DETELINA MARINOVA, SUNIL K. SINGH, and JAGDIP SINGH*

This study examines the impact of frontline employees' problem solving on customer satisfaction (CSAT) during ongoing interactions prompted by service failures and complaints. Using outsourced regulation theory, the authors predict negative moderating effects of frontline relational work and displayed affect on the dynamic influence of frontline solving work on CSAT. Frontline employees' verbal (nonverbal) cues provide the basis to identify solving and relational work (displayed affect). The authors test hypotheses with data from video recordings of real-life problem-solving interactions involving airline customers as well as a controlled experimental study. They find that frontline solving work has a positive effect on CSAT, and it increases in magnitude as the interaction unfolds. However, this positive effect becomes *weaker* for relatively higher levels of frontline relational work or displayed affect and, conversely, *stronger* for relatively lower levels over time. In summary, overdoing relational work and overdisplaying positive affect diminish the efficacy of problem-solving interactions, a finding that provides implications for theory and practice.

Keywords: customer satisfaction, service recovery, complaint handling, video recording, verbal/nonverbal

Online Supplement: http://dx.doi.org/10.1509/jmr.15.0243

Frontline Problem-Solving Effectiveness: A Dynamic Analysis of Verbal and Nonverbal Cues

Service problems of varying natures and intensities occur every day at the frontlines of organizations (Bitner, Booms, and Mohr 1994; Smith, Bolton, and Wagner 1999). In airline settings, for example, problems may arise as a result of service failures (e.g., lost baggage), externally caused service interruptions (e.g., weather-related delays), customer noncompliance (e.g., late for check-in), or problems anticipated in advance (e.g., overbooked flights). Two main streams of research address such frontline problem-solving situations: consumer dissatisfaction and complaint-handling literature, which focuses on dissatisfied consumers' responses or complaints and the resolution efforts of companies and their employees (Gelbrich and Roschk 2011; Richins 1983), and service recovery literature, which examines service failures (whether voiced or not) and companies' efforts to return consumers to a satisfied state (Maxham and Netemeyer 2002; Smith, Bolton, and Wagner 1999).

Across these different literature streams, problem-solving interactions consistently show several common features. First, they cannot be scripted easily and often involve on-the-spot improvisations to address specific service problems as they arise (Heritage and Maynard 2006). Second, they tend to be emotionally charged and marked by customer frustration, which increases the potential for miscommunication and

*Detelina Marinova is Frances Ridge Gay MBA Professor of Marketing, Robert J. Trulaske Sr. College of Business, University of Missouri (email: marinovad@missouri.edu). Sunil K. Singh is Assistant Professor of Marketing, College of Business, University of Nebraska–Lincoln (email: ssingh10@unl.edu); Sunil completed the research work as a doctoral student at the University of Missouri. Jagdip Singh is AT&T Professor of Marketing, Weatherhead School of Management, Case Western Reserve University (email: Jagdip.singh@case.edu). All three authors contributed equally. This work was supported by a Marketing Science Institute Grant (4-177) and Trulaske College of Business Large Grant. The authors thank Paige Anderson, Ashley Cole, Sabina Shrestha, Kathryn Plummer, Kat Birkenbeuel, Nina Amedin, Sabrina Soenksen, Erin Gomez, and Abbene Johnston for their help with coding of verbal and nonverbal cues. Jeffrey Inman served as associate editor for this article.

ISSN: 0022-2437 (print)
1547-7193 (electronic)

Journal of Marketing Research
Vol. LV (April 2018), 178–192
DOI: 10.1509/jmr.15.0243

misperception (Groth and Grandey 2012). Third, customers—and increasingly, public citizens in general (Stack 2017)—perceive them as critical events that leave enduring memory traces and trigger recalibration of relationships with the service provider (Bitner, Booms, and Mohr 1994). Because problem-solving interactions are uncertain, salient, emotionally charged, and demanding, many companies invest significant resources to get them right (Spector and McCarthy 2005). Firms with reputations for exemplary customer problem solving, such as Southwest and Nordstrom, have consistently high customer satisfaction (CSAT) ratings (Anderson and Sullivan 1993; Mittal and Frennea 2010; Oliver 2010).

However, most research has examined problem solving at the frontline by studying customers' response states either before the problem-solving effort, such as their causal attributions, emotions, expectations, or actions (e.g., complaint; Kelley and Davis 1994; Richins 1983; Ringberg, Odekerken-Schröder, and Christensen 2007), or after problem solving, with a focus on the nature (e.g., compensation, apology), fairness (e.g., distributive, procedural, interactional), and outcomes (satisfaction) of frontline employees' (FLEs') actions (Smith, Bolton, and Wagner 1999; Tax, Brown, and Chandrashekaran 1998). Largely overlooked are the problem-solving *processes*—that is, dynamics of frontline work and displayed affect that take place *during* problem-solving interactions as FLEs come to understand the problems and construct solutions in real time. Shifting attention from states to processes parallels the shift from cross-sectional to longitudinal analyses of time-varying effects during problem-solving interactions, as evident in related literature (DeChurch, Mesmer-Magnus, and Doty 2013).[1]

To examine problem-solving processes and effectiveness in frontline interactions, we address three outstanding questions. First, does FLEs' relational work (e.g., empathy) and positive affect (e.g., smile) help or hinder problem-solving effectiveness? Most services marketing research has emphasized the positive impact of FLEs' relational work, including empathy, courtesy, and apology, on CSAT (Zeithaml, Berry, and Parasuraman 1996). However, recent meta-analytic studies of service failure have concluded that relational work is less helpful in failure situations that do not involve psychological loss (Roschk and Gelbrich 2014). Although positive affect may help soothe customers in distress, Rafaeli et al. (2017) find that it can be counterproductive in contexts marked by time pressures.

Second, what are the dynamics of frontline problem solving? Contradictory findings in extant research may be attributable, at least in part, to a common reliance on cross-sectional approaches that aggregate time to examine the effects of frontline work on postexchange CSAT. These studies confound the FLE relational work that may be helpful in earlier stages but unhelpful in later phases of the problem-solving interaction. Theorizing about time-varying effects of FLE work in problem-solving interactions is critical to advance prior literature.

Third, do FLE nonverbal cues influence problem-solving effectiveness? Customers use nonverbal cues to infer affective qualities of FLEs (Puccinelli, Motyka, and Grewal 2010), but prior studies have overlooked the role of these nonverbal cues for determining problem-solving effectiveness. To the extent that FLE nonverbal cues are salient and distinct input to customer evaluations, studies of problem-solving effectiveness may suffer from misspecification biases.

To address these questions, we examine frontline problem solving in real time during ongoing, face-to-face interactions in which solutions are developed and negotiated under time pressure. Drawing on outsourced regulation theory, we theorize the moderating effects of FLE relational work and displayed affect on the time-varying relationship between FLE solving work and CSAT. With a mixed-method design, we test these predictions in two studies. Study 1 includes a longitudinal panel of field data from fly-on-the-wall (FoTW) video recordings of problem-solving interactions involving airline travel that occur naturalistically at actual airports. In Study 2, we conduct a controlled study using actual airline passengers.

Our study makes four main contributions. First, we conceptualize and empirically isolate the dynamic and interactive influence of FLE work and displayed affect on CSAT. In a novel approach, we conceive of solving and relational work as separate dimensions of FLE work, which customers infer from verbal cues. We further use displayed affect to indicate FLE displayed emotion, which customers infer from nonverbal facial, body, and gestural cues. We validate separate dictionaries for the distinctive verbal cues associated with FLE solving and relational work, as well as for the nonverbal cues FLEs display to signal affect. Second, we show that FLE solving work positively affects CSAT, and this effect increases in magnitude during the interaction. Third, we demonstrate that the influence of FLE solving work on CSAT remains significant, even if service recovery is not feasible. That is, customers appear to separate problem-solving processes from solution outcomes and value FLE efforts to develop a range of varied solution options available for customer selection, even if the selected outcome is less satisfying. Fourth, the positive association between FLEs' frontline solving work and CSAT becomes weaker for relatively higher levels of FLE relational work or displayed affect but stronger for relatively lower levels of relational work or displayed affect over time. Thus, overdoing relational work or positive affect is counterproductive in problem-solving interactions.

CONCEPTUAL DEVELOPMENT AND HYPOTHESES

Frontline Problem-Solving Interactions and Work

For effective problem solving, firms often rely on detailed scripts and routines to guide and direct their FLEs' actions. However, to address emergent and unanticipated customer problems, FLEs must use their own discretion and mindfulness to enact behaviors that may deviate from or extend prevailing role scripts, or else they creatively construct behavioral patterns that differ from role expectations. In this sense, it is important to distinguish between role *expectations*, as coded in norms and rules, and *behaviors* enacted in situ, which we refer to as work (Okhuysen et al. 2013). Enacted behaviors are observable, indicate employee agency and effort, inform customer inferences (e.g., helpful/not helpful), and serve as input to customer responses, so they are key to understanding customer outcomes in ongoing service interactions (Bradley et al. 2013). Instead of trying to access what FLEs think or intend, we focus on the work that FLEs actually perform and display during customer interactions.

[1]We thank an anonymous reviewer for noting the state versus process distinction in team conflict literature. DeChurch, Mesmer-Magnus, and Doty (2013) show that team conflict states account for just 2% of the incremental variance in team performance, but team processes account for 13%.

To conceptualize FLE work, we first consulted service quality literature, in which scales such as SERVQUAL (Parasuraman, Zeithaml, and Berry 1988) have been developed to capture customers' cumulative postconsumption experience (but not within-interaction FLE behaviors). Some dimensions of SERVQUAL, such as reliability, are not relevant for studying problem-solving interactions because many service problems are failures of reliability, for which FLEs often must improvise or construct solutions on the spot. Thus, consistent with our focus on the process, rather than the state, of problem-solving interactions, we draw on role theory (Biddle 1986) and service interaction research (Bitner, Booms, and Mohr 1994; Bradley et al. 2013). In this domain, Homburg, Müller, and Klarmann (2011) observe that FLEs demonstrate customer orientation by blending functional (task-oriented) and relational (relationship-oriented) role dimensions. Similarly, in a service recovery context, Liao (2007) proposes that the role expectations of FLEs include instrumental (prompt handling, explaining, resolving concerns) and relational (listening, apologizing, helping, being courteous) dimensions. In a recent, detailed analysis of service interaction research, Bradley et al. (2013) identify two meta-categories of behaviors: task behaviors focused on core service delivery to customers (e.g., competence) and relational behaviors focused on the emotional relationship with customers (e.g., empathy).

Conceptually, the task and relational dimensions of FLE behaviors correspond to the psychological constructs of competence and warmth (Abele and Wojciszke 2014), as recently studied in contexts of branding, conspicuous consumption (Scott, Mende, and Bolton 2013), and service provider choice (Kirmani et al. 2017). This conceptual correspondence is useful for drawing linkages to broader marketing literature, but FLE task and relational behaviors demonstrate several notable nuances. For example, customers must depend on FLEs to resolve the problem, which is distinct from the relatively unconstrained process in other contexts (e.g., service provider choice in Kirmani et al. [2017]). In addition, problems are solved in real time, and customers form evaluations on the spot, unlike the typical search process in choice decisions.

Accordingly, we define FLE solving work as verbal cues that indicate the FLE's competence (e.g., knowledge, skills) and action orientation (e.g., engaged, proactive) toward effective problem solving. We define FLE relational work as verbal cues that indicate the FLE's compassion (e.g., empathy, caring) and agreeableness (e.g., courtesy, respect) to support effective customer bonding.[2] Verbal cues signal a communication partner's attitudes and motivation, as well as message content. Nonverbal cues, including facial, bodily, and hand gestures (Aviezer, Trope, and Todorov 2012; Bonoma and Felder 1977), may reinforce or contradict verbal cues, but they also provide additional information about the sender's affect, whether aligned or not with the more consciously managed verbal cues (Puccinelli, Motyka, and Grewal 2010). We define FLE displayed affect as the nonverbal cues displayed by the FLE during problem-solving interactions that indicate his or her feeling state (i.e., positive, negative, or neutral). In practice, verbal and nonverbal cues may vary fluidly and systematically, in line with problem-solving progress, which typically involves three phases: sensing (e.g., comprehending problems), seeking (e.g., generating solutions), and settling (e.g., implementing solutions). However, these phases are neither demarcated cleanly nor ordered systematically.

Frontline Solving Work and CSAT

We view problem solving as a process motivated by goal pursuit, and we predict that dissatisfied customers monitor the present situation, relative to some internal standard for satisfaction, in accord with self-regulation theory (Carver and Scheier 1990). Similar to self-regulated goal pursuit, a problem-solving interaction is initiated by a customer with a goal to resolve a pressing problem; however, unlike self-regulated goal pursuit, the locus for the problem resolution is the FLE's actions. Without an FLE's problem-solving actions, a dissatisfied customer cannot attain the goals (s)he seeks. The notion of separating goals and actions is anticipated by outsourced self-regulation theory in interpersonal contexts (Fitzsimons and Finkel 2011). Specifically, goal pursuit is sourced out to instrumental others, who provide effort and resources and engage in actions to facilitate that goal pursuit and attainment.

In line with outsourced regulation theory, we also theorize that customers actively and continuously monitor the outsourced (to FLEs) solving work, relative to some internal standard of expected discrepancy reduction at any particular point in time, to assess whether FLEs' outsourced actions are moving toward goal attainment. If not, a feedback loop prompts increasing frustration and dissatisfaction. Conversely, if the actions exceed expectations, the feedback loop yields positive satisfaction. The outsourced regulation mechanism also suggests tracking the rate of discrepancy reduction in goal pursuit, which can evoke anticipatory feelings of satisfaction (dissatisfaction) if the rate exceeds (lags) an internal standard for progress. Therefore, feedback monitoring is sensitive to both the level and rate of discrepancy reduction at a particular point in time, given the time already invested in goal pursuit (Fishbach and Finkelstein 2012).

In practice, customers rely on cues available in FLEs' language to monitor the level of discrepancy reduction achieved at any point in the problem-solving interaction, as well as the progress achieved toward problem-solving goals. Verbal cues include words and phrases that FLEs use to seek information, communicate options, and explain solutions. Customers use these cues to evaluate problem-solving progress and effectiveness (Groth and Grandey 2012). Problem-solving outcomes, such as compensation and distributive justice, have been identified as the best means to restore transaction-specific CSAT, according to two meta-analyses of service failures (Roschk and Gelbrich 2014) and complaint handling (Gelbrich and Roschk 2011). Although static postconsumption evaluations of problem solving have been widely studied (e.g., 142 studies, 2 meta-analyses), FLEs' solving work during interactions, as reflected in their verbal cues, has not been examined. In social communications, vocabularies are powerful mechanisms of influence. For example, words reflecting a professional logic vocabulary (e.g., practice, quality, lasting) enhance the likelihood that architects win project bids, compared with words signaling a business logic vocabulary (e.g., client, works, needs; Jones and Livne-Tarandach 2008). In customer service contexts, Sturdy and Fleming (2003) show that firms can train FLEs to engage in "verbal labor" by emphasizing a service vocabulary with words that promote positive customer inferences and outcomes. We know of no study examining a vocabulary of effective problem-solving words.

[2]Bradley et al. (2013) also identify a self-referent category, related to the actor's own goals, needs, and interests, but that category is not relevant for the current study.

Customers' dissatisfaction (satisfaction) should increase if their assessment of observed verbal cues indicates that the FLEs' solving work is ineffective (effective). Furthermore, FLEs' solving work should influence CSAT throughout the problem-solving interaction, such that its effect increases with time. Outsourced regulation considerations vary across the sensing, seeking, and settling phases because of the distinct expectations in each phase. Sensing, which occurs early in customer interactions, usually requires FLEs to gather information to understand the nature of customer problems. From a customer perspective, sensing does little to signal how the problem will be solved, so regulatory feedback indicates that the FLE, as the instrumental other, has made little progress toward discrepancy reduction. Customer dissatisfaction should remain largely unaltered. However, progress is likely discernible during seeking activities because the FLE focuses on generating feasible options to address the problem. The FLE also communicates with the customer to seek additional information, construct relevant options, and explore the customer's willingness to accept different options. Consistent with outsourced regulation theory, customers actively monitor these verbal cues to infer progress toward problem solving. They should discern positive progress in their goal pursuit when seeking work is effective, which prompts a positive change in their satisfaction. Finally, during settling activities, FLEs communicate one or more solution options, respond to objections by reworking solutions, and implement the ultimate solution. The concreteness of these solution options and alacrity of solution implementation provides tangible evidence of progress, which should increase CSAT. Thus,

> H_1: FLEs' solving work has a positive effect on CSAT, and this effect increases in magnitude during the problem-solving interaction.

Relational Work Moderates the Influence of Frontline Solving Work on CSAT

Prior studies have recognized the positive role of relational work in frontline problem solving (Fang, Luo, and Jiang 2013; Smith, Bolton, and Wagner 1999). Although it cannot directly solve customer problems, relational work features prosocial behaviors that can enhance the effectiveness of FLEs' solving efforts. Prosocial behaviors signal that FLEs understand customers' problems and are interested in solving them (e.g., "If you don't apologize and don't make customers know you care, it's very difficult to recover the customer afterward"; Stoller 2005).

However, in their meta-analysis of complaint-handling research, Gelbrich and Roschk (2011, p. 36) conclude that interactional justice, a concept akin to relational behavior, exerts "a negligible impact, if at all, on transaction-specific satisfaction." Moreover, Menon and Dubé (2007) argue that relational work may be less useful in situations in which customers seek satisfactory solutions to a service problem that has caused some unexpected, often intolerable, inconvenience. In these situations, customers may perceive that relational actions dilute or divert FLEs' focus from their solving work. Thus, relational work can trigger contrast effects between customers' expectations that FLEs should focus on problem resolution (solving work) and their observation of unhelpful prosocial actions. According to Rafaeli and Sutton (1987), the relationship between retail store sales and FLEs' emotional work—captured by relational actions such as greeting and thanking customers—is moderately but significantly negative. These authors argue that when a store is busy, with long lines that signal time pressure, displays of emotional (relational) work are counterproductive and frustrate customers who leave without completing their purchases. According to Menon and Dubé (2007), customers under time pressure evaluate their interactions with FLEs more positively if the FLEs focus on instrumental actions (solving work), but less so if they engage in emotional work (relational work). Thus, a low level of relational work may be effective, but moderate or high levels induce contrast effects.

We therefore predict that FLE relational work negatively moderates the effect of FLE solving work on CSAT, and this negative effect grows in significance (becomes more negative) over the course of the interaction. During sensing, at the beginning of the interaction, FLE relational work likely includes empathetic talk (e.g., "I understand," "I am sorry"), which customers perceive as customary and reasonable. It also might help diffuse customers' negative emotions, so the FLE can more readily understand the problem and establish a common ground. In this stage, some relational work could enhance the efficacy of solving work, but vigorous relational work involving small talk (e.g., "Isn't it just freezing today?") is unlikely to be helpful. During seeking, customers want FLEs to focus on solving and have little tolerance for distraction, so the range of acceptable relational work likely narrows. Even customary relational work (e.g., repeatedly apologizing, constantly empathizing) may raise customers' concerns about timely progress toward effective problem solving. That is, the negative moderating effect of relational work likely increases in the seeking phase relative to the sensing phase. Finally, effective settling requires the FLE to work out the solution details, adapting them to customers' preferences and executing the solution with minimum delays. Attention to detail, focused action, and completeness in solving work are prominent criteria. We expect this emphasis on solving work in the settling phase to crowd out the need and tolerance for FLE relational work. Thus,

> H_2: FLEs' relational work negatively moderates the impact of their solving work on CSAT over time, such that the positive association between FLE solving work and CSAT weakens (strengthens) at higher (lower) levels of FLE relational work.

Displayed Affect Moderates the Influence of Frontline Solving Work on CSAT

Bonoma and Felder (1977) emphasize that facial (e.g., smiling, nodding, eye contact), bodily (e.g., personal distance), and gestural (e.g., touch, wave) cues are just as prevalent and salient as verbal cues in interpersonal interactions. Studies of nonverbal cues in diverse settings—including client presentations, training, service relationships, financial services, and retail settings—consistently show that nonverbal cues are actively perceived and processed in face-to-face interactions. Customers tend to perceive nonverbal cues as more authentic or less prone to impression management relative to the more consciously managed verbal cues (Puccinelli, Motyka, and Grewal 2010) and process them to infer the affective qualities of the FLE. That is, FLEs' authentic affective states leak through their nonverbal cues, and customers use those cues to evaluate FLEs' internal affect toward them and the problem.

Affect inferred by customers from FLE nonverbal cues should conform to the contrast mechanism outlined previous for relational work. Because customers perceive nonverbal cues as more authentic and diagnostic than verbal cues (Bonoma and Felder 1977), the moderating effect of FLEs' displayed affect is expected to be stronger (more negative) than that of FLEs' relational work. Customers who actively monitor FLEs' problem-solving actions likely have limited tolerance for overly positive displayed affect, such that affective states that would be appropriate in typical or routine customer interactions evoke contrast effects and are perceived less favorably by customers. Customers also likely perceive positive displayed affect as less conducive to effortful and diligent problem solving. As Paul, Hennig-Thurau, and Groth (2014) show, FLEs' nonverbal cues during dining experiences have stronger effects on customers' service quality perceptions than do verbal cues, though their study considers business-as-usual service interactions. We know of no study that examines these effects in a problem-solving context.

In terms of dynamic effects, the negative moderating influence of FLE displayed affect is expected to strengthen as the problem-solving interaction progresses from sensing to settling. In the initial stages, customers likely perceive FLE positive displayed affect as an acceptable norm for initiating interactions, but we expect displays of positive affect in the seeking and settling phases to appear increasingly inappropriate and insensitive to customer problems. Thus,

> H_3: FLE displayed affect negatively moderates the impact of solving work on CSAT over time, such that the positive association between FLE solving work and CSAT weakens (strengthens) for higher (lower) levels of displayed affect.

STUDY 1: AIRLINE FIELD STUDY

Research Setting

To test our hypotheses and their cross-contextual generalizability, we need longitudinal, in-situ data about ongoing problem-solving interactions between FLEs and customers. Prior research has advocated a prospective, naturalistic, observational design (Ma and Dubé 2011) to mitigate the recall and desirability biases of retrospective self-report studies. To overcome both obtrusiveness (e.g., observers hinder natural interactions) and incompleteness (e.g., observers miss details) concerns, video-recorded observations of real-time interactions are effective (Echeverri 2005). However, recording customers raises privacy concerns, and firms rarely use video recording for purposes other than safety, theft, and criminal control. Therefore, we turned to FoTW video recordings of problem-solving interactions to obtain observational data in natural settings. This method captures events in their naturalistic settings without scripting but with consent of the involved parties. Prior research uses FoTW video recordings to investigate media (Doyle 1998) and communication (Nabi et al. 2003) topics. It prioritizes a naturalistic setting but relaxes the observational condition by securing customer consent to record during actual service experiences. This consent process might limit authenticity and foster impressionistic behavior, but customers and employees are quick to acclimatize after the consent phase, such that awareness of the video recording tends to recede (Penner et al. 2007).

We secured the FoTW series "Airline" by purchasing original data from a broadcasting company (ITV UK). These data are particularly suitable for our study. First, the primary focus is problem-solving interactions during daily "business-as-usual" FLE–customer interactions at check-in, at the departure gates, and in flight. They include easyJet's operations at the Liverpool and Luton airports ("Airline UK") and Southwest's operations at the Chicago and Los Angeles airports ("Airline US"). Second, the data are substantial. The series includes 100 U.K. video-recorded episodes during 1998–2006 and 18 U.S. episodes during 2004; each episode includes multiple problem-solving interactions (usually two or three per episode). Third, the Airline FoTW series captures problem-solving interactions in a naturalistic setting with no scripting. To check for data validity, we conducted structured interviews with the series producers and editor.

Data Quality Assessment

In structured interviews with two producers and one editor of Airline UK, we asked about the integrity of the problem-solving interactions in the video recordings, criteria used to identify which interactions to record, and any constraints that guided the recording and editing of the interactions. The producers noted that they randomly selected real-life customer interactions as they occurred, without any interference, during a regular business day. One camera crew was assigned to each airport, to limit tendencies to pick and choose interactions. Typically, the camera crew waited near a check-in counter and started shooting an event once a customer presented a problem and gave permission to record the event (refusal incidence was <10%). The camera crew was also passport ready and sometimes flew with the customer to complete a story. The producers confirmed that their central objective was to capture authentic interactions; the camera crew was specifically trained not to intrude in the problem-solving event.

The series editor presented the protocols for capturing and cutting video recordings, as independently verified by the producers. The camera crew was instructed to capture the problem-solving interactions in as complete a form as possible. Shooting time ranged from 30 minutes to more than 3 hours per interaction. The established protocols helped trim the recorded content to 10 minutes or less by eliminating content that did not feature direct interactions between the customer and an airline employee. Voice-overs filled in details about nonfocal events, and the story line had to be clear and authentic. The broadcasting organization also reviewed the content and provided input, but editorial control remained entirely with the series producers and editors. Thus, the Airline FoTW series offers robust quality and is relevant for our study.

Sampling

We sampled 111 interactions from the 138 total interactions derived from the series using several criteria. First, to ensure sufficient longitudinal data for the dynamic analysis, we selected interactions with duration of at least 3 minutes, which excluded 12 interactions. Second, prior research has indicated that a mix of long and close-up shots is needed to observe nonverbal cues, which requires at least 25% content dedicated to close-up and long shots; this resulted in the loss of another 12 interactions. Third, using a cutoff threshold of 60% for a content focus on customer–FLE communications, we obtained 111 usable interactions. We set aside 9 interactions as a test sample for grounded research, including building and validating a dictionary of verbal and nonverbal cues related to the study constructs. The remaining 102 interactions served as the analysis sample for hypothesis testing. The test sample did not differ from the analysis sample in length (t = 1.42, $p > .10$) or number of episodes per interaction (t = .83, $p > .10$).

For the dynamic analysis, we used a segment as the unit of analysis. A segment is a slice of each problem-solving interaction, spliced at naturally occurring breaks in the events. For our data, each segment was 20–60 seconds in duration, and each interaction comprised 2–5 segments, with time-specific tags to capture their sequential order. Ambady and Rosenthal (1992) indicate that 20-second slices are sufficient to draw conclusions about displayed behaviors. Studies of nonverbal cues require sampling at a lower order of analysis (i.e., thin slices) that occur for very brief periods (1–5 seconds). Coding nonverbal cues requires precise codes of facial, hand, and bodily movements that can change quickly in a 20-second segment. Therefore, we spliced each segment into 1–9 thin slices of 5–10 seconds in duration. To capture the fluidity of nonverbal cues, we also included 2 seconds of content before and after each thin slice. Thus, our usable sample of 102 interactions resulted in 373 segments and 803 (991) thin slices for FLE displayed affect (CSAT) assessments (Web Appendix A).

Measurement Libraries

Video recordings customarily are coded separately for audio (verbal cues) and visual (nonverbal cues) content, using dictionaries that correspond to the specific constructs of interest (Hill, White, and Wallace 2014). Validated dictionaries of verbal and nonverbal cue representations are available for a wide range of conceptual phenomena, such as the Harvard Inquirer, a dictionary of 11,788 words commonly used in English and categorized in 26 macro and 182 micro categories; Whissell's (2009) Revised Dictionary of Affect in Language, which categorizes 8,000 English words into positive or negative valence; and Ekman and Friesen's (2003) Facial Action Coding System for categorizing facial expressions into action units that indicate specific emotional states. Although these general use dictionaries often lack contextual relevance, they are useful as starting points for contextual refinement and development, which is how we deploy them in this study.

For our dictionary development process, we separate each segment into two components: audio without video for verbal cues (FLE solving and relational work) and visual without audio for nonverbal cues (CSAT and FLE displayed affect). For the verbal cues, we use existing dictionaries and our test sample to develop, refine, and validate the dictionary of words that correspond to solving and relational work. These dictionaries then support an automated extraction of measures for each slice of the problem-solving interaction in the analysis sample (see Figure 1). Before automation, we examine the face, convergent, and discriminant validity of the measurement dictionaries for verbal cues.

For nonverbal cues, the process accommodates video features, such as repeated uses of zooms, pans, close-ups, cutaways, and other video-journalistic styles that aim to engage the audience and capture authentic emotions/events. This makes approaches that require relatively fixed video capture (e.g., the Facial Action Coding System) less relevant. Nonverbal coding should represent how nonverbal cues are interpreted by the observers in the context in which they appear, so a grounded approach is needed to mimic this interpretation. We devised such an approach and relied on human coders to provide construct measures for each thin slice of the analysis sample.

FLE solving work. We initially reviewed the Harvard Inquirer library to identify relevant microcategories associated with "knowing," "assessing," "problem-solving," "interpersonal interaction," and "work" to develop an initial set of 3,305 words for use in customer problem-solving contexts. To ground the list, we asked two domain experts to sort the words in terms of their meaningfulness for solving work (definition provided). This step reduced the dictionary to 620 words after three iterations (interrater reliability = .83). We supplemented these words with an inductive refinement. Using the test sample of nine interactions, we generated 65 frequently used (>5 times) words by FLEs to communicate solving work, then cross-compared them with our dictionary to obtain 29 additional words, resulting in an updated solving work dictionary of 649 words. Two research assistants classified each word into one of two dimensions identified from a grounded analysis (interrater reliability = .86 after three iterations): (1) 315 "competence" words, indicating FLE skill and expertise to comprehend, analyze, and communicate information related to problem solving (usually adjectives and conjunctions: why, when, what, while, because) and (2) 334 "action" words, indicating FLE effort and engagement in finding solutions (usually verbs; e.g., "go," "do," "offer," "transfer," "send," "investigate," "provide").

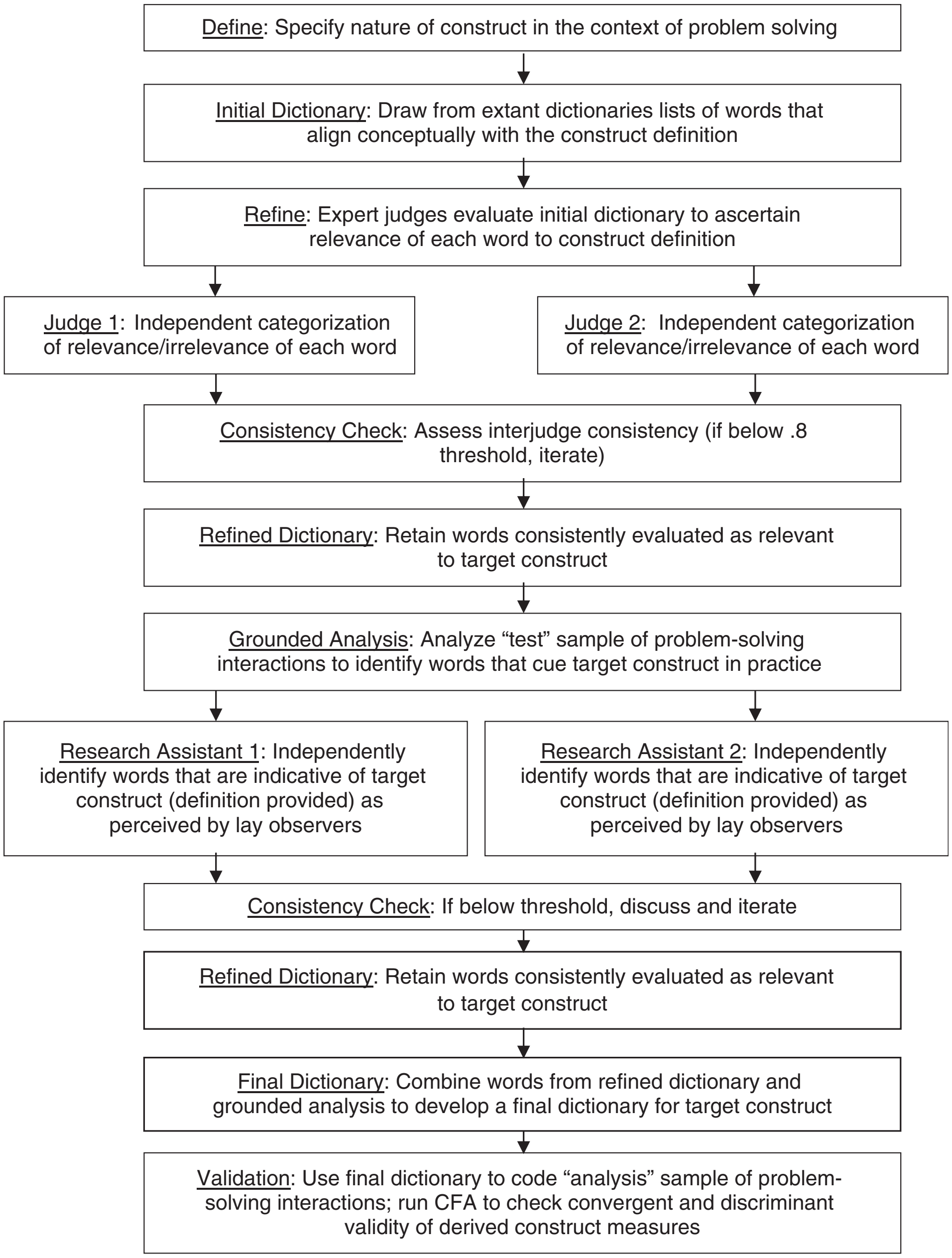
Next, we accounted for the cue strength intensity in the individual dictionary words. Some words such as "investigate" and "because" offer stronger cues of solving work than words such as "send" or "while." We developed a coding scheme by asking respondents (219 undergraduate students from a large Midwestern U.S. university) to rate each word on a 1–3 scale (1 = "low intensity," and 3 = "high intensity"), in terms of their everyday use in service interactions. The scores for each word (≥10) were averaged and divided by the standard deviation (SD) across respondents to arrive at a weighted intensity score.[3]

To operationalize solving work dimensions, we multiplied the occurrence (frequency = 0/1) of each competence and action word by its weighted intensity score (1–3) and obtained a score for any given segment of the problem-solving interaction (per the transcribed audio content). To account for varying segment and interaction length, we normalized the scores by dividing the time the FLE took to communicate the sentence (using time stamps) and obtained a weighted solving work measure (Web Appendix B, Tables B1.1 and B2).

FLE relational work. Relational work involves expressions of compassion and agreeableness to strengthen relationship bonds with customers. A common feature of these words is their approach or avoidance meaning for recipients. Whissell's (2009) Revised Dictionary of Affect in Language provides our starting point. Not all 8,000 words in this dictionary are relevant to problem solving. Using a procedure similar to the one used for the solving dictionary, we identified 244 relational words with acceptable consistency (interrater reliability = .88) and supplemented this dictionary with 20 words we obtained from an inductive analysis of words that raters judged as indicative of FLE relational work in the test sample. Two research assistants classified each word in the relational dictionary into two dimensions (interrater reliability = .89, after two iterations): (1) 88

[3]Because 8% of the solving words and 16% of the relational words had SDs of 0, we added 1 to all SDs to avoid dividing by 0. Therefore, words with 0 SD earn a score equivalent to the mean score, and the denominator exceeds 1, resulting in a weighted word intensity measure ranging from 0 to 3. Shah, Kumar, and Kim (2014) also use this approach.

Figure 1
PROCEDURES FOR VALIDATING DISCTIONARIES FOR FLE SOLVING AND RELATIONAL WORK



"agreeable" words, indicating FLE expressions of a good nature, courtesy, respect, helpfulness, and cooperativeness, often including adjectives, interjections, and verbs (e.g., "yeah," "agree," "calm," "help," "hear"), and (2) 176 "compassion" words that indicate expressions of kindness, tenderness, empathy, warmth, sympathy, and caring (Goetz, Keltner, and Simon-Thomas 2010) that include adverbs, adjectives, interjections, and verbs (e.g., "apologize," "sorry," "regret," "appreciate," "love," "hello"). Finally, we extracted the relational work measures for the analysis sample by multiplying the frequency of each relational word in each segment of the analysis sample (1 = present) by its weighted intensity score (1–3 scale) (per 219 respondents, with ≥10 ratings per word) and normalizing the score by the time-to-verbalize measure (Web Appendix B, Tables B1.2 and B2).

FLE displayed affect. For the nonverbal cues of FLE displayed affect, we used the test sample to develop coding rules, including identifying the valence and salience of each nonverbal cue in each thin slice (positive/neutral/negative) and isolating the cue source as facial (i.e., smiling, raising eyebrows, head shaking), bodily (i.e., distance and posture), or gestural (i.e., touching, tapping, and waving). This advances extant service research, which has largely focused on isolated or single nonverbal cues (e.g., type of smile) (Grandey et al. 2005; Wang et al. 2013). Two expert judges viewed thin slices from the test sample to identify 20 specific nonverbal cues associated with FLE feeling states (7 positive, 13 negative). They rated these slices for valence (1 = "extremely negative," and 7 = "extremely positive"), as well as for salience, by allocating 100 points across the salient nonverbal cue categories according to their significance (face, body, or hand gestures). We refined this procedure for clarity and consistency until we achieved acceptable interjudge reliability (.95). Then we trained six research assistants to code the thin slices from the analysis sample for FLE displayed affect. The interrater reliability was .92 (Web Appendix B, Tables B1.3 and B3).

CSAT. Consistent with Day (1983), we operationalize CSAT as an emotional response, manifested in customers' feeling states of positive fulfillment in situations involving dissatisfaction responses. Affective measures of CSAT are relevant to problem-solving experiences, because they disrupt usage experiences and degrade hedonic qualities, resulting in emotionally charged experiences (Oliver 1993; Westbrook 1981). Affective responses also are salient and diagnostic in conditions of cognitive constraints and time pressure, uncertain outcomes, and information asymmetry. Because nonverbal cues offer more authentic measures of affective states than do self-reports (Leigh and Summers 2002), they provide a reliable assessment of CSAT in problem-solving interactions. To develop nonverbal cues to measure CSAT, we used procedures parallel to those for FLE displayed affect. Six research assistants coded customers' nonverbal cues from thin slices in the test and analysis samples (interrater reliability = .93 training and .95 final coding).[4]

Control variables. We detail the control variables in Web Appendix C.

Hypothesis Testing Model

In the nested panel structure of the data, sequentially time-ordered segments (ST) are nested within problem-solving interactions. Both CSAT and its drivers (FLEs' work and affect) are segment specific, and the latter are hypothesized to have time-dependent (dynamic) effects. Therefore, we employ a random parameters model (Greene 2012), as follows:

$$
\begin{aligned}
(1)\quad \mathrm{CSAT}_{jkt} = {} & \beta_0 + \beta_{1kt}\mathrm{ST}_{jkt} + \beta_2\mathrm{SOLVING}_{jkt} \\
& + \beta_3\mathrm{RELATION}_{jkt} + \beta_4\mathrm{AFFECT}_{jkt} \\
& + \beta_5\mathrm{ST}_{jkt} \times \mathrm{SOLVING}_{jkt} \\
& + \beta_6\mathrm{ST}_{jkt} \times \mathrm{RELATION}_{jkt} \\
& + \beta_7\mathrm{ST}_{jkt} \times \mathrm{AFFECT}_{jkt} \\
& + \beta_8\mathrm{SOLVING}_{jkt} \times \mathrm{RELATION}_{jkt} \\
& + \beta_9\mathrm{SOLVING}_{jkt} \times \mathrm{AFFECT}_{jkt} \\
& + \beta_{10}\mathrm{RELATION}_{jkt} \times \mathrm{AFFECT}_{jkt} \\
& + \beta_{11}\mathrm{ST}_{jkt} \times \mathrm{SOLVING}_{jkt} \times \mathrm{RELATION}_{jkt} \\
& + \beta_{12}\mathrm{ST}_{jkt} \times \mathrm{SOLVING}_{jkt} \times \mathrm{AFFECT}_{jkt} \\
& + \beta_{13}\mathrm{CUSG}_j + \beta_{14}\mathrm{CUSR}_j + \beta_{15}\mathrm{CUSA}_j \\
& + \beta_{16}\mathrm{CSAT}_{jk(t-1)} + \varepsilon_{jkt},
\end{aligned}
$$

where $\varepsilon_{jkt} \sim \mathrm{iid}(0, \sigma^2)$.

$$
\begin{aligned}
(2)\quad \beta_{1kt} = {} & \alpha_0 + \alpha_1\mathrm{EMPG}_k + \alpha_2\mathrm{EMPR}_k \\
& + \alpha_3\mathrm{EMPA}_k + \alpha_4\mathrm{EMPD}_k + \zeta_{kt},
\end{aligned}
$$

where $\zeta_{kt} \sim N(0, \sigma^2)$.

In these equations, t = time, j = customer, and k = FLE; ST = segment/time for collecting repeated measures (from 2 to 5); SOLVING = FLE solving work, RELATION = FLE relational work, AFFECT = FLE displayed affect, CUSG/EMPG = customer/employee gender (0 = female, 1 = male), CUSR/EMPR = customer/employee race (0 = Caucasian, 1 = other), CUSA/EMPA = customer/employee age (0 = less than 30 years, 1 = more than 30 years), and EMPD = employee dress (0 = commonly dressed, 1 = well-dressed).

Endogeneity. The FLE–customer interaction yields temporally ordered and contemporaneous measures of the study variables. Typical dynamic panel data models, such as the Arellano–Bond specification, are not appropriate, because they require the presence of time-varying exogenous variables, which our data and research setting do not provide. Thus, to address endogeneity, we included a lagged dependent variable in our model to control for state dependence and also employed instruments (Germann, Ebbes, and Grewal 2015) (see Web Appendix D).

Multicollinearity. Relational and solving work correlate at .64. We regressed solving work on relational work, saving the residual, and then used the residual as an instrument for relational work in the hypothesis testing model (Cronbach and Furby 1970). The variance inflation factors are uniformly less than 5 (range = 1.46–5.18) (Neter, Wasserman, and Kutner 1989).

Results

Measure validity. With a confirmatory factor analysis (CFA), we examine the convergent and discriminant validity of the FLE solving and relational work measures (Web Appendix E). The CFA model produced statistics with acceptable fit ($\chi^2 = 3.49$, d.f. = 1, $p < .06$, confirmatory fit index [CFI] = .99, Tucker–Lewis index [TLI] = .97, root mean square error of approximation [RMSEA] = .08, $p > .05$). In support of convergent validity, the composite reliabilities for the solving and relational constructs are .85 and .75, respectively, and loadings are high (>.55) and significant ($p < .001$). Their average variance extracted (AVE) values are .77 and .64, respectively, which exceeds the shared variance of .48, indicating discriminant validity. Finally, we obtain factor scores for the solving and relational work constructs using the Bentler–Yuan optimal generalized least squares estimation. We also checked the expected pattern of interaction progression from sensing to seeking to settling activities (Web Appendix F).

Consistent with research that suggests senders use nonverbal cues uniquely (Aviezer, Trope, and Todorov 2012), we computed composite measures for FLE displayed affect (and CSAT) using unweighted combinations of facial, bodily, and gestural cue measures. Because FLE displayed affect and CSAT use common nonverbal cues, discriminant validity is a concern. However, the measures are not collinear (variance inflation factor < 2), sharing less than 12% of their variance. In addition, to test that FLE displayed affect precedes CSAT, we examine the interactive effect of FLE displayed affect and segment (time) on CSAT and find it to be significant (.11, $p < .1$), in support of the nomological validity of the two measures.

[4]We assessed cognitive (Maxham and Netemeyer 2002) and affective (Westbrook 1981) CSAT at the end of the interaction, using existing scales. These measures correlate at .76, indicating evidence of consistency.

Finally, we report the descriptive statistics and intercorrelations for the field study measures in Web Appendix G.

Model fit. We test different functional forms for ε_{ijk} in Equation 1 to identify the best-fitting model. Using the Akaike information criterion (AIC) to compare nonnested models, we identify normal and logistic pdf as the best-fitting parametric forms; the logistic specification outperforms the normal one (AIC = 837.4 vs. AIC = 838.1). We also compare the hypothesized model against a model that contains only control variables (Web Appendix H). The likelihood ratio test shows that the hypothesized model offers superior fit over the controls-only model ($\chi^2(12) = 335.66$, $p < .001$), a finding confirmed by the lower AIC for the hypothesized model (AIC = 837.4) compared with the control model (AIC = 1,149.1).

Hypotheses tests. As Table 1 shows, FLE solving work has a positive and significant impact on CSAT (.13, $p < .05$). According to a Wald test, the impact of solving work on CSAT increases steadily from .11 ($p > .29$) at the beginning (segment 1) to .37 ($p < .05$; segment 3) in the middle, and to .64 ($p < .03$) at the end of the interaction (segment 5), in support of H_1. In addition, FLE relational work negatively and significantly interacts with solving work and the segment (–.09, $p < .05$). We follow Spiller et al. (2013) and assess the impact of solving work on CSAT using a range (–2 SD to +2 SD) of relational work and segments (1–5). The results in Figure 2, Panel A, show that relational work significantly diminishes the influence of solving work on CSAT, all else being equal. In the beginning of the interaction (segment 1), the impact of solving work on CSAT decreases from .28 ($p < .05$) when relational work is low (–2 SD) to .10 ($p > .29$) at the mean to –.07 ($p > .62$) when relational work is high (+2 SD). Then, in the middle and end of the FLE–customer interaction, the patterns are similar, such that the effects diminish from .91 ($p < .02$) and 1.54 ($p < .01$) when relational work is low, to .37 ($p < .05$) and .64 ($p < .03$) when relational work is average, and finally to –.16 ($p > .65$) and –.25 ($p > .64$) when relational work is high. Thus, solving work has a positive effect on CSAT when relational work is low (<.1 SD) but is nonsignificant when relational work is greater than .1 SD. In support of H_2, FLEs' relational work interferes with the perceived efficacy of their solving work.

In addition, the influence of FLE solving work on CSAT is negatively and significantly moderated by FLE displayed affect over time (–.13, $p < .03$). We again use Wald tests over a range (–2 SD to +2 SD) of displayed affect and segments (1–5), all else being equal. In Panel B of Figure 2, we show that at the beginning of the interaction, the impact of solving work on CSAT decreases from .37 ($p < .02$) at low displayed affect (–2 SD), to .10 ($p > .29$) at the mean, and to –.15 ($p > .29$) at

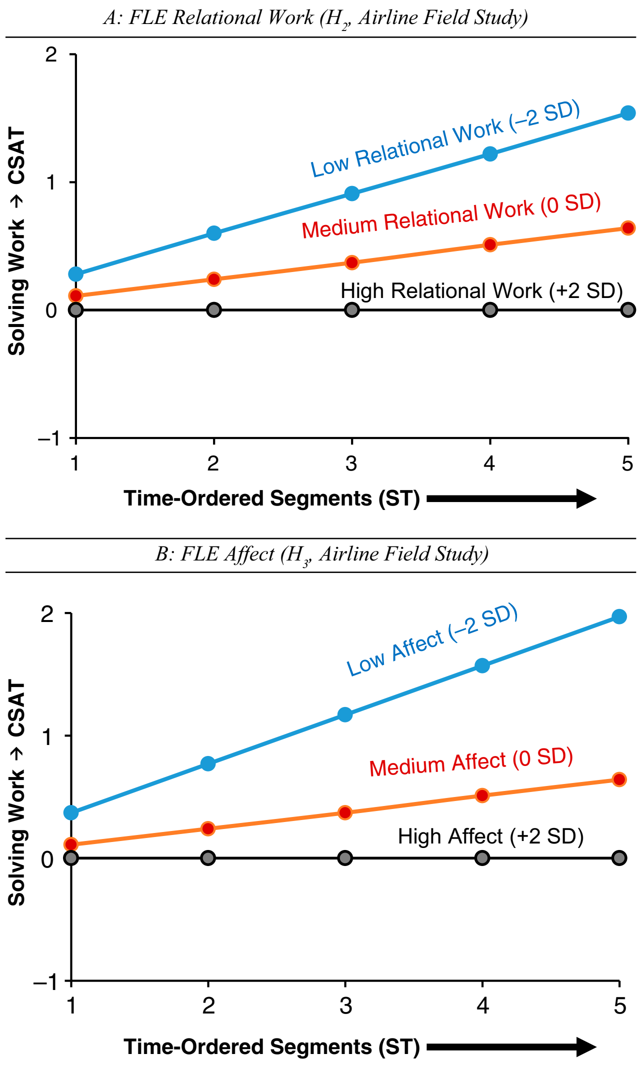
Table 1
MODEL ESTIMATION RESULTS FOR AIRLINE FIELD STUDY

Dependent Variable: CSAT	*Hypothesized Model*[a]	*Model with 90 Interaction*[b]	*Problem Severity*[c]	*Remove Highly (Un)pleasant Words*[d]
Intercept	1.15 (.47)*	1.13 (.41)**	1.53 (.49)**	1.13 (.46)**
Solving work	–.03 (.09)	–.01 (.09)	–.02 (.09)	–.02 (.09)
Relational work	–.07 (.09)	–.12 (.10)	–.08 (.09)	–.08 (.09)
Displayed affect	.25 (.12)*	.19 (.12)	.23 (.11)*	.24 (.11)*
Solving work × ST	**.13 (.05)***	**.10 (.04)***	**.14 (.05)****	**.13 (.06)***
Relational work × ST	.01 (.05)	.11 (.06)	.01 (.05)	.03 (.05)
Displayed affect × ST	.03 (.07)	.04 (.06)	–.04 (.08)	.03 (.07)
Solving work × Relational work	–.01 (.08)	.05 (.06)	.01 (.08)	.01 (.08)
Solving work × Displayed affect	–.07 (.06)	–.05 (.07)	–.07 (.06)	–.07 (.06)
Relational work × Displayed affect	.10 (.07)	.09 (.08)	.09 (.07)	.09 (.08)
Solving work × Relational work × ST	**–.09 (.04)***	**–.10 (.06)**	**–.08 (.04)***	**–.10 (.05)***
Solving work × Displayed affect × ST	**–.13 (.05)***	**–.16 (.06)****	**–.14 (.05)****	**–.12 (.05)***
ST	–.64 (.16)***	–.58 (.16)***	–.57 (.15)***	–.62 (.16)***
Problem severity			–.06 (.02)**	
Problem severity × Solving work × ST			–.03 (.02)	
Problem severity × Relational work × ST			–.03 (.02)	
Problem severity × Displayed affect × ST			–.01 (.02)	
Lag CSAT	.27 (.12)*	.29 (.10)**	.25 (.12)*	.27 (.13)*
Customer gender	.03 (.13)	.01 (.14)	.07 (.13)	.04 (.13)
Customer race	.14 (.15)	.04 (.17)	.14 (.14)	.15 (.15)
Customer age	–.11 (.13)	–.23 (.13)	–.13 (.13)	–.13 (.14)
Employee gender × ST	.05 (.09)	–.02 (.08)	.02 (.08)	.03 (.08)
Employee race × ST	–.70 (.28)*	–.66 (.28)*	–.78 (.27)**	–.71 (.27)*
Employee age × ST	–.14 (.09)	–.07 (.09)	–.15 (.09)	–.13 (.09)
Employee dress × ST	–.21 (.10)*	–.18 (.10)	–.27 (.10)**	–.22 (.10)*
AIC	837.40	750.00	841.10	837.20
Log-likelihood (d.f.)	–395.72 (23)	–352.01 (23)	–393.55 (27)	–395.58 (23)

*$p < .05$.
**$p < .01$.
***$p < .001$.
[a]Hypothesized model with 102 interactions.
[b]Robustness check model that contains only 90 interactions of more than 80% video content focusing on customer–FLE interactions.
[c]Robustness check model that controls for problem severity.
[d]Robustness check model in which FLE relationship work adjusts for overly pleasant or unpleasant words.
Notes: Two-tailed tests of significance. Boldfaced cells indicate hypothesized effects.

Figure 2

EFFECT OF FLE SOLVING WORK ON CSAT AT DIFFERENT LEVELS OF AFFECT



high displayed affect (+2 SD). In the middle of the interaction, the pattern is even stronger, such that the influence of solving work diminishes from 1.17 ($p < .01$) when displayed affect is low, to .37 ($p < .05$) when displayed affect is average, and to –.42 ($p > .27$) when displayed affect is high. At the end of the interaction, the decrements go from 1.97 ($p < .01$) to .64 ($p < .03$) to –.68 ($p > .28$). Solving work thus has a positive and significant influence on CSAT when displayed affect is low (<.1 SD) but a nonsignificant effect when it is greater than .1 SD. In support of H_3, FLEs' displays of positive affect diminish the perceived efficacy of solving work.

Robustness checks. To evaluate the sensitivity of these results, we compared the obtained parameter estimates with those from alternative specifications that (1) include interactions where 80% (vs. 60%) of video content focuses on customer–FLE communication (N = 90 interactions, column 2, Table 1), (2) control for problem severity (column 3, Table 1), and (3) adjust FLE relational work for overly pleasant or unpleasant words (column 4, Table 1). In all cases, the statistical inferences about the hypothesized effects remain unchanged.

STUDY 2: AIRLINE EXPERIMENTAL STUDY

We examine key findings from the airline field study in an experimental study that allows us to control potential extraneous causal factors. Three research questions guide Study 2's design. First, does FLE relational work negatively moderate the relationship between FLE solving work and CSAT? The field study findings indicate that high levels of relational work are counterproductive when FLEs solve problems under time pressure. Prior research has indicated an unconditional and beneficial effect of FLE relational work in problem-solving situations (i.e., "more is better") but has not examined this assertion with a dynamic analysis of problem-solving interactions. Despite support from the proposed contrast effects, we seek to test this result by controlling for alternative explanations. Second, does FLE relational work increase CSAT? Our field work challenges the notion that FLE relational work is beneficial, while giving prominence to FLE solving work in problem-solving interactions. We test this by comparing the effects of FLE solving and relational work on CSAT in a controlled setting. Third, after we control for the resolution of the service problem, does FLE solving work exert a positive effect on CSAT? That is, we recognize that the service problem resolution implies the choice of a single option in the specific conditions that define a situation. As an outcome, service problem resolution does not necessarily reflect the process for generating solution options. Our intuition from Study 1 is that FLE efforts to generate and present multiple viable solution options for customer selection is key to increased CSAT. With this experimental study, we test this question directly.

We designed a 2 (high vs. low) solving work × 2 (high vs. low) relational work × 2 problem context[5] (missed flight [MF] vs. lost baggage [LB]) between-subjects experiment (Web Appendix I). Missed flights and lost baggage are common consumer problems in airline travel (www.transportation.gov/airconsumer/), as confirmed by our field data. For the stimuli, we orthogonally manipulated FLE relational and solving work across four scenarios, using the dictionaries from Study 1. Several factors were held constant: (1) number of customer and FLE interaction turns, (2) content and number of words used by the customer, (3) number of words (but not content) used by the FLE, (4) use of automated voices[6] for the customer and FLE, (5) pictorial image of the customer and FLE interacting (extracted from video data), and (6) problem situation. The scenario designs also ensured identical customer outcomes, to avoid confounds due to varying problem resolutions. All scenarios were pretested with 101 respondents.

Sample and Measures

We recruited 568 participants ($M_{age} = 46.1$ years, SD = 13.25; 56.8% women) from an online panel of the U.S. population (>20 years in age) who had flown for business or pleasure in the last two years. Each participant was randomly assigned to

[5]We thank the associate editor and an anonymous reviewer for this suggestion.

[6]We use automated female voices from NeoSpeech (http://neospeech.com/), a text-to-speech provider. The customer and agent voices had distinct tones and tempo. Gender was nonsignificant in Study 1, so we were agnostic about the gender of the voices while recording the stimulus.

one of eight scenarios, with 275 participants in the LB context (M_{age} = 44.5 years, SD = 12.80; 52.7% women) and 293 participants in the MF context (M_{age} = 47.4 years, SD = 13.53; 60.8% women). Each participant listened to a brief audio file embedded in an airline problem-solving setting and provided responses from a consumer's perspective, on the following measures: (1) CSAT using a three-item, seven-point semantic differential scale anchored by "very displeased/very pleased," "very unhappy/very happy," and "terrible/delighted"(Westbrook 1981)[7]; (2) solving (Sirdeshmukh, Singh, and Sabol, 2002) and relational (Mattila and Enz 2002) work using two separate three-item, seven-point "strongly agree/strongly disagree" Likert scales; and (3) problem severity using a two-item, ten-point semantic differential scale anchored by "not at all distressing/ highly distressing" and "not at all stressful/highly stressful" (Maxham and Netemeyer 2002; see Web Appendix J).

The CFA of all multi-item constructs for both pooled and individual data produced reasonable fit statistics for the lost baggage (χ^2 = 427.2, d.f. = 202, $p < .001$; CFI = .95, TLI = .95, RMSEA = .06; 90% confidence interval [CI] = [.05, .07]; pclose = .05) and missed flight (χ^2 = 571.9, d.f. = 224, $p < .001$; CFI = .94, TLI = .93, RMSEA = .07; 90% CI = [.06, .08], pclose = .07) contexts. The composite reliabilities for all the constructs were consistently high (≥.7, $p < .001$), and all the AVEs exceeded .5. In support of discriminant validity, the AVE values also were greater than the shared variance between any pair of constructs. We extracted factor scores for each construct to use in our subsequent analysis.

Data and Manipulation Checks

We checked participants' airline travel experience, problem severity, and scenario realism. On average, participants had traveled by airline 2.98 times in the LF scenario and 2.67 times in the MF context in the previous two years. They evaluated problem severity after reading a description of each problem context but before being exposed to the experimental stimuli. Both contexts prompted above average problem severity perceptions (M_{LB} = 8.52, SD = 1.70; M_{MF} = 8.05, SD = 1.92; $p < .001$). In terms of scenario realism, participants indicated that both problem contexts were above average in realism (M_{LB} = 8.57, SD = 1.57; M_{MF} = 8.45, SD = 1.92, $p < .001$; 1 = "unrealistic," and 10 = "realistic"). Finally, manipulation checks showed that participants believed that the high–solving work scenarios (M_{LB} = 5.25, M_{MF} = 5.69) indicated greater solving work than the low conditions (M_{LB} = 4.37, $p < .001$; M_{MF} = 4.98, $p < .001$). Similarly, the high–relational work scenarios (M_{LB} = 5.79, M_{MF} = 5.88) indicated more relational work than the low conditions (M_{LB} = 5.03, $p < .001$; M_{MF} = 5.47, $p < .05$) (Web Appendix K).

Results

With CSAT as the dependent variable, our results show that experimentally manipulated FLE solving and relational work (and their interaction) explain significant incremental variance in CSAT beyond the effect of the control variables, including problem context, problem severity, age, gender, education, and airline travel frequency ($F(11, 556) = 10$, $p < .01$). Furthermore, in both severity contexts, FLE solving work has a significant simple effect (LB = .78, $p < .001$; MF = .94, $p < .001$), whereas FLE relational work has a nonsignificant simple effect (LB = .18, $p > .14$; MF = .07, $p > .69$) but a significant negative interaction effect with FLE solving work (LB = –.56, $p < .01$; MF = –.50, $p < .02$). The Wald test (Spiller et al. 2013) shows that solving work exerts a positive effect on CSAT when relational work is low (LB = .78, $p < .001$; MF = .94, $p < .001$), but when relational work is high, it has a nonsignificant effect in the LB context (.21, $p > .17$) and a significant but highly attenuated effect in the MF context (.44, $p < .01$). The negative moderating effect of FLE relational work affirms our first research question.

To address the second research question, we examine the influence of FLE relational work at different levels of solving work. Using Wald tests, we find that FLE relational work has a statistically nonsignificant effect on CSAT at low solving work (LB = .18, $p > .21$; MF = .07, $p > .69$) but a significant negative effect at high solving work (LB = –.39, $p < .01$; MF = –.43, $p < .01$). Thus, relational work is not responsible for an increase in CSAT, but it decreases CSAT when solving work is high, providing further support of the proposed contrast effects.

We also obtain the reported effects of FLE relational and solving work after controlling for the problem resolution (outcome). In both LB and MF contexts, customers choose the same outcome; neither the positive direct effect of solving work nor the negative moderating effect of relational work stems from differences in problem solutions or outcomes. Finally, we show that the findings are robust to variations in perceived problem severity.

GENERAL DISCUSSION

Unlike research that focuses on service recovery either before customers voice problems to seek resolution or after organizational resolution efforts to examine service recovery effectiveness, this study examines the dynamic influence of frontline work on CSAT *during* problem-solving interactions. We predicted a positive effect of FLE solving work on CSAT but also anticipated attenuating effects of FLE relational work and displayed affect. With two studies, featuring longitudinal panel data of real-life problem-solving interactions as well as a causal analysis using experimental data, we demonstrate that FLE solving work exerts increasingly positive effects on CSAT over the course of the interaction, but this influence is neutralized when FLEs display verbal cues that indicate high relational work or nonverbal cues that signal high positive affect. A distinct feature of this study is that we depart from previous research to extract theoretically well-grounded concepts of frontline work from the observed verbal and nonverbal cues FLEs display during problem-solving interactions, rather than from self-reported or experimentally manipulated data.

Limitations

Our study contains several limitations. We consider an airline setting, so further research might investigate disparate contexts with heterogeneous problem severity conditions. We do not include intonation cues, because there was not FLE routine interaction voice available to establish a baseline (Mayew and Venkatachalam 2012). Other methodologies also may be available for extracting FLE work and affect measures, such as machine learning techniques. Customer outcomes, such as loyalty, might help broaden our study's insights as well.

[7]We also included a three-item, seven-point Likert scale for an alternative measure of CSAT as a robustness check. We thank an anonymous reviewer for this suggestion.

Additional experimental manipulations might incorporate video content or more complex designs to examine the underlying processes and their boundary conditions. For example, Study 1 implies that FLE relational work has a marginally significant, positive effect at low levels of solving work but increasingly negative influences on CSAT at high levels of solving work, as the interaction evolves. A detailed examination of these dynamic effects would be a fruitful avenue for research. Finally, our study conceives of higher-order constructs of FLE solving and relational work to examine their time-varying, dynamic effects on CSAT during a problem-solving interaction, rather than the finer-grained constructs used by prior research to examine static (before or after) effects. A challenge for further research thus will be to resolve these trade-offs to produce dynamic analyses with fine-grained constructs.

Theoretical Implications

Frontline problem-solving constructs for process studies. This research conceptualizes and develops empirically validated dictionaries of novel frontline constructs that can inform process studies of problem-solving interactions. Solving work, relational work, and displayed affect offer theoretically useful, empirically distinct constructs that are conceptually grounded in the verbal cues communicated or the nonverbal cues displayed by FLEs during interactions. These constructs extend prior research examining either verbal (e.g., Ma and Dubé 2011) or nonverbal (e.g., Mattila and Enz 2002) cues in isolation when assessing service interactions (Puccinelli, Motyka, and Grewal 2010). The current research is novel, in that it provides frontline constructs for the simultaneous analysis of both verbal and nonverbal cues. Our grounded efforts to operationalize these focal constructs and establish empirically validated dictionaries of verbal and nonverbal cues support their application in further studies of frontline problem solving, which should facilitate consistent conceptualizations and operationalizations of the key constructs. In particular, in Study 2 we show that the validated dictionaries support robust manipulations of frontline problem-solving constructs and achieve discriminant validity. Thus, they offer a reasonable foundation for studying how frontline problem solving can lead to effective CSAT outcomes, which may be useful for training FLEs.

Effective problem solving: high solving work, low relational work, and displayed positive affect. By moving beyond a static effect (Smith, Bolton, and Wagner 1999), our dynamic analysis reveals that the influence of FLE solving work on CSAT grows six-fold in magnitude over time (Study 1: from .10 during sensing, $p > .29$, to .64 during settling, $p < .03$) when relational work and displayed affect remain constant and at average levels. This increasing influence of frontline solving work indicates that customers vigilantly monitor solving efficacy and impose severe punishments if the solving work continues to be ineffective in later phases of a problem-solving interaction. To estimate the potential financial penalty, we draw on Knox and Van Oest's (2014) calculations that the average value of service recovery to a firm for new and existing customers is \$36.50, given average purchases of \$57.32 per year. Extrapolating to the airline context, in which the average ticket price in the airports we studied was \$273 at the time of our investigation (U.S. Department of Transportation; www.rita.dot.gov), the average service recovery value may be as much as \$175.20 per customer. Service recovery requires an increase of 3.12 in the CSAT (from 2.92 at the beginning to 6.04 at the end of the interaction, on average, in our study). In our data, a 1 SD (16.6%) increase in FLEs' solving work can produce such service recovery, in that it lifts CSAT by 3.16 points in the settling phase, when relational work and displayed affect are low. If a FLE handles 20 customer problems each day, a low 25% effectiveness rate would imply a \$2,628 daily loss. If an airline has 200 problem-solving agents working approximately 20 days per month, the loss due to ineffective solving work would be \$10.5 million each month.[8] Training and technological aids thus need to help FLEs improve their solving work; even a minor improvement would mitigate these losses.

Customers also discount solving work accompanied by a high level of relational work. This discounting effect is nontrivial and consistent across our studies. In Study 2, designed to isolate causal effects, the significant and substantial influence of solving work on CSAT in both problem contexts at low relational work levels (.94, $p < .001$; .78, $p < .001$) decreases or becomes nonsignificant at high levels (.44, $p < .01$; .21, $p > .17$). In the field Study 1, we find that this discounting effect increases with time. In the early (sensing) phase of the interaction, the influence of solving work on CSAT is discounted for high (relative to low)[9] relational work (from .28 to 0); during later (seeking/settling) phases, the effect is five times higher (from 1.54 to 0), consistent with predictions of contrast effects. This consistent support across studies suggests that scholars and practitioners must reconsider their conventional beliefs about the role of relational behaviors for solving problems under time pressure. For example, Delta Airlines "gives its agents freedom to be chatty and personal" when solving customer problems (McCartney 2014), but our study suggests such freedom is risky, especially if FLEs take it as a recommendation to engage in high relational work while solving problems. Such actions will tend to backfire in practice.

Our finding that a negative moderating effect appears at above-average levels of FLE relational work has parallels in studies of the competence and warmth dimensions of human behavior (Abele and Wojciszke 2014). According to Kirmani et al. (2017), when consumers choose service providers, they prefer competent rather than moral (or high-warmth) providers. These extant findings from an adjacent literature stream are consistent with our study, but our research offers several unique contributions. Specifically, because we study the interactive and dynamic effects of solving and relational behaviors on CSAT during problem-solving interactions, we can establish not only that customers give prominence to solving over relational work, consistent with Kirmani et al.'s (2017) findings, but also that customers significantly discount the effect of solving work on CSAT in the presence of relational work. This discounting effect is not trivial; it is sufficient to neutralize, or even reverse, the positive effect of solving work. The discounting effect of relational behaviors (warmth) on the positive influence of solving behaviors (competence) on

[8]Our loss computation represents first-order effects and does not account for higher-order losses that result from word of mouth, network, and social propagation of ineffective problem solving. The recent widespread public exposure of ineffective problem solving by several leading airline service providers indicates that the total losses can be significantly higher by an order of magnitude.

[9]Low relational work (e.g., –2 SD) is not *negative* relational work; it means positive relational work at a low level.

CSAT during service interactions has not been examined before, to the best of our knowledge.

Finally, FLE displays of positive affect attenuate the impact of their solving work in the early phase (from .37 to 0), and this attenuation increases by a factor of five later in the interaction (from 1.97 to 0). Low levels of displayed affect during the sensing state seem sufficient to pacify customers and gain information to facilitate solving work. Subdued positive displayed affect during the seeking and settling phases promotes customers' positive assessment of problem-solving effectiveness. These findings parallel Sutton and Rafaeli's (1988) finding that customers "sanction" FLEs who display positive emotions during busy periods in retail checkout counters not because positive emotions are bad but because their display wastes time. On this point, our findings contrast with those reported by Grandey et al. (2005), who find in a slow-paced context that FLE smiles (nonverbal cue) enhance the effect of FLE task performance on postencounter CSAT. However, their study does not measure employees' actual behavior or its dynamic effect on CSAT during the interaction as the current study does.

CSAT in problem-solving interactions. Conceptually, CSAT has been defined as a postconsumption or post-interaction outcome, based on customers' evaluations of the degree to which the totality of the interaction meets or disconfirms expectations (Anderson and Sullivan 1993; Mittal and Kamakura 2001; Oliver 2010). Variations in the level of CSAT during the interaction or how CSAT changes as it unfolds over time usually are neglected (Verhoef, Antonides, and De Hoog 2004). We conceptualize CSAT on the basis of nonverbal cues that capture customers' evolving affective responses to FLE behaviors during problem-solving interactions. Further research might address the theoretical mechanisms that explain these changes in CSAT during problem-solving interactions and alter the direction of the interaction outcomes.

Managerial Implications

Many companies strive to improve CSAT and routinely record problem-solving interactions in call centers for qualitative reviews or individual FLE training. However, few use these data to derive generalizable insights for practice improvement. This research provides a practical way for companies to analyze such recorded data.

For companies that seek effective problem-solving approaches (e.g., McCartney 2014), our studies also offer compelling recommendations. When problem solving under time pressure, FLEs' solving work is critical to increasing CSAT—more important than their relational work or positive displayed affect, especially in later phases of the interaction, such that it can yield nontrivial rates of return (e.g., 1 SD increase in solving work = $10.5 million per month for 200 agents currently with a 25% effective solving rate). Moreover, we provide a library of validated dictionaries that managers can use for cue-based training of FLEs; these dictionaries also might feed into automated, technology-enabled systems for dynamic, live FLE assistance interfaces.

Effective solving work is best not confused with problem-solving outcome. The former pertains to FLE competence and action to generate options for customers. The latter refers to a single choice of a solution, as negotiated between the FLE and customer. Customers credit frontline efforts that increase the quality and quantity of solution options presented to them, regardless of the solution choice. Customers also discredit frontline efforts that appear to deviate from their expectations for problem-solving interactions. As our findings show, high relational work and overly pleasant affective displays during solving work create contrast effects, so that customers perceive that FLEs are distracted and discount the effectiveness of their solving work. The continued use of relational work and pleasant affect displays in later phases even prompts penalties, regardless of how competent or action-oriented the FLE solving work might be. To ensure effective problem solving, managers must realize that when it comes to relational work and displays of pleasant affect, sometimes less is more. Our study thus calls for shifts in the frame, conception, and practice of frontline problem solving: from service to problem-solving, from static to dynamic, and from "more is better" to "less is more."

REFERENCES

Abele, Andrea E., and Bogdan Wojciszke (2014), "Communal and Agentic Content in Social Cognition: A Dual Perspective Model," in *Advances in Experimental Social Psychology*, James M. Olson and Mark P. Zanna, eds. Waltham, MA: Elsevier, 195–255.

Ambady, Nalini, and Robert Rosenthal (1992), "Thin Slices of Expressive Behavior as Predictors of Interpersonal Consequences: A Meta-Analysis," *Psychological Bulletin*, 111 (2), 256–74.

Anderson, Eugene W., and Mary W. Sullivan (1993), "The Antecedents and Consequences of Customer Satisfaction for Firms," *Marketing Science*, 12 (2), 125–43.

Aviezer, Hillel, Yaacov Trope, and Alexander Todorov (2012), "Body Cues, Not Facial Expressions, Discriminate Between Intense Positive and Negative Emotions," *Science*, 338 (6111), 1225–29.

Biddle, Bruce J. (1986), "Recent Development in Role Theory," *Annual Review of Sociology*, 12, 67–92.

Bitner, Mary Jo, Bernard H. Booms, and Lois A. Mohr (1994), "Critical Service Encounters: The Employee's Viewpoint," *Journal of Marketing*, 58 (4), 95–106.

Bonoma, Thomas V., and Leonard C. Felder (1977), "Nonverbal Communication in Marketing: Toward a Communicational Analysis," *Journal of Marketing Research*, 14 (May), 169–80.

Bradley, Graham L., Beverley A. Sparks, Dieter Zapf, Janet R. McColl-Kennedy, and Nerina L. Jimmieson (2013), "Task-Relationship-Self: A Framework for Understanding Service Encounter Behaviors," *Psychology and Marketing*, 30 (6), 512–28.

Carver, Charles S., and Michael F. Scheier (1990), "Origins and Functions of Positive and Negative Affect: A Control-Process View," *Psychological Review*, 97 (1), 19–35.

Cronbach, Lee J., and Lita Furby (1970), "How We Should Measure 'Change': Or Should We?" *Psychological Bulletin*, 74 (1), 68–80.

Day, George S. (1983), "The Next Step: Commonly Accepted Constructs for Satisfaction Research," in *International Fare in Consumer Satisfaction and Behavior*, R.L. Day and K.H. Hunts, eds. Bloomington: Indiana University Press.

DeChurch, Leslie A., Jessica R. Mesmer-Magnus, and Dan Doty (2013), "Moving Beyond Relationship and Task Conflict: Toward a Process-State Perspective," *Journal of Applied Psychology*, 98 (4), 559–78.

Doyle, Aaron (1998), "'Cops': Television Policing as Policing Reality," in *Entertaining Crime: Television Reality Programs*, M. Fishman and G. Cavender, eds. New York: Walter de Gruyter, 95–116.

Echeverri, Per (2005), "Video-Based Methodology: Capturing Real-Time Perceptions of Customer Processes," *International Journal of Service Industry Management*, 16 (2), 199–209.

Ekman, Paul, and Wallace V. Friesen (2003), *Unmasking the Face: A Guide to Recognizing Emotions from Facial Clues*. Los Altos, CA: Ishk.

Fang, Zheng, Xueming, Luo, and Minghua, Jiang (2013), "Quantifying the Dynamic Effects of Service Recovery on CSAT: Evidence from Chinese Mobile Phone Markets," *Journal of Service Research*, 16 (3), 341–55.

Fishbach, Ayelet, and Stacey R. Finkelstein (2012), "How Feedback Influences Persistence, Disengagement, and Change in Goal Pursuit," in *Goal-Directed Behavior*, I.H. Aarts and A.J. Elliot, eds. New York: Psychology Press.

Fitzsimons, Gráinne M., and Eli J. Finkel (2011), "Outsourcing Self-Regulation," *Psychological Science*, 22 (3), 369–75.

Gelbrich, Katja, and Holger Roschk (2011), "Do Complainants Appreciate Overcompensation? A Meta-Analysis on the Effect of Simple Compensation vs. Overcompensation on Post-Complaint Satisfaction," *Marketing Letters*, 22 (1), 31–47.

Germann, Frank, Peter Ebbes, and Rajdeep Grewal (2015), "The Chief Marketing Officer Matters!" *Journal of Marketing*, 79 (3), 1–22.

Goetz, Jennifer L, Dacher Keltner, and Emiliana Simon-Thomas (2010), "Compassion: An Evolutionary Analysis and Empirical Review," *Psychological Bulletin*, 136 (3), 351–74.

Grandey, Alicia A., Glenda M. Fisk, Anna S. Mattila, Karen J. Jansen, and Lori A. Sideman (2005), "Is 'Service with a Smile' Enough? Authenticity of Positive Displays During Service Encounters," *Organizational Behavior and Human Decision Processes*, 96 (1), 38–55.

Greene, William H. (2012), *Econometric Analysis*. Englewood Cliffs, NJ: Pearson Education.

Groth, Markus, and Alicia Grandey (2012), "From Bad to Worse: Negative Exchange Spirals in Employee–Customer Service Interactions," *Organizational Psychology Review*, 2 (3), 208–33.

Heritage, John, and Douglas W. Maynard (2006), "Problems and Prospects in the Study of Physician-Patient Interaction: 30 Years of Research," *Annual Review of Sociology*, 32, 351–74.

Hill, Aaron D., Margaret A. White, and J. Craig Wallace (2014), "Unobtrusive Measurement of Psychological Constructs in Organizational Research," *Organizational Psychology Review*, 4 (2), 148–74.

Homburg, Christian, Michael Müller, and Martin Klarmann (2011), "When Should the Customer Really Be King? On the Optimum Level of Salesperson Customer Orientation in Sales Encounters," *Journal of Marketing*, 75 (2), 55–74.

Jones, Candace, and Reut Livne-Tarandach (2008), "Designing a Frame: Rhetorical Strategies of Architects," *Journal of Organizational Behavior*, 29 (8), 1075–99.

Kelley, Scott W., and Mark A. Davis (1994), "Antecedents to Customer Expectations for Service Recovery," *Journal of the Academy of Marketing Science*, 22 (1), 52–61.

Kirmani, Amna, Rebecca W. Hamilton, Debora V. Thompson, and Shannon Lantzy (2017), "Doing Well Versus Doing Good: The Differential Effect of Underdog Positioning on Moral and Competent Service Providers," *Journal of Marketing*, 81 (1), 103–17.

Knox, George, and Rutger Van Oest (2014), "Customer Complaints and Recovery Effectiveness: A Customer Base Approach," *Journal of Marketing*, 78 (5), 42–57.

Leigh, Thomas W., and John O. Summers (2002), "An Initial Evaluation of Industrial Buyers' Impressions of Salespersons' Nonverbal Cues," *Journal of Personal Selling & Sales Management*, 22 (1), 41–53.

Liao, Hui (2007), "Do It Right This Time: The Role of Employee Service Recovery Performance in Customer-Perceived Justice and Customer Loyalty After Service Failures," *Journal of Applied Psychology*, 92 (2), 475–89.

Ma, Zhenfeng, and Laurette Dubé (2011), "Process and Outcome Interdependency in Frontline Service Encounters," *Journal of Marketing*, 75 (3), 83–98.

Mattila, Anna S., and Cathy A. Enz (2002), "The Role of Emotions in Service Encounters," *Journal of Service Research*, 4 (4), 268–77.

Maxham, James G., III, and Richard G. Netemeyer (2002), "A Longitudinal Study of Complaining Customers' Evaluations of Multiple Service Failures and Recovery Efforts," *Journal of Marketing*, 66 (4), 57–71.

Mayew, William J., and Mohan Venkatachalam (2012), "The Power of Voice: Managerial Affective States and Future Firm Performance," *Journal of Finance*, 67 (1), 1–43.

McCartney, Scott (2014), "The Art of Airline Apology," *The Wall Street Journal* (July 9), http://www.wsj.com/articles/the-art-of-the-airline-apology-1404947542.

Menon, Kalyani, and Laurette Dubé (2007), "The Effect of Emotional Provider Support on Angry Versus Anxious Consumers," *International Journal of Research in Marketing*, 24 (3), 268–75.

Mittal, Vikas, and Carly Frennea (2010), *Customer Satisfaction: A Strategic Review and Guidelines for Managers*. Cambridge, MA: Marketing Science Institute.

Mittal, Vikas, and Wagner A. Kamakura (2001), "Satisfaction, Repurchase Intent, and Repurchase Behavior: Investigating the Moderating Effect of Customer Characteristics," *Journal of Marketing Research*, 38 (1), 131–42.

Nabi, Robin L., Erica N. Biely, Sara J. Morgan, and Carmen R. Stitt (2003), "Reality-Based Television Programming and Psychology of Its Appeal," *Media Psychology*, 5 (4), 303–30.

Neter, John, William Wasserman, and Michael H. Kutner (1989), *Applied Linear Regression Models*. Homewood, IL: Richard D. Irwin Inc.

Okhuysen, Gerardo A., David Lepak, Karen L. Ashcraft, Giuseppe J. Labianca, Vicki Smith, and H.K. Steensma (2013), "Theories of Work and Working Today," *Academy of Management Review*, 38 (4), 491–502.

Oliver, Richard L. (1993), "Cognitive, Affective, and Attribute Bases of the Satisfaction Response," *Journal of Consumer Research*, 20 (3), 418–30.

Oliver, Richard L. (2010), *Satisfaction: A Behavioral Perspective on the Consumer*. New York: ME Sharpe Inc.

Parasuraman, Anantharanthan, Valarie A. Zeithaml, and Leonard L. Berry (1988), "SERVQUAL: A Multiple-item Scale for Measuring Consumer Perceptions of Service Quality," *Journal of Retailing*, 64 (1), 12–40.

Paul, Michael, Thorsten Hennig-Thurau, and Markus Groth (2014), "Tightening or Loosening the 'Iron Cage'? The Impact of Formal and Informal Display Controls on Service Customers," *Journal of Business Research*, 68 (5), 1062–73.

Penner, Louis A., Heather Orom, Terrance L. Albrecht, Melissa M. Franks, Tanina S. Foster, and John C. Ruckdeschel (2007), "Camera-Related Behaviors During Video Recorded Medical Interactions," *Journal of Nonverbal Behavior*, 31 (2), 99–117.

Puccinelli, Nancy M., Scott Motyka, and Dhruv Grewal (2010), "Can You Trust a Customer's Expression? Insights into Nonverbal Communication in the Retail Context," *Psychology and Marketing*, 27 (10), 964–88.

Rafaeli, Anat, Daniel Altman, Dwayne D. Gremler, Ming-Hui Huang, Dhruv Grewal, Bala Iyer, A. Parasuraman, and Ko de Ruyter (2017), "The Future of Frontline Research: Invited Commentaries," *Journal of Service Research*, 20 (1), 91–99.

Rafaeli, Anat, and Robert I. Sutton (1987), "Expression of Emotion as Part of the Work Role," *Academy of Management Review*, 12 (1), 23–37.

Richins, Marsha L. (1983), "Negative Word-of-Mouth by Dissatisfied Consumers: A Pilot Study," *Journal of Marketing*, 47 (1), 68–78.

Ringberg, Torsten, Gaby Odekerken-Schröder, and Glenn L. Christensen (2007), "A Cultural Models Approach to Service Recovery," *Journal of Marketing*, 71 (3), 194–214.

Roschk, Holger, and Katja Gelbrich (2014), "Identifying Appropriate Compensation Types for Service Failures: A Meta-Analytic and Experimental Analysis," *Journal of Service Research*, 17 (2), 195–211.

Scott, Maura L., Martin Mende, and Lisa E. Bolton (2013), "Judging the Book by Its Cover? How Consumers Decode Conspicuous Consumption Cues in Buyer–Seller Relationships," *Journal of Marketing Research*, 50 (3), 334–47.

Shah, Denish, Vipin Kumar, and Kihyun Hannah Kim (2014), "Managing Customer Profits: The Power of Habits," *Journal of Marketing Research*, 51 (6), 726–41.

Sirdeshmukh, Deepak, Jagdip Singh, and Barry Sabol (2002), "Consumer Trust, Value, and Loyalty in Relational Exchanges," *Journal of Marketing*, 66 (1), 15–37.

Smith, Amy K., Ruth N. Bolton, and Janet Wagner (1999), "A Model of Customer Satisfaction with Service Encounters Involving Failure and Recovery," *Journal of Marketing Research*, 36 (4), 356–72.

Spector, Robert, and Patrick D. McCarthy (2005), *The Nordstrom Way to Customer Service Excellence: A Handbook for Implementing Great Service in Your Organization*. Hoboken, NJ: John Wiley & Sons.

Spiller, Stephen A., Gavan J. Fitzsimons, John G. Lynch Jr., and Gary H. McClelland (2013), "Spotlights, Floodlights, and the Magic Number Zero: Simple Effects Tests in Moderated Regression," *Journal of Marketing Research*, 50 (2), 277–88.

Stack, Liam (2017), "United Airlines Offers Refunds as Outrage at a Violent Removal Continues," *The New York Times* (April 12), https://www.nytimes.com/2017/04/12/business/united-dragged-passenger-video-chicago.html.

Stoller, Gary (2005), "Companies Give Front-Line Employees More Power," *USA Today* (June 26), https://usatoday30.usatoday.com/money/companies/management/2005-06-26-service-usat_x.htm.

Sturdy, Andrew, and Peter Fleming (2003), "Talk as Technique: A Critique of the Words and Deeds Distinction in the Diffusion of Customer Service Cultures in Call Centres," *Journal of Management Studies*, 40 (4), 753–73.

Sutton, Robert I., and Anat Rafaeli (1988), "Untangling the Relationship Between Displayed Emotions and Organizational Sales: The Case of Convenience Stores," *Academy of Management Journal*, 31 (3), 461–87.

Tax, Stephen S., Stephen W. Brown, and Murali Chandrashekaran (1998), "Customer Evaluations of Service Complaint Experiences: Implications for Relationship Marketing," *Journal of Marketing*, 62 (2), 60–76.

Verhoef, Peter C., Gerrit Antonides, and Arnoud N. de Hoog (2004), "Service Encounters as a Sequence of Events the Importance of Peak Experiences," *Journal of Service Research*, 7 (1), 53–64.

Wang, Ze, Aaron D. Arndt, Surendra N. Singh, Monica Biernat, and Fan Liu (2013), "'You Lost Me at Hello': How and When Accent-Based Biases Are Expressed and Suppressed," *International Journal of Research in Marketing*, 30 (2), 185–96.

Westbrook, Robert A. (1981), "Sources of Consumer Satisfaction with Retail Outlets," *Journal of Retailing*, 57 (3), 68–85.

Whissell, Cynthia (2009), "Using the Revised Dictionary of Affect in Language to Quantify the Emotional Undertones of Samples of Natural Language," *Psychological Reports*, 105 (2), 509–21.

Zeithaml, Valarie A., Leonard L. Berry, and A. Parasuraman (1996), "The Behavioral Consequences of Service Quality," *Journal of Marketing*, 60 (2), 31–46.