FRONTLINE PROBLEM-SOLVING EFFECTIVENESS: A DYNAMIC ANALYSIS OF VERBAL AND NONVERBAL CUES

Detelina Marinova
Frances Ridge Gay MBA Professor of Marketing
Robert J. Trulaske, Sr. College of Business, University of Missouri-Columbia
424 Cornell Hall, Columbia, MO - 65211
marinovad@missouri.edu

Sunil K. Singh
Assistant Professor of Marketing
College of Business, University of Nebraska-Lincoln
345E Hawks Hall, 730 N. Vine St., P.O. Box 880405, Lincoln, NE-68588
ssingh10@unl.edu

Jagdip Singh
AT&T Professor of Marketing
Weatherhead School of Management, Case Western Reserve University
Peter B Lewis Building 221
11119 Bellflower Road
Cleveland, OH-44106
Jagdip.singh@case.edu

FRONTLINE PROBLEM-SOLVING EFFECTIVENESS: A DYNAMIC ANALYSIS OF VERBAL AND NONVERBAL CUES

Abstract

This study examines the impact of frontline employees' problem solving on customer satisfaction (CSAT) *during* ongoing interactions prompted by service failures and complaints. Based on outsourced regulation theory, we predict negative moderating effects of frontline relational work and displayed affect on the dynamic influence of frontline solving work on CSAT. Frontline employee's verbal cues provide the basis for identifying solving and relational work, and nonverbal cues for identifying their displayed affect. We test hypotheses with data from video-recordings of real-life problem-solving interactions involving airline customers, as well as a controlled experimental study. We 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 sum, overdoing relational work and over-displaying positive affect diminish the efficacy of problem-solving interactions, which provides implications for theory and practice.

Keywords: problem solving, service recovery, complaint handling, frontline employees, dynamic, verbal cues, nonverbal cues, CSAT, solving, relational, affect

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 due to 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 consumers' dissatisfied 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 feature 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 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 recalibrated 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, enjoy consistently high customer satisfaction (CSAT) ratings (Anderson and Sullivan 1993; Mittal and Frennea 2010; Oliver 2010).

However, most research examines problem solving at the frontline by studying customers' response states either ***prior*** to the problem-solving effort, such as their causal attributions, emotions , expectations , or actions (e.g., complaint; Kelley and Davis 1994; Richins 1983; Ringberg, Odekerken-Schroder, and Christensen 2007), or ***subsequent*** to 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*—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 emphasizes the positive impact of frontline employee's relational work, including empathy, courtesy, and apology, on CSAT (Zeithaml, Berry, and Parasuraman 1996). But recent meta-analytic studies of service failure conclude that relational work is less helpful in failure situations that do not involve psychological loss (Roschk and Gelbrich 2014). Although positive affect may help sooth customers in distress, Rafaeli et al. (2017) find that it can be

[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%.

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 post-exchange 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 advancing 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 past studies overlook 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 get developed and negotiated under time pressure. Based 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 satisfactory. 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* (cf. 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.

To conceptualize FLE work, we first consulted service quality literature, in which scales such as SERVQUAL (Parasuraman, Zeithaml, and Berry 1988) are developed to capture customers' cumulative post-consumption experience, not *within*-interaction *FLE behaviors*. Some dimensions of SERVQUAL, such as reliability, are not relevant for studying problem-solving interactions since 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 process, rather than 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 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 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 accord with problem-solving progress, which typically involves three phases: *sensing* (e.g., problem comprehension), *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

Viewing problem-solving as a process motivated by goal pursuit, we predict that

[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.

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; unlike it though, 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 he or she 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 it exceeds 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 post-consumption evaluations of problem solving have been widely studied (e.g., 142 studies, two 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 that examines a vocabulary of effective problem-solving words.

Customer dissatisfaction (satisfaction) should grow if their assessment of observed verbal cues indicates that the FLE's 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, due to 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. But 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 customers' satisfaction. Thus:

> **H$_1$**: FLEs' solving work has a positive effect on customer satisfaction, and this effect increases in magnitude during the course of the problem-solving interaction.

Relational Work Moderates the Influence of Frontline Solving Work on CSAT

Prior studies recognize the positive role of relational work in frontline problem solving (Fang, Luo, and Jiang 2013; Smith, Bolton, and Wagner 1999). It cannot directly solve customer problems, but 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 problem solving. For example, "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, with 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 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. This emphasis on solving work in the settling phase is expected to crowd out the need and tolerance for FLE relational work. Thus:

> **H_2**: FLE 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 above 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 displays of positive affect in the seeking and settling phases are expected 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 cross-contextual generalizability, we need longitudinal, in-situ data about ongoing problem-solving interactions between FLEs and customers. Prior research advocates 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 fly-on-the-wall

(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, 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 non-focal 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 interaction with duration of at least 3 minutes, which excluded 12 interactions. Second, prior research indicates 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 cut-off 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 hypotheses 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 duration. 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 Enquirer, a dictionary of 11,788 words commonly used in English and categorized in 26 macro and 182 micro categories; Whissell's (2009) RDAL, which categorizes 8000 English words into positive or negative valence; and Ekman and Friesen's (2003) FACS 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.

Insert Figure 1 about here

For nonverbal cues, the process accommodates video features, such as repeated uses of zooms, pans, close-ups, cutaways, and other video-journalistic styles that seek to engage the audience and capture authentic emotions/events. This makes approaches that require relatively fixed video capture (e.g., FACS) 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 rely on human coders to provide construct measures for each thin slice of the analysis sample.

FLE solving work. We initially reviewed the Harvard Enquirer library to identify relevant micro-categories associated with "knowing," "assessing," "problem-solving," "interpersonal interaction," and "work" to develop an initial set of 3305 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 9 interactions, we generated 65 frequently used (>5 times) words by FLEs to communicate solving work, then cross-compared them with earlier 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 3 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: go, do, offer, transfer, send, investigate, and 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, 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 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

[3] Because 8% of the solving and 16% of the relational words had standard deviations (SD) 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.

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) RDAL provides our starting point. Not all 8000 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 2 iterations): (1) 88 "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 eye brows, head shaking), bodily (i.e., distance and posture), or gestural (i.e., touching, tapping, and waving). This advances extant service research, which largely focuses 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.

Hypotheses 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:

(1) $$\begin{aligned}CSAT_{jkt} = {} & \beta_0 + \beta_{1kt}ST_{jkt} + \beta_2\, SOLVING_{jkt} + \beta_3 RELATION_{jkt} + \beta_4 AFFECT_{jkt} + \beta_5 ST_{jkt} \times SOLVING_{jkt} + \beta_6 ST_{jkt} \times RELATION_{jkt} + \beta_7 ST_{jkt} \times AFFECT_{jkt} + \beta_8\, SOLVING_{jkt} \times RELATION_{jkt} \\ & + \beta_9\, SOLVING_{jkt} \times AFFECT_{jkt} + \beta_{10}\, RELATION_{jkt} \times AFFECT_{jkt} + \beta_{11} ST_{jkt} \times SOLVING_{jkt} \times RELATION_{jkt} + \beta_{12} ST_{jkt} \times SOLVING_{jkt} \times AFFECT_{jkt} + \beta_{13} CUSG_j + \beta_{14} CUSR_j + \beta_{15} CUSA_j + \beta_{16} CSAT_{jk(t\text{-}1)} + \varepsilon_{jkt},\end{aligned}$$

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

(2) $$\beta_{1kt} = \alpha_0 + \alpha_1 EMPG_k + \alpha_2 EMPR_k + \alpha_3 EMPA_k + \alpha_4 EMPD_k + \zeta_{kt},$$

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

[4]We assessed cognitive (Maxham and Netemeyer 2002) and affective (Westbrook 1980) CSAT at the *end* of the interaction, using existing scales. These measures correlate at .76, indicating evidence of consistency.

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 hypotheses testing model (Cronbach and Furby 1970). The VIF 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 ($\chi^2 = 3.49$, df = 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 GLS 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 (VIF < 2), sharing less than 12% of their variance. Also, 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.

Finally, we report the descriptive statistics and inter-correlations 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 AIC to compare non-nested models, we identify normal and logistic *pdf* as the best fitting parametric forms; the logistic specification outperforms the normal one (AIC = 837.4 versus 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 controls only model (χ^2 (12) = 335.66 $p < .001$), a finding confirmed by the lower AIC for the hypothesized (AIC = 837.4) compared with the control (AIC = 1149.1) model.

Insert Table 1 about here

Hypotheses tests. As shown in Table 1, 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), 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 (–2SD to +2SD) 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 (–2SD) to .10 (p >.29) at the mean to –.07 (p > .62) when relational work is high (+2SD). 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 (below .1SD) but non-significant when it is greater than .1SD. In support of H_2, FLEs' relational work interfere with the perceived efficacy of their solving work.

Insert Figure 2a-b about here

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 (–2SD to +2SD) 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 (–2SD), to .10 (p > .29) at the mean, and to -.15 (p > .29) at high displayed affect (+2SD). 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 (< .1SD) but a non-significant effect when it is greater than

.1SD. 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 against those from alternative specifications that (a) include interactions where 80% (vs. 60%) of video content focuses on customer–FLE communication (N = 90 interactions, column 2, Table 1), (b) control for problem severity (column 3, Table 1), and (c) 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 control over potential extraneous causal factors. Three research questions guide Study 2 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 indicates 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, does FLE solving work exert a positive effect on CSAT, after controlling for the resolution of the service problem? 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 multiple viable solution options and present them 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 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

[5] We thank the AE and an anonymous reviewer for this suggestion.
[6] We use automated female voices from Neo speech (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.

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 3-item, 7-point semantic differential scale anchored by "very displeased/very pleased," "very unhappy/very happy," and "terrible/delighted"(Westbrook 1980).[7] (2) Solving (Sirdeshmukh, Singh, and Sabol, 2002) and relational (Mattila and Enz 2002) work using two separate 3-item, 7-point, "strongly agree/strongly disagree" Likert scales, and (3) Problem severity using a 2-item, 10-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, *df* = 202, *p* < .001, CFI = .95, TLI = .95, RMSEA = .06; 90% confidence interval [CI] = [.05, .07]; PClose = .05) and missed flight contexts (χ^2 = 571.9, *df* = 224, *p* < .001, CFI = .94, TLI = .93, RMSEA = .07; 90% CI = [.06, .08], PClose = .07). The composite reliabilities for all the constructs were consistently high (≥ .7, *p* < .001), and all the AVE 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 average problem severity perceptions (M_{LB} = 8.52, SD = 1.70; M_{MF} = 8.05,

[7] We also included a 3-item, 7-point Likert scale for an alternative measure of CSAT as a robustness check. We thank an anonymous reviewer for this suggestion.

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, 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 non-significant 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 non-significant 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 gets 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 past research to extract theoretically well-grounded concepts of frontline work from the observed verbal and nonverbal cues frontline employees 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. Other customer outcomes, such as loyalty, might help broaden our study's insights too. 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 past research to examine their 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 that examines 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, with 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, where the average ticket price in the airports we studied was $273 at the time of our investigation (US DoT, 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

1SD (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 \$2628 daily loss. If an airline has 200 problem-solving agents, working about 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 non-significant 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 gets 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

[8] Our loss computation represents first order effects, and do 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., –2SD) is not *negative* relational work; it means positive relational work at a low level.

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 finds parallels in studies of the competence and warmth dimensions of human behavior (Abele and Wojcisze 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 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 on the positive influence of solving behaviors on 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 appear 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 check-out 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 post-encounter CSAT. However, their study does not measure employees' actual behavior or its dynamic effect on CSAT during the interaction as the current study does.

Customer satisfaction in problem-solving interactions. Conceptually, CSAT has been defined as a post-consumption or -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 based on non-verbal 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 customer satisfaction and routinely record problem-solving interactions in call centers for qualitative reviews or individual FLE training. Few use these data to derive generalizable insights for practice improvement though. 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., 1SD 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.

Table 1
MODEL ESTIMATION RESULTS FOR AIRLINE FIELD STUDY

Dependent Variable: CSAT	Hypothesized Model	Model with 90 Interactions	Problem Severity	Remove highly (un)pleasant. words
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 x ST	-.21 (.10)*	-.18 (.10)	-.27 (.10)**	-.22 (.10)*
Akaike information criterion	837.4	750.0	841.1	837.2
Log-likelihood (df)	-395.72 (23)	-352.01 (23)	-393.55 (27)	-395.58 (23)

*$p < .05$, **$p < .01$, ***$p < .001$ (two-tailed tests); Hypothesized. Model – Hypothesized Model with 102 interactions; Robustness Check Models: 90 interaction – Model contains only 90 interaction that have > 80% of video content focusing on customer–FLE, Problem Severity - Controlled for problem severity, Removing (un)pleasant words - FLE relational work adjusts for overly pleasant or unpleasant words.

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

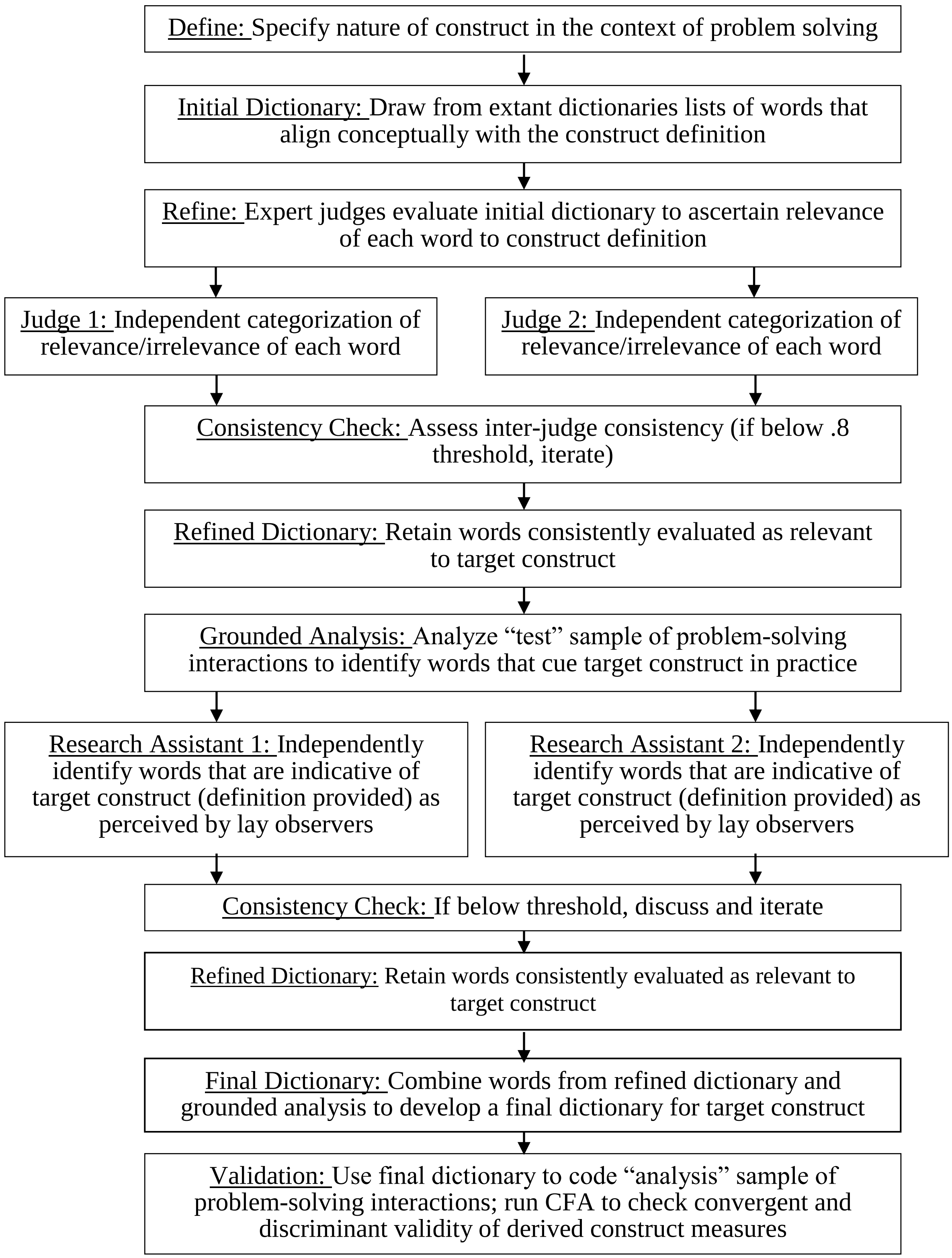
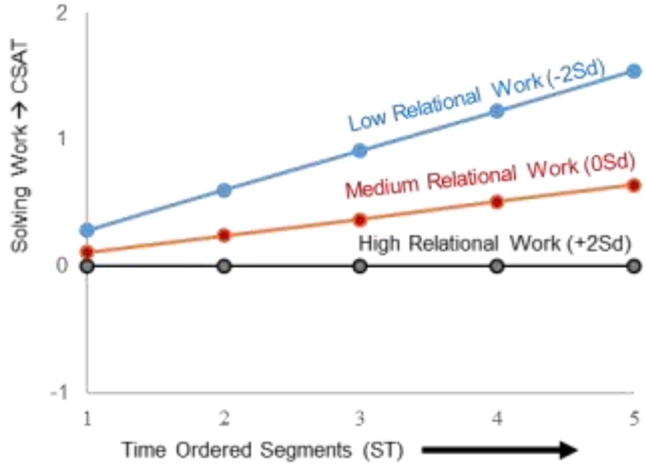


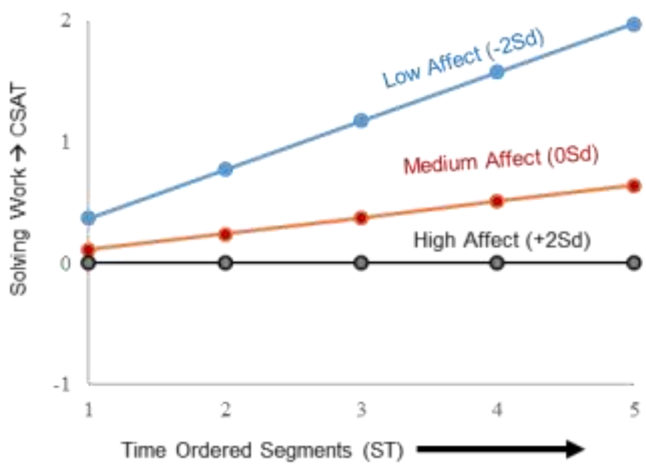
Figure 2

EFFECT OF FLE SOLVING WORK ON CSAT AT DIFFERENT LEVELS

a. FLE Relational Work (H_2, airline field study)



b. FLE Affect (H_3, airline field study)



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*, Olson, James M. and Mark P. Zanna, eds. Elsevier.

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), 1225-9.

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 (October), 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 & 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,* Day, R.L. 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,* Fishman, Mark and Gary Cavender, eds., New York:Walter de Gruyter.

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: 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*, Aarts I.H 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 (March), 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), 103-117.

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 (May), 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 III, James G. 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," (accessed March 8, 2015), [available at 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,” in *MSI Fast Forward Series*. 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 & 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), 334-347.

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 (August), 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 (April), 277-88.

Stack, Liam (2017), (accessed 05/11/2017), [available at https://www.nytimes.com/2017/04/12/business/united-dragged-passenger-video-chicago.html?_r=0]

Stoller, Gary (2005), "Companies Give Front-line Employees More Power," (accessed November 26, 2014), [available at http://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.

WEB APPENDIX A

Key Terms

Key Terms	Definition
Solving work	The competence and actions displayed by FLEs during problem-solving interactions that are indicative of FLEs' efforts to resolve customers' problems.
Relational work	The compassion and agreeableness displayed by FLEs during problem-solving interactions that are indicative of FLEs' efforts to foster relational bonds with customers.
Displayed affect	The facial, bodily, and gestural cues displayed by FLEs during problem-solving interactions that are indicative of their feeling states (positive/negative/neutral).
CSAT	The facial, bodily and gestural cues displayed by customers during problem-solving interactions that are indicative of their feeling state (positive/negative/neutral).
Problem-solving interaction	An encounter in which a customer communicates with a frontline employee to address a dissatisfaction, question, or concern related to the firm's product or service offerings. This study focuses on face-to-face encounters, but they also can be mediated by technology.
Segment	A section of a problem-solving interaction obtained by splicing at naturally occurring turn-taking events during an interaction. In the Airline study, problem-solving interactions typically involve 4–5 segments, each 20-60 seconds in duration.
Thin-slice	A section of a problem-solving segment that is sufficient to capture, accurately and meaningfully, nonverbal cues related to facial, bodily, or gestural expressions by the customer or FLE at any point in time. In our study, each segment is spliced into 1–9 thin slices of 5–10 seconds duration each.
Verbal cues	Audible words used in the communications between the FLE and customer.
Nonverbal cues	Facial expressions, bodily posture, and gestural displays used in the communications between the customer and FLE.
Test sample	A subset of problem-solving interactions randomly sampled from the full set of problem-solving interactions for grounded research purposes, to develop and validate measures that are contextually meaningful.
Analysis sample	The remaining subset of problem-solving interactions (full set minus test sample) used to test the hypotheses.

WEB APPENDIX B
Table B1.1
Coding Procedures for FLE Solving Work

Interaction Segment {Interaction #_Segment #} [Start time – End Time]	**Solving Dictionary Words** (Int*) Competence	Action	**Solving Work Score =** ∑(Frequency×Intensity) of solving words/(Time taken in seconds)
{41_2} ***Customer***: [1:27 – 1: 35] The child is less than 2 years and we provided the information when buying the tickets. Why are you not letting us through now? ***FLE***: [1:37 – 1:57] Sir, I ***understand***, but we ***need*** a birth certificate for the child. If you don't ***have*** it then you have two ***choices*** – ***buy*** a new ticket now and ***go*** but ***do send*** us the boarding pass and I ***guarantee*** that we will ***send*** the refund or we can ***call*** the hospital and ***resolve*** the age issue now.	Guarantee (2.9) Resolve (2.8) Understand (2.6) Issue (2.5)	Buy (2.7) Call (2.6) Go (2.5) Send (2.4) Do (2.2) Need (2) Have (1.5)	(30.6)/(20) = **1.53**
{77_3}***Customer*:** [1:55 – 2:01] I need to get to San Diego today. Please do something. ***FLE*:** [2:02 – 2:21] If you definitely ***need*** to ***get*** to San Diego today then only way it is ***possible*** is if I ***put*** you on a flight to Kansas and then ***get*** you a ***transfer*** at Phoenix, though you will ***reach*** home only by Mid-night. If you ***prefer doing*** so, then you ***have*** to ***spend*** an additional $59.	Prefer (2.2) Possible (1.2)	Get (2.4) Spend (2.2) Do (2.2) Put (2.1) Transfer (2.1) Reach (2) Need (2) Have (1.5)	(22.3)/(19) = **1.17**
{23_3}***Customer*:** [1:21 – 1:33] Can you check with the pilot or whoever if I can go. I don't understand why I can't go when the plane is standing there. ***FLE*:** [1:34 – 1:53] I ***got*** off the phone with the gate agent. He ***said*** that he won't ***allow*** any more passengers onboard as the load sheets are already ***submitted***. And once that is ***done***, it is not ***possible*** for anyone to board the aircraft. You can rebook for the next available flight.	Allow (2.2) Possible (1.2)	Got (2.4) Submit (2.3) Said (2.3) Done (1.7)	(12.1)/(19) = **.64**
{12_1}***Customer*:** [0:39 – 0:47] I don't have my passport. But I have my library card that has my picture on it. ***FLE***: [0:48 – 1:08] It is not ***possible*** to board the flight without the passport. Only ***chance*** you ***have*** of travelling today is if you ***get*** your passport and are ***here*** in next two hours. If you are ***late*** then you will ***miss*** the flight. However, I am not sure if I can ***keep*** a seat for you till then.	Chance (1.4) Miss (1.3) Late (1.3) Possible (1.2)	Get (2.4) Have (1.5) Here (1)	(10.1)/(20) = **.50**

*Int refers to the intensity of solving dictionary words, based on ratings obtained from 219 undergraduates (1= low, 3 = high).

Table B1.2
Coding Procedures for FLE Relational Work

Interaction _Segment {Interaction #_Segment #} [Start time – End Time]	**Relational Dictionary Words (**Int*) (pleasant/unpleasant)* Agreeable	Compassion	**Relational Work Score =** ∑ (Frequency×Intensity) of relational words / (Time taken in seconds)
{34_4} ***Customer***: [2:13 – 2: 21] Am I going out today? Please say yes. ***FLE***: [2:22 – 2:30] ***Yes Sir***, you are going on the 7:30 flight. I am ***glad*** it worked out. It is ***great*** to be with family on holidays.	Yes (2.9)	Glad (3) Great (3) Sir (2.6)	(11.44)/(8) = **1.44**
{59_2}***Customer*****:** [0:45 – 0:54] Where is my luggage. How can you miss it? ***FLE*****:** [0:55 – 1:05] ***Sir***, ***no*** need to ***shout***. ***Please calm*** down. I understand that you have a sick passenger, and I am trying to get her medication box as the top ***priority***.	No (1.3) Calm (2.4) Shout (1.4)	Please (2.7) Sir (2.6) Priority (2.4)	(12.8)/(10) = **1.28**
{11_4}***Customer*****:** [2:49 – 2:53] I am sorry for the smell. ***FLE*****:** [2:54 – 3:04] ***No*** need to ***apologize***. We were ***only*** concerned about other passengers. However, this is a $12 voucher you ***can*** use for redeeming food/drinks that you ***can*** even take with you.	Can (2.6) Only (1.7) No (1.3)	Apologize (2.4)	(10.6)/(10) = **1.1**
{83_3}***Customer*****:** [1:31 – 1:41] This is not done. You need to get me my reimbursement for missing 1 day from my cruise. ***FLE***: [1:42 – 1:54] ***No***, I ***don't*** think it is possible. And ***please don't*** use ***curse*** words. We know you are ***upset*** but it is not our ***mistake*** that the flight is delayed.	No (1.3) Don't (1.1)	Please (2.7) Upset (1.4) Mistake (1.4) Curse (1.3)	(10.7)/(12) = **.89**

*The relational dictionary word intensity refers to degree of pleasantness/unpleasantness (1= unpleasant, 3 = pleasant), based on ratings from 219 undergraduates.

Table B1.3
Coding Procedures for FLE Displayed Affect

Nonverbal Cues {Interaction No._Segment No.}	**Hand Gesture**	**Body**	**Face**	**Displayed Affect Score = ∑ (Frequency × Intensity) of Nonverbal Cues/# of Nonverbal Cues**
{18_2}				
Nonverbal Cue	Touch (Self)	Posture (Closed)	Gaze (Avert)	
Cue Intensity*	3	3	2	= (8)/3 = **2.67**
{79_4}				
Nonverbal Cue	Touch (Others)	Posture (Open)	Gaze (Maintain)	
Cue Intensity*	7	5	6	= (18)/3 = **6.0**

*Nonverbal cue intensity is measured on a 7-point scale (1 = extremely negative, 4 = neutral, 7 = extremely positive).

Table B2
Dictionary Words for Solving and Relational Work (Upper and Lower 30th Percentiles)

Solving Work	High Intensity*			Low Intensity*		
Competence	Choice	Show	Aware	Whether	How	Miss
	Decision	Support	Correct	Know	Depart	Why
	Understand	Damage	Always	Matter	Reason	Late
	Carry	Allow		Mean	What	Possible
	Guarantee	Issue		Tell	Because	Chance
	Resolve	Free		Away	Delay	
Action	Investigate	Order	Come	Wait	Therefore	
	Buy	Provide	Get	Ask		
	Call	Report	Hold	Release		
	Find	Pay	See	Bit		
	Speak	Verify	Send	Let		
	Leave	Go	Take	Pass		
	Make	Change	Talk	Quick		
Relational Work	**Pleasant ****			**Unpleasant****		
Agreeable	Believe	Alleviate	Help	Adamant	Don't	
	Respect	Care	Calm	Only	Whatever	
	Agree	Best	Will	Didn't	Only	
	Yes	Can	Opinion	Argue	Error	
	Polite	Definitely	Remember	No	Won't	
	Relax	Right		Can't	Shout	
Compassion	Glad	Welcome	Madam	Debate	Accuse	
	Great	Hope	Sir	Stop	Cry	
	Happy	Love	Pray	Upset	Hate	
	Nice	Sorry	Priority	Mistake	Rude	
	Smile	Please	Apology	Worry	Curse	
	Appreciate	Trust	Goodbye	Embarrass	Worse	
	Hello	Accommodate	Thank	Afraid		

*Intensity is measured on a 1–3 point scale. High intensity refers to ratings of 2.4–3 (upper 30%), and low intensity refers to ratings of 1–1.7 (lower 30%).
** Degree of agreeableness/compassion is measured on a three-point scale. Pleasant refers to ratings of 2.4–3 (upper 30%), and unpleasant refers to ratings of 1–1.7 (lower 30%).

Table B3
Prototypical Nonverbal Cues for Training Coders

FLE Displayed Affect	Prototypical Nonverbal Cues
Positive	Face – smile/laugh, maintain gaze, nod (in agreement) Body – lean forward/open posture, open arms Hand Gestures - Thumbs up, wave (good bye, take care)
Negative	Face – avert gaze (look down/ look away), frown, roll eyes, twitch lips, shake head sideways Body – lean backwards/closed posture, raise shoulders Hand Gestures - Fold hands, steeple hands, hit the countertop, put hands in pocket/hips, play with fingers, touch self (body, face)
CSAT	**Prototypical Nonverbal Cues**
Positive	Face – smile/laugh, maintain gaze, nod (in agreement) Body – lean forward/open posture, open arms Hand Gestures - Thumbs up, wave (good bye, take care)
Negative	Face – avert gaze (look down/ look away), frown, roll eyes, twitch lips, shake head sideways Body – lean backwards/closed posture, raise shoulders Hand Gestures - Fold hands, steeple hands, hit the countertop, put hands in pocket/hips, play with fingers, touch self (body, face)

WEB APPENDIX C

Controls (Airline Field Study)

To control for alternative predictors of FLE behavior, we measure and control for FLE demographic variables, including age, gender, and race (Khan, Chawla, and Devine 1996; Schepers et al. 2012). We operationalize the variables on dichotomous scales: customer/employee gender (0 = female, 1 = male), customer/employee race (0 = Caucasian, 1 = other), and customer/employee age (0 = less than 30 years, 1 = more than 30 years). Race is often operationalized as multi-category variable, but in our study setting, we aggregate all races other than Caucasian as "other" because the data are insufficient to model the effect of each race individually. Each variable was coded by six raters (interrater reliability ≥ .95). We also use employee dress as a control variable. Rafaeli and Pratt (1993) suggests that employees dress is a visual cue that likely affects customers' behavioral response. We operationalize employee dress as a dichotomous variable (0 = poorly dressed, 1 = well-dressed), with interrater reliability of .98.

Problem severity refers to the magnitude of service failure, harm, or inconvenience perceived by customers at the time they experience the service problem (McCollough, Berry, and Yadav 2000; Smith and Bolton 1998). Therefore, we assess perceived severity according to the expressions of customer dissatisfaction at the beginning of the problem-solving interaction. The greater the severity, the greater the anxiety, distress, or agitation expressed by the customer. Available video data allow us to capture customers' perceptions of the problem severity from their nonverbal expressions of experienced dissatisfaction, instead of verbal cues that customers may regulated to match social norms (Puccinelli, Motyka, and Grewal 2010). Seven judges independently reviewed each customer's nonverbal cues from the first slice of each problem-solving interaction, then recorded the degree of problem severity on a two-item, 10-point scale (Web Appendix J). Based on Hayes and Krippendorff (2007), the interrater reliability of problem severity coding is .92 and .93 for the two items; given this acceptable reliability, we compute a perceived problem severity score for each interaction (M = 6.83, SD = 2.15). Then we control for problem severity in the model estimation by including both its simple and interaction effects with all hypothesized variables. Table 1 (Column 3) summarizes the results. Although problem severity has a significant simple effect on CSAT (–.06, $p < .05$), none of the interaction terms involving problem severity achieve significance. In addition, the statistical inference about the hypothesized effects remains unchanged.

WEB APPENDIX D
Endogeneity Correction (Instrumental Variable Creation and Testing)

Steps	Description	Reference
1. Include lagged dependent variable in the model	To account for state dependence and time-varying unobserved effects, we include the lagged dependent variable in the model (eq. 1).	Germann, Ebbes, and Grewal, (2015); Rossi(2014) Jacobson (1990); Liu and Yang (2009); Mizik and Jacobson (2008); Fair (1970)
2. Create instruments	To create valid instruments for temporally ordered and contemporaneous measures of FLE work and displayed affect, we follow the guidance provided by prior research in marketing. The current value of a contemporaneous variable is regressed on its past values, lagged one period, as well as on customer satisfaction, lagged one period. The predicted scores then serve as instruments and satisfy the relevance and exclusion criteria; they correlate with the current values of the predictor variables that they preceded but are not influenced by contemporaneous unobservable variables in equation 2.	
3. Establish instrument validity	Sargan test for over-identifying restrictions to assess if the instruments created are uncorrelated with the residuals. $N \times R^2 = .313$, $\chi^2_{3df} = 6.25$ for $p < .1$. Thus, for our instruments, we cannot reject the null prediction that the instruments and residuals are uncorrelated.	Sotgiu and Gielens (2015); Wooldridge (2010)
4. Assess strength of the instruments	In a two-step process, we (1) regress the endogenous variable on all exogenous variables, and (2) add instruments to step 1 to perform an incremental F-test and assess their incremental explanatory power. Incremental F statistics greater than 10 suggest the instruments are strong. Incremental F-statistics (df = 12, 14) of 108.7 (FLE solving work), 91.30 (FLE relational work), and 65.22 (FLE displayed affect), all significant at $p < .001$.	Sotgiu and Gielens (2015); Wooldridge (2010)
5. Test for endogeneity in segment concurrent measures	Hausman test for endogeneity confirms that endogeneity exists ($\chi^2_{5df} = 17.2$, $p < .01$), and the segment-concurrent measures are inconsistent.	Greene (2012)

WEB APPENDIX E
Estimated Coefficients from Confirmatory Factor Analysis of Study Constructs (Airline Field Study, n = 373)

	Loading[a]	t-Value	Reliability[b]	Average Variance Extracted[c]	Maximum Variance Shared
Solving Work			.85	.77	.48
Competence	.69	13.84			
Action	.92	18.95			
Relational Work			.75	.64	.48
Agreeable	.90	16.56			
Compassion	.57	10.75			

[a]Estimated standardized coefficients with corresponding t-values in the adjacent column were obtained with a maximum likelihood solution.
[b]Estimated composite reliability per Fornell and Larcker (1981).
[c]Estimated variance extracted by the latent construct from its hypothesized indicators per Fornell and Larcker (1981).

WEB APPENDIX F

Interaction Phases in Airline Field Study

We examined the expected pattern of progression from sensing and seeking to settling activities in four steps. First, randomly select a subset of problem-solving interactions. Second, categorize the segments in each interaction into three groups, based on the time order of progression: G1 = segment 1, G2 = segments 2 and 3, and G3 = segments 4 and 5. Third, administer the individual segments (randomized across groups and interactions) to a customer sample to evaluate whether activities in a given segment correspond to sensing (S1) (e.g., problem articulation, Brashers, Goldsmith, and Hsieh 2002), seeking (S2) (e.g., solution development, Kellas and Trees 2006), or settling (S3) (e.g., decision, Kieren, Maguire, and Hurlbut 1996). Fourth, test the extent to which G1 segments are dominated by S1 activities, G2 by S2 activities, and G3 by S3 activities. In all, 12 problem-solving interactions were randomly selected to yield 47 segments, grouped into three categories, and evaluated by 15 respondents who were unaware of the study's purpose, as well as by segment grouping to provide judgments of observed activities. The respondents' judgments yielded an overall interrater reliability of 93.3%. For all categories, respondents' dominant assignment of observed activities was consistent with the expected pattern: 95% of G1 segments were evaluated as S1 activities, 94% of G2 segments as S2 activities, and 92% of G3 segments as S3 activities (all $p < .01$), in support of the proposed sensing–seeking–settling sequence of interaction phases during problem solving.

WEB APPENDIX G

Descriptive Statistics and Correlations of Constructs in Airline Field Study

	Variable	1	2	3	4	5	6	7	8	10	11	12	13	14
1	CSAT	1												
2	ST	.26**	1											
3	Displayed affect	.66***	.07	1										
4	Solving work	-.15**	.01	-.17**	1									
5	Relational work	-.05	.17*	-.13**	.64***	1								
6	Customer gender	.04	-.00	-.11	.04	.15**	1							
7	Customer race	-.02	.00	.10	-.02	.16**	.05	1						
8	Customer age	-.07	-.00	-.08	.06	.09	.18**	.12**	1					
9	Employee gender	.08	.03	.20**	.04	-.04	-.14*	.06	-.11	1				
10	Employee race	-.13**	-.05	-.17**	-.05	.01	.14**	.28***	-.01	-.11	1			
11	Employee age	-.06	.09	-.03	-.05	-.06	.07	.10	.07	-.03	-.11**	1		
12	Employee dress	-.19**	-.07	-.31***	.02	-.07	.04	.07	-.06	-.16**	.06	.10	1	
13	Problem severity	-.17*	.02	-.09	.05	.03	-.19**	-.02	-.08	.04	-.17**	.00	-.19**	1
	Mean	3.22	2.42	3.45	6.26	4.62	.61	.20	.48	.38	.14	.35	.61	6.83
	SD	1.33	1.16	1.16	4.96	4.75	.49	.40	.50	.49	.35	.47	.48	2.15

**p* < .05, ***p* < .01, ****p* < .001 (two-tailed tests).

WEB APPENDIX H

Controls Only Model Estimation Results

Dependent Variable: CSAT	Controls Only
Intercept	2.01 (.13)***
Solving work	
Relational work	
Displayed affect	
Solving work × ST	
Relational work × ST	
Displayed affect × ST	
Solving work x Relational work	
Solving work x Displayed affect	
Relational work x Displayed affect	
Solving work × Relational work × ST	
Solving work × Displayed affect × ST	
ST	-.82 (.12)***
Lag CSAT	
Customer gender	.42 (.10)***
Customer race	-.01 (.12)
Customer age	.06 (.10)
Employee gender × ST	.24 (.08)**
Employee race × ST	-1.31 (.46)**
Employee age × ST	-.44 (.12)**
Employee dress × ST	-.51 (.09)***
AIC	1149.1
Log-likelihood (df)	-563.55 (11)

*$p < .05$, **$p < .01$, ***$p < .001$ (two-tailed tests).

WEB APPENDIX I
Stimuli for Study 2 (Airline Experimental Study)

	Missed Flight [high solving, high relational in red; low solving, high relational in blue]	**Lost Baggage** [high solving, low relational in red; low solving, low relational in blue]
Agent	Hi! May I help you?	Hi! How are you doing?
Joanna	I have a ticket for the 9 am Miami flight. Can you check me in?	I came on the 6 am flight from New York via Atlanta, but my checked baggage, is not here.
Agent	I am so sorry, but I am afraid, that is not possible. You are late for check-in, and the gate is now closed. Unfortunately, your ticket is cancelled I apologize, for the inconvenience).	Let me check this right away. May I have your boarding pass, and baggage tag?
Joanna	You don't understand. There was a major backup on I-294. It is not my fault. I need to be in Miami, as I have an important meeting to attend.	Sure.
Agent	Sorry for the inconvenience. I would be upset too (I understand your urgency and see you are upset).	I see, that there was a weather related delay. Your baggage did not make the Atlanta flight due to insufficient connection time (and is not here at this time). (Were you late for the Atlanta flight?)
Joanna	I sure am. You can't do this.	That's unacceptable. I have a job interview at 1pm, and my baggage has all the materials.
Agent	I am so sorry, for the inconvenience. I can't get you on the 9 am flight, but I will check right away, and see what I can do, to get you to Miami today (I am so sorry for the inconvenience. I sympathize with you. I really wish, I could help you get on your Miami flight. But unfortunately, you will have to go on a later flight).	Let me see, what I can do, to get it here for you, as soon as possible… (You don't understand. Your bag is still in Atlanta, and not here. Weather related delays are unavoidable...)
Joanna	That is so unfair!	This is so unfair!
Agent	I am so sorry. I was in a similar situation once due to traffic… I am checking, how to get you to Miami (Missing flights, is the worst. I know, how you feel). Now, there is another direct flight at 3:30 pm.	I understand (You don't understand). Let me see, how to get your bag, here at the earliest.… (Your bag, is in Atlanta, and would not arrive in Miami, until 2:25 pm. This is when, the next flight from Atlanta, gets here). Okay, I have a few options. I can have your bag, on the next direct flight, at 2:25 pm, and delivered by 5:30 pm. (The bag, would be delivered to your address, by 5:30 pm.)
Joanna	What time does it get into Miami?	That won't work. I need my bag before my 1 pm interview.
Agent	5:42 pm	Ok. Once the bag arrives, I can expedite delivery, for a $25 fee, which I will waive, but you still won't get the bag, until 3:30 pm. If that doesn't work, I have some other options (We cannot do much, when the baggage delay is weather related. That is why, we advise

		passengers, to not pack their important materials in the checked baggage. Weather related delays, are not in our control).
Joanna	Well… That's not going to work. I will miss my meeting.	Why can't you get my bag on an earlier flight?
Agent	That is unpleasant (and stressful)…, We (really) don't want you to miss your meeting. Let me see. I could try, to get you on another carrier, for an 11 am flight that connects in Atlanta, and gets into Miami at 5:12 pm (And, we apologize, that you are in this situation. But, unfortunately, I cannot do anything else. The next flight is at 3:30 pm).	Yes, an earlier option I have, is an Atlanta-Houston connection, that will get your bag in Miami, by 1:47 pm. If I expedite, you will have it by 2:30 pm (That is not possible. As I explained to you earlier, there are no direct flights from Atlanta that arrive in Miami, before 2:25 pm. I can't do anything.).
Joanna	That doesn't help. I won't make my meeting.	That does not help. I have my interview at 1 pm.… I guess, the 2:25 pm direct flight will have to do.
Agent	I know (I understand, that), you have an important meeting (and I feel for you). Let me try Midway. We have a 1 pm direct flight that gets into Miami at 4 pm but you will have to go to Midway, which means going back into Chicago traffic (I wish, I could help you, get on your original flight. Unfortunately, the only thing I can do, is to re-book you on the 3:30 pm flight).	Great. Please complete this claim form with a delivery address.
Joanna	I don't know.... They all make me late.	Okay. Here it is….
Agent	I understand how you feel, believe me. But which option would work: 3:30 pm direct, 11 am connecting, or 1 pm from Midway? (However, the 3:30 pm flight is direct and will get into Miami at 5:42 pm.)	Your baggage, will arrive on the 2:25 flight, and we will call you before delivering. Have a good day.
Joanna	OK, 3:30 seems to be the best option right now.	
Agent	Excellent choice. I will book your ticket immediately … may I get your credit card?	
Joanna	Yes. Here it is. Thanks.	
Agent	You are all set for the 3:30 departure. The boarding starts at 2:50. Have a good flight.	

Notes: For brevity, we show two conditions for each context. Text in black is common to both contexts. The different colors specify the differences, as indicated.

WEB APPENDIX J
Estimated Coefficients from Confirmatory Factor Analysis of Study Constructs (Airline Experimental Study)

Items	Standardized Loadings[a]	t-Value	Composite Reliability[b]	Average Variance Extracted[c]
Customer Satisfaction				
Upon conclusion of the interaction, how would you be feeling				
1. Very Displeased - Very Pleased	.88/.93	55.77/82.16	.93/.94	.82/.85
2. Very Unhappy - Very Happy	.93/.95	74.21/101.8		
3. Terrible – Delighted	.90/.88	61.57/56.51		
Relational Work				
The extent to which the airline agent in the interaction…				
1. Spoke politely.	.74/.65	25.23/17.68	.89/.88	.74/.72
2. Listened carefully to Joanna's situation.	.88/.95	55.34/75.18		
3. Paid attention to Joanna's concerns.	.94/.91	71.85/63.58		
Solving Work				
The extent to which the airline agent in the interaction...				
1. Suggested feasible options.	.75/.77	26.62/28.66	.90/.87	.75/70
2. Acted competently to solve Joanna's problem.	.91/.86	65.17/45.62		
3. Showed active problem solving	.93/.86	76.87/43.95		
Problem Severity				
Thinking of yourself as Joanna, would you say that the situation is ...				
1. Not at All Distressing - Highly Distressing	.81/.88	5.51/10.48	.69/.83	..53/.71
2. Not at All Stressful - Highly Stressful	.86/.81	5.69/10.22		

[a]Estimated standardized coefficients with corresponding t-values in the adjacent column were obtained with a maximum likelihood solution.
[b]Estimated composite reliability per Fornell and Larcker (1981).
[c]Estimated variance extracted by the latent construct from its hypothesized indicators per Fornell and Larcker (1981).
Notes: For the standard loadings, t-value, composite reliability, and AVE, we report the statistics for MF, followed by LB (MF/LB).

WEB APPENDIX K

Estimated Means for Manipulated Constructs and Outcomes (Airline Experimental Study)

	Missed Flight (MF)			Lost Baggage (LB)		
Conditions	Solving Work	Relational Work	CSAT	Solving Work	Relational Work	CSAT
HSHR	5.55 (1.26)	5.94 (.93)	3.55 (1.19)	5.31 (1.08)	5.73 (.88)	3.14 (1.62)
HSLR	5.83 (.89)	5.65 (1.04)	4.10 (1.10)	5.18 (1.09)	5.48 (1.11)	3.74 (1.55)
LSHR	5.03 (1.34)	5.82 (.99)	3.06 (1.34)	4.69 (1.43)	5.85 (1.14)	2.67 (1.46)
LSLR	4.93 (1.43)	5.29 (1.18)	2.86 (1.35)	4.05 (1.42)	4.57 (1.21)	2.47 (1.15)

Notes: HSHR = high solving, high relational; HSLR = high solving, low relational; LSHR = low solving, high relational; LSLR = low solving, low relational.

REFERENCES (for citations appearing in web appendix and footnotes)

Auh, Seigyoung, Stavroula Spyropoulou, Bulent Menguc, and Aypar Uslu (2014), "When and How Does Sales Team Conflict Affect Sales Team Performance?" *Journal of the Academy of Marketing Science*, 42 (6), 658-79.

Brashers, Dale E., Daena J. Goldsmith, and Elaine Hsieh (2002), "Information Seeking and Avoiding in Health Contexts," *Human Communication Research*, 28 (2), 258-71.

Fair, Ray C (1970), "The Estimation of Simultaneous Equation Models with Lagged Endogenous Variables and First Order Serially Correlated Errors," *Econometrica*, 38(3), 507-16.

Fornell, Claes and David F. Larcker (1981), "Structural Equation Models with Unobservable Variables and Measurement Error: Algebra and Statistics," *Journal of Marketing Research,* 18(3), 382-88.

Jacobson, Robert (1990), "Unobservable Effects and Business Performance," *Marketing Science,* 9(1), 74-85.

Kellas, Jody K. and April R. Trees (2006), "Finding Meaning in Difficult Family Experiences: Sense-Making and Interaction Processes During Joint Family Storytelling," *Journal of Family Communication*, 6 (1), 49-76.

Khan, Zafar U., Sudhir K. Chawla, and Elton A. Devine (1996), "Impact of Gender, Race, and Dress on Choice of CPA's," *Journal of Applied Business Research*, 13(1), 53-68.

Kieren, Dianne K., Thomas O. Maguire, and Nancy Hurlbut (1996), "A Marker Method to Test a Phasing Hypothesis in Family Problem-Solving Interaction," *Journal of Marriage and the Family*, 58 (2), 442-55.

McCollough, Michael A., Leonard L. Berry, and Manjit S. Yadav (2000), "An Empirical Investigation of Customer Satisfaction After Service Failure and Recovery," *Journal of Service Research*, 3 (November), 121–37.

Mizik, Natalie (2014), "Assessing the Total Financial Performance Impact of Brand Equity With Limited Time-Series Data," *Journal of Marketing Research*, 51(6), 691-706.

Mizik, Natalie and Robert Jacobson (2008), "The Financial Value Impact of Perceptual Brand Attributes," *Journal of Marketing Research*, 45(1), 15-32.

Rafaeli, Anat and Michael G. Pratt (1993), "Tailored Meanings: On the Meaning and Impact of Organizational Dress," *Academy of Management Review*, 18(1), 32-55.

Rossi, Peter E (2014), "Even the Rich Can Make Themselves Poor: A Critical Examination of IV Methods in Marketing Applications," *Marketing Science*, 33 (5), 655-72.

Schepers Jeroen, Tomas Falk, Ko de Ruyter, Ad de Jong, and Maik Hammerschmidt (2012), "Principles and Principals: Do Customer Stewardship and Agency Control Compete or Complement When Shaping Frontline Employee Behavior?" *Journal of Marketing*, 76 (6), 1-20.

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

Smith, Amy K. and Ruth N. Bolton (1998), "An Experimental Investigation of Customer Reactions to Service Failure and Recovery Encounters Paradox or Peril?," *Journal of Service Research*, 1 (1), 65-81.

Sotgiu, Francesca and Katrijn Gielens (2015), "Suppliers Caught in Supermarket Price Wars: Victims or Victors? Insights from a Dutch Price War," *Journal of Marketing Research*, 52 (6), 784-800.

Srinivasan, Raji, Gary L. Lilien, and Shrihari Sridhar (2011), “Should Firms Spend More on Research and Development and Advertising during Recessions?” *Journal of Marketing*, 75(3), 49-65.

Westbrook, Robert A. (1980), “A Rating Scale for Measuring Product/Service Satisfaction,” *Journal of Marketing*, 44 (4), 68-72.

Wooldridge, Jeffrey M. (2010), *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.