. The nature of these interactions was very much the same as for total regulation, with negative slopes for low surface actors ($\gamma_{11} = -.40$) and low deep actors ($\gamma_{11} = -.45$) and slightly positive slopes for high surface actors ($\gamma_{11} = .07$) and high deep actors ($\gamma_{11} = .15$). Thus, Hypothesis 1B was also supported.
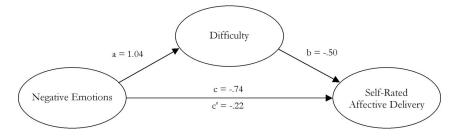
Hypotheses 2A, 2B, and 3

We hypothesized that negative emotions experienced during the sessions would be positively related to perceptions of difficulty maintaining display rules (Hypothesis 2A) and negatively related to perceptions of affective delivery (Hypothesis 2B). Furthermore, Hypothesis 3 suggests that the perceptions of difficulty maintaining display rules during the session should mediate the relation between negative emotion and self-ratings of affective delivery for the session. To test these hypotheses, we followed procedures outlined by Kenny et al. (2003). These authors pointed out that traditional tests of mediation (e.g., Baron & Kenny, 1986) might not always be accurate in multilevel models because each path in the model may be a random effect. Figure 1 shows the path diagram for this test of mediation. Because Paths a and b both were found to be random effects, we used the alternative computational procedures outlined by Kenny et al. (2003). As can be seen in the figure, Paths a, b, and c all were significant, supporting all preconditions for mediation as well as supporting the predictions of Hypotheses 2A and 2B. Path c′ represents the relation between

negative emotions and self-rated affective delivery, with the effects of difficulty maintaining display rules controlled. Although this coefficient remained significant, a Sobel test, modified to account for random effects, found the drop in magnitude to be significant ($z = -2.58$, $p < .05$). In fact, 70% of the relation between negative emotions and self-rated affective delivery was mediated by difficulty maintaining display rules. Thus, Hypothesis 3 was supported as well.
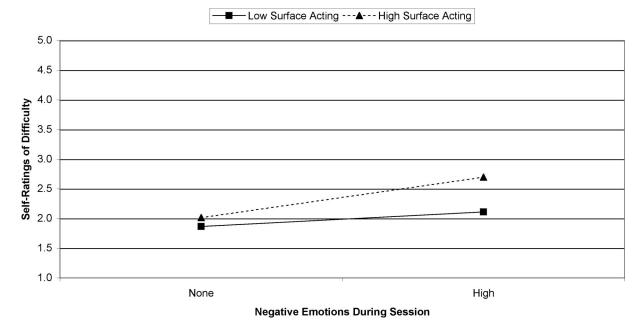
Hypotheses 4A and 4B

Hypothesis 4A suggested that surface acting would moderate the relation between negative emotions and perceptions of difficulty maintaining display rules, such that as surface acting increased, the relation between negative emotions and difficulty maintaining display rules would increase. To test this hypothesis, we ran a multilevel model with surface acting predicting the negative emotion–difficulty slopes. As Figure 2 shows, maintaining display rules became increasingly difficult with the experience of negative emotions during a session, regardless of the level of surface acting. The moderating effect, however, was still significant and in the predicted direction. So, as level of surface acting increased, it became increasingly difficult to maintain display rules during sessions in which negative emotions were experienced ($\gamma_{11} = .42$), $t(67) = 2.11$, $p < .05$. Thus, Hypothesis 4A was supported.

Hypothesis 4B was similar to Hypothesis 4A; the only difference was that the criterion was self-rated affective delivery as opposed to self-rated difficulty. Therefore, we used a similar analysis to test this hypothesis. In this case, however, the moderating effect of surface acting on the slope between negative emotions and self-ratings of affective delivery was not significant ($\gamma_{11} = -.14$), $t(67) = -1.22$, $p = .23$. Hypothesis 4B was therefore not supported.

We conducted exploratory analyses similar to those used to test Hypotheses 4A and 4B to investigate whether deep acting moderated the relation between negative emotions and self-ratings of difficulty and affective delivery. Recall that our initial thoughts were that deep acting would not have as great a moderating effect compared with surface acting. Our analyses revealed that both high and low deep actors displayed a slight positive relation between negative emotion and difficulty, but these relations did not differ across levels of deep acting. A similar pattern emerged for the moderating effects of deep acting on the relation between negative

Figure 1. Level 1 mediating effect of difficulty maintaining display rules on the relation between negative emotion experienced during a session and self-rated affective delivery for that session. Paths are unstandardized and all significant at the $p < .05$ level.

Figure 2. Cross-level moderating effect of surface acting on the relation between negative emotion experienced during a session and self-ratings of difficulty maintaining display rules for that session.

emotion and affective delivery (i.e., a slight negative relation). There was, however, a moderating effect that approached marginal levels of significance ($\gamma_{11} = .22$), $t(67) = 1.62$, $p = .11$, suggesting a slight possibility that deep acting buffers the relation between negative emotions and affective delivery.

Hypotheses 5A and 5B

Hypothesis 5A predicted that surface acting would be related to lower trait levels of positive emotions and greater trait levels of negative emotions. Consistent with this hypothesis, trait positive emotions were significantly related to surface acting ($r = -.31$, $p < .05$), whereas deep acting held no relation ($r = .02$, $p > .05$). Trait negative emotions, despite being in the hypothesized direction, did not attain conventional levels of significance for surface acting ($r = .19$, $p > .05$), and they held virtually no relation to deep acting ($r = .08$, $p > .05$).

In accordance with existing theory, Hypothesis 5B predicted core self-evaluations to be negatively related to surface acting. Using Judge et al.'s (2003) measure of core self-evaluations, we observed a negative correlation with surface acting and no relation to deep acting (surface acting, $r = -.41$, $p < .05$; deep acting, $r = -.17$, $p > .05$). Overall, these results portray the continued use of surface acting as associated with the tendency to experience fewer positive emotions and with a negative evaluation of the role the self plays in those experiences.

Exploratory Analysis of the Effects of Surface Acting on Affective Delivery

Recall that Hypothesis 4B, which predicted that surface acting would moderate the relation between negative emotions and self-rated affective delivery, was not supported. We were surprised by this result and its apparent inconsistency with the mediational chain supported in Hypothesis 3. That is, if an increase in difficulty leads to lower self-ratings of affective delivery, then why would high surface actors not also rate themselves even lower on affective delivery? To investigate this issue more closely, we examined how the mediational patterns changed across levels of surface acting. Specifically, as discussed in Baron and Kenny (1986), we regressed self-rated affective delivery on negative emotions, difficulty, surface acting, the cross-level interaction between negative emotions and surface acting, and the cross-level interaction between difficulty and surface acting. Two pieces of evidence from this analysis are required to demonstrate moderated mediation: First, the mediational effect observed without surface acting was still present, with the gamma dropping from $-.74$ to $-.23$. Second, the effect of difficulty on self-ratings of affective delivery was significantly moderated by surface acting ($\gamma_{21} = .08$), $t(67) = 2.25$, $p < .05$. Thus, it appears that although surface actors experienced more difficulty as a result of their preferred strategy, they did not perceive this difficulty to have an effect on their ability to maintain display rules.

Discussion

This study used an episodic within-person design to examine the relations among emotional experiences, emotion regulation strategies, the difficulty in using those strategies, and self- and supervisor-rated affective delivery. We found that the effect of negative emotion on supervisor ratings of affective delivery was moderated by efforts at regulation, regardless of the type of regulation strategy. In particular, negative emotion had a direct neg-

ative relation to affective delivery, but this relation was attenuated and, indeed, became positive when instructors engaged in either surface or deep acting strategies. In contrast to this moderated relation for supervisor ratings, we found evidence for a mediated relation among negative emotion, difficulty, and self-ratings of affective delivery. Specifically, our data suggest that the influence of negative emotions on self-ratings occurred through its effect on perceived difficulty.

Although amount of regulation, as opposed to the type of strategy, was found to be important for supervisors, type of strategy mattered greatly for self-perceptions. Because surface acting is perceived as a more difficult strategy, the mediational chain between negative emotion, difficulty, and self-ratings should be moderated by the endorsement of surface acting. We did find evidence of moderation with surface acting, but the nature of the relations was more complicated than we had first expected. Specifically, surface actors did have more difficulty maintaining display rules, but high surface actors also exhibited a weaker connection between difficulty and ratings of affective delivery. As a result, we observed moderated mediation: For high surface actors, difficulty was not as strong a mediator of the negative emotion–affective delivery link as it was for low surface actors.

Before discussing the implications and limitations of the results, we should, perhaps, make mention of the unique context in which the study was conducted. We realize that the job of cheerleading instructor is rather atypical; for several reasons, however, we believe that the setting of our research provides an exceptional context for examining the issues of emotional labor. First, the job in question is one in which display rules are strong and widely understood by both employees (instructors) and their supervisors. As with other emotionally laborious jobs described in the literature, these cheerleader instructors were expected, even required, to maintain high levels of positive emotional display regardless of their own emotional state. Second, these instructors were real employees of a bona fide organization. The camps are run by a national, for-profit organization. Instructors are paid and circulate among the camps. Finally, the structure of the training experience provided a perfect opportunity to examine emotional labor in an episodic within-person manner. Each day was broken up into discrete training sessions. Affective delivery in each session was individually observed and evaluated by supervisors.

By examining within-person changes in emotional regulation processes across a series of episodes, this research provides a unique view of emotional labor and offers new insights into how employees regulate emotion. For example, the within-person design means that the findings relevant to negative emotions refer not to how frequently people experience negative emotions compared with other people but rather to how each person responds when in one emotional state compared with how that same person responds when in a different emotional state. The distinction may seem subtle, but it eliminates the unwarranted influence of individual differences in tendencies to experience high or low levels of emotions. Another reason why our within-person design is useful is that it allows us to examine momentary affective processes with a time frame that matches the ongoing experiences of the employee (Beal et al., 2005). Below, we explore how these methods and findings build our understanding of emotional labor processes.

Supervisor Perceptions

We hypothesized and found that, from the perspective of an observer, surface and deep acting strategies can be equally effective in combating the influence of negative emotion on affective delivery, at least during a single episode. This episodic time frame mirrors the experiences of customers in that there is little opportunity to get to know the employees' habitual uses of one strategy or another and therefore generalizes to situations in which actual customers make evaluations of employees' behavior. That is, customers have only brief episodes of contact with most service employees, and, during these encounters, they typically are not engaging in overly scrutinizing observations of emotional expressions.

We think there are three possibilities why supervisors similarly rated surface and deep actors. First, supervisors simply may be poor observers of the instructors' behavior. We find this possibility unlikely, however. First, supervisors had ample time and opportunity to observe and evaluate each instructor. Second, the supervisors were all well-trained employees who had at least 2 years' worth of experience as instructors and were now tasked with ensuring that all cheerleaders enjoyed their experiences at the camps. Finally, if simple poor observation were the explanation for these findings, then we would expect to find *no* relations between predictor variables and supervisory ratings. Instead, we found a complex pattern of results that was consistent with our theoretically derived hypotheses.

A second possibility is that even if supervisors can tell the difference, they do not feel that the sincerity of the expressed emotion is important for evaluating affective delivery. So, positive displays are positive displays, regardless of authenticity, and supervisors who may have been able to tell the difference between authentic and inauthentic positive displays did not adjust their ratings along these lines. Although our study cannot eliminate this possibility, it seems unlikely that supervisors rating affective delivery would ignore the authenticity of the expressed emotions as a factor in their ratings. In fact, Grandey, Fisk, Mattila, Jansen, and Sideman (2005) have shown that the extent to which observers believe an employee to be authentic weighs heavily in their ratings of affective delivery.

The third possibility for the similarity of surface and deep acting is that, for any given episode, the particular strategy choice is unimportant for this sort of affective delivery setting. In other words, perhaps supervisors (and, presumably, campers) only see the positive displays of emotion, missing out on the subtlety of the strategy used to overcome negative emotions. Thus, it seems possible that supervisors may not always require the heartfelt authenticity of a service employee's emotional displays. Although we did not have reports from the campers themselves, it seems likely that they, too, may not have discerned the subtleties of emotion regulation strategy.

Recently, however, Grandey et al. (2005) have found that given the right circumstances, customers may be able to detect authenticity of expressed emotion. This research has manipulated and measured authenticity in ways different from our own, so it is unclear how comparable the studies are with each other. For example, Grandey et al.'s research explicitly incorporates the possibility that behaviors other than emotional expressions can

influence overall customer perceptions of affective delivery. Our study, however, directed supervisors' attention away from such global perceptions and focused it specifically on emotional expression. Future research will need to determine when people can and when they cannot perceive authenticity.

Aside from issues concerning preference for surface or deep acting strategies, our results suggest that the total amount of regulation is much more important. Those who endorsed the higher levels of total regulation actually were rated better as their emotional state worsened. One possible interpretation of this result is that as instructors become aware of their negative emotions, they initiate efforts to behave positively. Those who are particularly apt to regulate may overcompensate, resulting in more positive emotional displays than when no negative emotions are experienced at all. Obviously, this interpretation is speculative; future research should more thoroughly investigate whether regulation efforts surpass levels of emotional expression displayed in the absence of a strong affective experience.

Self-Ratings

Unlike its role in supervisory perceptions, surface acting was involved in a complex pattern of self-perceptions. First, negative emotions were related both to perceived difficulty and self-ratings of affective delivery. In addition, difficulty was found overall to mediate the relation between negative emotions and affective delivery. These simple relations, however, were affected by the level of surface acting: Those who reported higher levels of surface acting found negative emotion episodes much more difficult to manage than those who reported lower levels of surface acting. Because of the previously discussed mediation, we suspected that this very effect would also lead to lower self-ratings of affective delivery. This was not the case. Surface acting did not moderate the relation between negative emotions and self-ratings of affective delivery.

At first blush, this finding was somewhat confusing. Why would high surface actors, who found the experience of negative emotions to be much more difficult, not also consider their own levels of affective delivery to be worse? Our supplemental analyses, however, appear to shed light on this pattern of results. Specifically, surface acting served to moderate the mediated effect. As surface acting increased, the mediating effect of difficulty became weaker. The reason for this weakening mediation appears to be that the connection between difficulty and self-ratings of affective delivery was not as strong for those high in surface acting. Put differently, those who chronically engaged in surface acting strategies recognized the difficulty of their chosen strategy but nevertheless felt it was effective for maintaining positive display rules during the sessions. Those who rarely used surface acting strategies, however, appeared to use difficulty as a gauge for how well they maintained positive displays.

Regulation Strategy and Global Views of Self

Our correlational findings concerning surface acting and overall views of self suggest a problematic pattern for those who engage in surface acting. Specifically, surface acting but not deep acting was negatively related to positive affect and core self-evaluations.

The cross-sectional nature of these data makes it difficult to determine causal ordering of the variables in question. It is possible that those who generally engage in surface acting strategies come to have fewer positive experiences and devalue themselves, but it is also possible that those with fewer positive experiences and lower self-evaluations then adopt surface acting strategies. Regardless of the particular direction of causality, it appears that although surface acting might be effective for a given episode of emotion regulation, it carries with it several more global, negative corollaries. Given that many of these constructs are relatively stable in the short term, a longitudinal study over a greater period of time might help elucidate the potential causal connections.

Although we did not hypothesize relations between self-views and deep acting, the lack of relations is somewhat surprising. It is worthy of note, however, that this measure achieved an alpha level lower than what is typically found. Thus, it is possible that the lack of observed correlations stems from psychometric issues with the scale. We modified the scale to refer to the camp as opposed to the more typical general framing of the scale. It is possible, then, that the framing of the items focused the instructors' attention on specific events and states that occurred in the recent past as opposed to more general and stable semantic memories of strategy use (cf. Robinson & Clore, 2002), resulting in less internal consistency.

Limitations

Although our assessments of regulation strategy were based on chronic and stable conceptualizations, it is entirely possible that these strategies also vary considerably within persons. That is, employees likely use a variety of different strategies to regulate their emotions—some surface and some deep. The two factors were moderately correlated across people ($r = .31$). So, although there is some degree of association between the strategies, there certainly is room for people to endorse primarily one or the other. Indeed, recent research has already begun to examine the extent to which people rely on only one form of emotion regulation or take a more eclectic approach (e.g., Beal, Trougakos, & Weiss, 2006; Totterdell & Holman, 2003). Regardless of whether the use of multiple strategies is common, our results verify that there is meaningful variability at the level of the individual, particularly with respect to other individual-differences constructs such as self-evaluations and trait affectivity.

Another possible limitation of our measures involves the labeling of our difficulty construct as pertaining to the difficulty maintaining display rules. For example, there is an implicit assumption that by reporting that it was difficult to behave in a spirited and enthusiastic manner, instructors were also reporting that they found it difficult to maintain display rules that were desired by the organization.[5] It is possible, then, that some instructors could report that they found it difficult to appear spirited and enthusiastic without considering organizationally desired display rules per se. Thus, a more appropriate name for this construct might be *difficulty maintaining affective delivery* rather than *difficulty maintain-*

---

[5] We thank an anonymous reviewer for bringing this possibility to our attention.

ing display rules. As discussed above, we were fairly certain that instructors were generally aware of the display rule requirements of the organization, but we cannot be certain that this awareness was present with each episode of the camp.

Another limiting factor in our analysis of negative emotions, regulation strategy, and affective delivery is that we have considered only two perspectives, that of the employee and that of the supervisor. We did not examine the role of the customer. In this research context, the experience, control, and expression of emotions are undoubtedly functions of the interaction between instructor and camper. This influence was not eliminated in our study, however. That is, although the influence of the campers was certainly present in the measures we included, we were unable to model it explicitly through camper data. A complete picture of this complex situation therefore awaits research that simultaneously incorporates the perspectives of the employee, the supervisor, and the customer and how these perspectives fluctuate from episode to episode.

Finally, although our sample allowed us a unique and rich context in which we could observe emotional labor processes, its unusual characteristics may have implications for generalizability. Certainly, it is an unusual sample in several respects: All employees were young workers who only worked for the organization during the summer months, typically returning to school during the rest of the year. The fact that our sample consisted of relatively young, seasonal labor should not have much bearing on the regulatory strategies used by these employees, the affective experiences they had during the sessions, or their ability to effectively regulate their emotions during any given episode of performance. Another unusual aspect of this sample was the fairly high levels of job satisfaction (a variable that was collected but is not reported here). Highly satisfied workers are likely to experience more frequent instances of positive emotions and fewer instances of negative emotions at their job relative to those lower in job satisfaction (Weiss, Nicholas, & Daus, 1999). Put differently, our sample may have restricted range on the amount of negative emotion that participants experienced. Thus, if this feature of our sample had any effect on our results, the most likely is an underestimation of the effects.

There were other notable characteristics of our sample, however. An examination of the means for affective delivery and difficulty maintaining display rules makes it obvious that instructors were usually effective at providing positive displays and that they did not find it terribly difficult to do so. Nevertheless, there is variability within people in these variables, and our results suggest that this variability is meaningful and explainable. Understanding the potential consequences of such high scores on affective delivery requires some knowledge of why the ratings were typically high (or low, for difficulty). One reason is that there may be a self-serving bias in the self-ratings and a leniency bias in the supervisor ratings. If this is the case, then, again, our interpretation is relatively unaffected except for potentially underestimating the effects. Presumably, if there are systematic relations among negative emotions, difficulty, and surface acting at the high end of affective delivery, then these relations would apply equally (if not more so) at lower ends of the spectrum. We cannot test to see whether this truly is the case, but there is no theoretical reason that we can think of why these relations would not operate at lower

levels of affective delivery. In our view, the relative range of the scales at which these variables occur does not diminish the theoretical relations between them. A second reason for the high or low ratings may be that poor affective displays from instructors were exceedingly rare. The question then becomes how useful it is to interpret variation in what is primarily good performance. Although a certain level of performance may be sufficient in some contexts, most utility analyses suggest that predicting any differences in performance provides a benefit to the organization (Schmitt & Chan, 1998).

Applications

We placed the bulk of our efforts on refining and extending theories of emotional experience and regulation in work settings; however, our results have much to say to practitioners as well. Training efforts in emotional labor jobs can be informed by the current study. The main point is that regulation strategies work, at least in the eyes of supervisors. To the extent that people engage in any regulation strategy, they will more effectively convey organizationally desired display rules. So, in efforts to train employees to maintain these display rules, organizations should focus on the use of regulation strategies to overcome dissonant emotional experiences. As for whether there is a preference for surface or deep acting, it appears that reliance on surface acting is not a good idea: Surface acting is a more difficult strategy, is associated with a variety of long-term negative consequences, and has the potential for burdening cognitive resources (Richards & Gross, 1999). Thus, it is probably not a good idea to encourage strategies that emphasize "grinning and bearing it." Instead, employees should focus more on reinterpreting a negative situation or focusing their thoughts on more positive experiences.

These results also have implications for selection. In jobs that require high levels of emotion regulation, those who do not use regulatory strategies allow their own emotions to get the better of them. If supervisors can detect the influence of negative emotional states on affective delivery, then it seems highly likely that customers can as well. Therefore, strategy use could very well predict the employees who are more likely to be successful in jobs involving emotional labor. Furthermore, in thinking about the individuals who endorsed neither regulation strategy, we were often reminded of the self-monitoring literature (Gangestad & Snyder, 2000). Although this literature covers more than just emotional experience, it certainly seems likely that those who do not feel it necessary to control their emotional states, even when it is an explicit part of their job, are low self-monitors. Various studies have pointed out the potential for high self-monitors to be effective members of organizations (e.g., Baron, 1989; Kilduff & Day, 1994). Future research should determine the extent to which dispositional levels of emotion regulation or self-monitoring more generally can be used as a selection tool in jobs with strong emotional labor components.

Our study helps our understanding of the complex relations among employees' experience of negative emotions, their efforts and difficulties in controlling these experiences, and the consequences of these efforts, both immediately and globally. Of course, it also highlights the overall importance of examining affective experiences at work (Weiss & Cropanzano, 1996), particularly

from a within-person, episodic framework (Beal et al., 2005). We have focused on a job particularly high in emotional labor, but we believe that these processes generalize to many jobs. Emotional experiences are not the sole province of employees who interact with customers, and, to the extent that employee emotions are regulated, our findings hold relevance.

References

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: W. H. Freeman.

Baron, R. A. (1989). Personality and organizational conflict: Effects of Type A behavior pattern and self-monitoring. *Organizational Behavior and Human Decision Processes, 44,* 281–296.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173–1182.

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology, 74,* 1252–1265.

Beal, D. J., Trougakos, J. P., & Weiss, H. M. (2006, May). The dynamics of emotion regulation strategies. In J. C. Wallace & E. M. Richard (Chairs), *New directions in emotional labor research.* Symposium conducted at the annual meeting of the Society for Industrial and Organizational Psychology, Dallas, TX.

Beal, D. J., & Weiss, H. M. (2003). Methods of ecological momentary assessment in organizational research. *Organizational Research Methods, 6,* 440–464.

Beal, D. J., Weiss, H. M., Barros, E., & MacDermid, S. M. (2005). An episodic process model of affective influences on performance. *Journal of Applied Psychology, 90,* 1054–1068.

Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6. pp. 1–62). New York: Academic Press.

Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology, 54,* 579–616.

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

Brotheridge, C. M., & Lee, R. T. (2003). Development and validation of the Emotional Labour Scale. *Journal of Occupational and Organizational Psychology, 76,* 365–379.

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

DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological Bulletin, 111,* 203–243.

Diener, E., Smith, H., & Fujita, F. (1995). The personality structure of affect. *Journal of Personality and Social Psychology, 69,* 130–141.

Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods, 5,* 155–174.

Erickson, R. J., & Wharton, A. S. (1997). Inauthenticity and depression: Assessing the consequences of interactive service work. *Work & Occupations, 24,* 188–213.

Farh, J., & Dobbins, G. H. (1989). Effects of self-esteem on leniency bias in self-reports of performance: A structural equation model analysis. *Personnel Psychology, 42,* 835–850.

Frank, M. G., Ekman, P., & Friesen, W. V. (1993). Behavioral markers and recognizability of the smile of enjoyment. *Journal of Personality and Social Psychology, 64,* 83–93.

Frijda, N. H. (1993). Moods, emotion episodes and emotions. In M. Lewis & J. M. Haviland (Eds.), *Handbook of emotions* (pp. 381–403). New York: Guilford Press.

Gangestad, S. W., & Snyder, M. (2000). Self-monitoring: Appraisal and reappraisal. *Psychological Bulletin, 126,* 530–555.

Grandey, A. A. (2000). Emotional 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 and deep acting as determinants of emotional exhaustion and peer-rated service delivery. *Academy of Management Journal, 46,* 86–96.

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.

Gross, J. J. (1998a). Antecedent- and response-focused emotion regulation: Divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology, 74,* 224–237.

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

Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology, 39,* 281–291.

Hochschild, A. R. (1983). *The managed heart: Commercialization of human feeling.* Berkeley: University of California Press.

Judge, T. A., Erez, A., Bono, J. E., & Thoreson, C. J. (2003). The Core Self-Evaluations Scale: Development of a measure. *Personnel Psychology, 56,* 303–331.

Kenny, D. A., Korchmaros, J. D., & Bolger, N. (2003). Lower level mediation in multilevel models. *Psychological Methods, 8,* 115–128.

Kilduff, M., & Day, D. V. (1994). Do chameleons get ahead? The effects of self-monitoring on managerial careers. *Academy of Management Journal, 37,* 1047–1060.

Kreft, I. G., de Leeuw, J., & Aiken, L. S. (1995). The effect of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research, 30,* 1–21.

Lane, J., & Herriot, P. (1990). Self-ratings, supervisor ratings, positions and performance. *Journal of Occupational Psychology, 63,* 77–88.

Mann, S., Vrij, A., & Bull, R. (2004). Detecting true lies: Police officers' ability to detect suspects' lies. *Journal of Applied Psychology, 89,* 137–149.

McWilliams, L. A., Cox, B. J., & Enns, M. W. (2003). Use of the Coping Inventory for Stressful Situations in a clinically depressed sample: Factor structure, personality correlates, and prediction of distress. *Journal of Clinical Psychology, 59,* 423–437.

Morris, J. A., & Feldman, D. C. (1996). The dimensions, antecedents, and consequences of emotional labor. *Academy of Management Review, 21,* 986–1010.

Pugh, S. D. (2001). Service with a smile: Emotional contagion in the service encounter. *Academy of Management Journal, 44,* 1018–1027.

Pugliesi, K. (1999). The consequences of emotional labor: Effects on work stress, job satisfaction, and well-being. *Motivation and Emotion, 23,* 125–154.

Raudenbush, S., Bryk, A., Cheong, Y. F., & Congdon, R. (2001). *HLM5: Hierarchical linear and nonlinear modeling.* Lincolnwood, IL: Scientific Software International, Inc.

Richards, J. M., & Gross, J. J. (1999). Composure at any cost? The cognitive consequences of emotion suppression. *Personality and Social Psychology Bulletin, 25,* 1033–1044.

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin, 128,* 934–960.

Schmitt, N., & Chan, D. (1998). *Personnel selection: A theoretical approach.* Thousand Oaks, CA: Sage

Schneider, B., & Bowen, D. E. (1995). *Winning the service game.* Boston: Harvard Business School Press.

Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: Further exploration of a prototype approach. *Journal of Personality and Social Psychology, 52,* 1061–1086.

Thoresen, C. J., Kaplan, S. A., Barsky, A. P., Warren, C. R., & de Chermont, K. (2003). The affective underpinnings of job perceptions and attitudes: A meta-analytic review and integration. *Psychological Bulletin, 129,* 914–945.

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.

Tsai, W., & Huang, Y. (2002). Mechanisms linking employee affective delivery and customer behavioral intentions. *Journal of Applied Psychology, 87,* 1001–1008.

Weiss, H. M., & Cropanzano, R. (1996). Affective events theory: A theoretical discussion of the structure, causes and consequences of affective experiences at work. In B. M. Staw & L. L. Cummings (Eds.),

Research in organizational behavior (Vol. 18, pp. 1–74). New York: Elsevier Science/JAI Press.

Weiss, H. M., Nicholas, J. P., & Daus, C. (1999). An examination of the joint effects of affective experiences and job beliefs on job satisfaction and variations in affective experiences over time. *Organizational Behavior and Human Decision Processes, 78,* 1–24.

Wells, L. E., & Sweeney, P. D. (1986). A test of three models of bias in self-assessment. *Social Psychology Quarterly, 49,* 1–10.

West, S. G., & Hepworth, J. T. (1991). Statistical issues in the study of temporal data: Daily experience. *Journal of Personality, 59,* 609–662.

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.

Instructions to Authors

For Instructions to Authors, please consult the January 2006 issue of the volume or visit www.apa.org/journals/apl and click on the "Instructions to authors" link in the Journal Info box on the right.