

---

# Getting Off on the Wrong Foot: The Timing of a Breach and the Restoration of Trust

**Robert B. Lount Jr.**

*The Ohio State University*

**Chen-Bo Zhong**

*University of Toronto*

**Niro Sivanathan**

*London Business School*

**J. Keith Murnighan**

*Northwestern University*

---

*Few interpersonal relationships endure without one party violating the other's expectations. Thus, the ability to build trust and to restore cooperation after a breach can be critical for the preservation of positive relationships. Using an iterated prisoner's dilemma, this article presents two experiments that investigated the effects of the timing of a trust breach—at the start of an interaction, after 5 trials, after 10 trials, or not at all. The findings indicate that getting off on the wrong foot has devastating long-term consequences. Although later breaches seemed to limit cooperation for only a short time, they still planted a seed of distrust that surfaced in the end.*

**Keywords:** *trust; trust repair; trust breach; cooperation; primacy effect*

In *Pride and Prejudice*, Lizzie Bennet's first meeting with Mr. Darcy leads her to conclude that he is arrogant, cold, and rude: It is hatred at first sight. In *To Have and Have Not*, however, Steve (Humphrey Bogart) and Slim (Lauren Bacall) connect with each other immediately: decidedly different beginnings, but very similar endings. In fact, fiction, cinema, and everyday life, all pay considerable attention to the risks and potential benefits of initial interpersonal interactions. Unlike everyday life, literature and movies typically end happily. As these two examples indicate, however, the process that leads to these happy endings is anything but consistent. We all know many stories that support

contrast models—a bad start portends a more positive end—or primacy and impression management models—a good start is needed to even have an end. Although initial interactions provide a foundation for future relationships, the nature of those initial events is clearly not definitive.

Trust, a central element in many interpersonal interactions, can also begin to flourish at the start of a relationship. Rousseau, Sitkin, Burt, and Camerer (1998) defined trust as a “psychological state comprising the intention to accept vulnerability based on positive expectations about another's behavior” (p. 395). Trust is “the single most important element of a good working relationship” (Fisher & Brown, 1988, p. 107). Arrow's (1974) early work indicated that, in organizations, action is predicated on trust. Solomon and Flores (2001) describe its necessity emphatically: “Without trust the corporation becomes not a community but a brutish state of nature, a war of all against all in which employment tends to be nasty, brutish, and short” (p. 5).

---

**Authors' Note:** This research was supported by a grant from the Dispute and Resolution Research Center at Northwestern University. We sincerely thank Eli Finkel, Peter Kim, and Madan Pillutla for their valuable feedback on earlier drafts of this article and Roderick Swaab for his advice about meta-analytic procedures. Please address correspondence to Robert B. Lount, Jr., Fisher College of Business, 720 Fisher Hall, 2100 Neil Avenue, Columbus, OH 43210; e-mail: Lount@fisher.osu.edu.

*PSPB*, Vol. 34 No. 12, December 2008 1601-1612

DOI: 10.1177/0146167208324512

© 2008 by the Society for Personality and Social Psychology, Inc.

Even when relationships are decidedly positive, few endure without one party violating the other's expectations, and the consequences of a violation can be severe. Not only is the ability to build and maintain cooperative interactions critical, but so is the ability to restore a relationship after a breach (Lewicki & Bunker, 1996). Although research has examined how apologies and offers of penance can facilitate the restoration of trust (e.g., Bottom, Gibson, Daniels, & Murnighan, 2002; Kim, Ferrin, Cooper, & Dirks, 2004; Schweitzer, Hershey, & Bradlow, 2006; Tomlinson, Dineen, & Lewicki, 2004), few have investigated the lingering impact of the timing of a breach. Thus, the current research investigates the effects of immediate (no prior interaction), early (several rounds of interaction), and late (considerable rounds of interaction) breaches on trust development and repair. A variety of theoretical models—rational trust, contrast, primacy, and impression formation—provides the essential conceptual foundations for our research; prisoner's dilemma interactions in two laboratory experiments provide repeated observations of positive and negative interpersonal choices.

### *Breaches*

Although models of trust development have many similarities, they differ in their predictions of the nature and consequences of a breach and its timing. Rational choice models predict that trust will grow gradually (e.g., Blau, 1964; Holmes, 1991; Kelley, 1979; Luhmann, 1979; Rempel, Holmes, & Zanna, 1985): As people repeatedly fulfill each others' expectations, confidence and trust can flourish (Rousseau & McLean Parks, 1993). Through this lens, individuals who engage in repeated interactions start with no trust and gradually build trust through mutually beneficial interactions. In this conceptualization, immediate or early breaches should be less damaging than later breaches because there is less trust to be violated.

Contrast theories (e.g., Sherif & Hovland, 1961) use a different conceptual logic to make the same prediction. They suggest that immediate or early breaches may make subsequent interactions seem more positive: Because current experiences create expectations, any changes or contrasts in future interactions become noteworthy and consequential. Thus, supervisors who had evaluated lower performing employees gave considerably better evaluations to other employees than did supervisors who had not made these previously low evaluations (Grey & Kipnis, 1976; Ivancevich, 1983; Sumer & Knight, 1996). Likewise, students who compared themselves with their peers in academically select schools had lower self-concepts than did equally qualified students in less select schools (Marsh, Kong, & Hau, 2000).

Bottom et al.'s (2002) results also suggest that early breaches are less damaging than later breaches. Their participants played a prisoner's dilemma game against an initially cooperative confederate who defected (twice) after 5 or 15 cooperative trials. The participants' messages suggested that early defections led to calm, cognitive, economically oriented reactions, but later defections led to surprised, emotional, trust-oriented reactions (i.e., more negative emotional and interpersonal impact). Later defections that were accompanied by an admission of intent ("I'm sorry but I couldn't help seeing the possibilities") led to even stronger negative reactions. Thus, the logic of rational choice and contrast theories suggests the following:

*Hypothesis 1a:* Later breaches will reduce trust more than immediate or earlier breaches.

Unlike rational choice models, McKnight, Cummings, and Chervany (1998) argued that, even at the start of an interaction, trust can be quite high, as institutional and interpersonal assumptions support trustworthy attributions. Thus, Kim et al. (2004) found that, even without previous contact, people reported considerable trust in others; allegations of untrustworthiness were needed before people would reduce these feelings of trust. Berg, Dickhaut, and McCabe (1995) have also shown that people repeatedly engage in potentially costly trusting acts toward anonymous counterparts whom they have never met. Thus, even before an interaction begins, people have expectations for their counterparts' behavior (Burgooon, Le Poire, & Rosenthal, 1995; Rink & Ellemers, 2007). Unfulfilled positive expectations are then interpreted as a trust breach, resulting in negative outcomes that are both observable (e.g., monetary costs or lost time) and emotional (e.g., anger and resentment). In this view, trust often depends on inherently fragile implicit interpersonal contracts (e.g., Bettenhausen & Murnighan, 1985), making expectation violations, even in the absence of prior interaction, serious events that can sever a potential relationship and obliterate future benefits (Bies, Shapiro, & Cummings, 1988; Bradfield & Aquino, 1999; McCornack & Levine, 1990; Murnighan, 1981).

Weber, Malhotra, and Murnighan's (2005) motivated attributions model is consistent with this logic, suggesting that individuals who are dependent on others for desired outcomes may reverse the assumptions of the rational model: Rather than establishing trust before taking large risks, they often engage in risky, trusting acts to establish trust. In this way, large, noticeably risky acts of trust become potent activators of positive attributions and substantive reciprocity.

If trust starts out high in a new interaction or if people use precipitous trust as a litmus test for trustworthiness (Kim et al., 2004), then immediate or early breaches will have significant consequences, more than later breaches. Immediate or early breaches will have a *primacy effect* and be more difficult to overcome (Lewicki & Bunker, 1996). Research dating back to Asch (1946) repeatedly shows that information acquired early tends to be weighted more heavily than information acquired later (e.g., Anderson & Barrios, 1961; Jones, Goethals, Kennington, & Severance, 1972). Primacy is particularly powerful in interpersonal interactions: Initial information creates a strong impression that resists subsequently discrepant information. In addition, negative first impressions have much more impact than positive first impressions (Skowronski & Carlston, 1989).

Komorita and Mechling (1967) also found that, in a task much like the one studied here, people who experienced early breaches were slower to return to mutual cooperation than people who experienced later defections. Thus, the logic behind models of primacy and impression formation suggests a competing hypothesis:

*Hypothesis 1b:* Immediate or early breaches will make it harder to restore trust than will later breaches.

To test these competing hypotheses, two experiments focused on three key interpersonal processes following a breach: (a) the development, maintenance, and restoration of trust; (b) the behavioral outcome; and (c) mutually cooperative action. Specifically, we charted the progress of repeated interactions that started cooperatively or noncooperatively. We interrupted some of the positive interactions with defections and investigated whether and when individuals overcame these breaches. Experiment 1 examined how the timing of a breach affected the likelihood of trust restoration; Experiment 2 investigated the underlying mechanisms behind trust restoration.

## EXPERIMENT 1

Interpersonal interactions often include the possibility for joint benefit or exploitation. Early on, the parties may have considerable uncertainty about each other's intentions and possible actions (Lewicki, McAllister, & Bies, 1998) and may realize that early cooperation entails the risk of exploitation (Trivers, 1971). If neither party acts cooperatively, however, the chances for future benefits can quickly diminish, reducing the likelihood that they will be able to optimize mutual benefits. The choices of individuals in a repeated, simultaneous-choice prisoner's dilemma game mirrors the dynamics

of these kinds of interactions (e.g., Kramer, Brewer, & Hanna, 1996; Murnighan, 1994): Parties must each choose to either cooperate or defect without knowing what their counterparts have chosen to do.

An iterated prisoner's dilemma game provides a context in which the parties' trust in each other can heavily influence their subsequent choices and the process of their interaction (Rapaport & Chammah, 1965). Early research by Wyer (1969) and more recent neuroimaging research (King-Casas et al., 2005) indicate that people generally want to cooperate and are willing to cooperate if they think that their counterparts will act similarly. Pruitt and Kimmel (1977) contend that people have an expectation that the other person will cooperate before they themselves choose to behave cooperatively. Not knowing whether this will be true, however, means that the choices in prisoner's dilemma games satisfy the definition of trust; that is, a cooperative choice represents a positive expectation while presenting vulnerability for the cooperator. Thus, choosing to cooperate in a prisoner's dilemma represents a behavioral manifestation of trust (Deutsch, 1958, 1973). For example, Kelley and Stahelski (1970) demonstrated that individuals who are willing to cooperate but do not trust their counterparts typically defect, illustrating that expectations of the counterpart's cooperativeness often bet one's own cooperative behavior.

This correspondence between trust and cooperation evolve as iterated prisoner's dilemma games progress (Yamagishi, Kanazawa, Mashima, & Terai, 2005). Early in an interaction, the risks of cooperating are only moderate because continued cooperation leads to better outcomes than does mutual defection (Pruitt & Kimmel, 1977). Once an endpoint is known, however, cooperative behavior decreases (Murnighan, 1981) and the risk of acting cooperatively increases. Game theory posits that rational players will expect their counterparts to defect on the last trial, as there is no chance for retribution (Luce & Raiffa, 1957). Because this same logic applies to the next-to-last and then to all previous trials, game theory predicts that backward induction will lead to rational, noncooperative choices on every trial after the endpoint is known (Rubenstein, 1982). Thus, although early cooperative choices may require some trust, cooperating after the endgame is announced symbolizes the presence of considerably more trust (Murnighan, 1981).

Thus, we used the interactions in iterated prisoner's dilemma games to investigate trust development and restoration (see Camerer, 2002, for a review). To examine the effects of the timing of a breach, we encouraged cooperative interactions and then systematically introduced a breach. More specifically, prior to beginning the prisoner's dilemma, participants read a short, cooperatively oriented

description of the game. Previous research (Bottom et al., 2002) indicates that this leads participants to expect a cooperative interaction. They then experienced a two-trial trust breach at one of three different times—immediately (Trials 1 and 2), early (Trials 6 and 7), or late (Trials 11 and 12). We also included a no-breach control condition.<sup>1</sup> The programmed confederate then cooperated on every trial until the end of the interaction, that is, for 30 consecutive trials. After 20 of these 30 trials, participants were told that 10 trials remained. This changed a repeated interaction with an uncertain end to one with a certain, finite end. We conceptualized endgame cooperation (i.e., cooperation rates during the last 10 trials) to be a behavioral measure of trust. The clearest indicator of trust in this context comes on the very last trial, when there was no future interaction and actors were most vulnerable. Thus, we operationally defined *ultimate trust* as a cooperative choice on the last round.

### Method

**Participants and design.** A total of 138 undergraduates (53% female) at a large Midwestern university volunteered to participate. They were offered \$10 for participating and an additional payoff based on their outcomes. Participants were randomly assigned to one of four trust breach conditions: Trials 1 and 2 versus Trials 6 and 7 versus Trials 11 and 12 versus a no-breach control. The experience of two defections on Trials 1 and 2 was designed to be an immediate and noteworthy negative experience.<sup>2</sup> Five rounds of cooperation prior to defections at Trials 6 and 7 were designed to create a short cooperative relationship; 10 rounds of cooperation prior to defections at Trials 11 and 12 were designed to create a stronger cooperative relationship prior to a breach.<sup>3</sup>

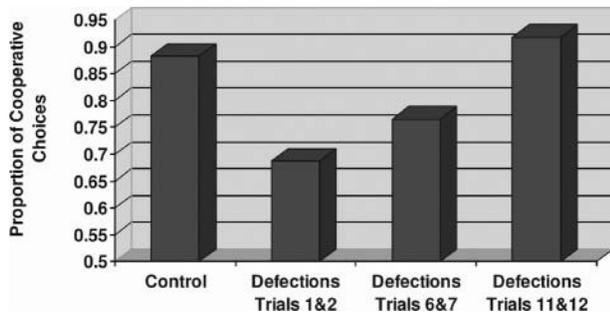
**Procedure.** Each session included approximately 6 participants who never saw or spoke to one another. Participants sat in private rooms at computers that, they were told, were networked together. In fact, participants interacted with a programmed counterpart. To increase the likelihood that they would begin with cooperative choices, participants read the “Tutorial on Cooperation” at the start of the experiment. This was a shortened version of text (Murnighan, 1991, pp. 13-27) that Bottom et al. (2002) used to describe the benefits of cooperation in repeated prisoner’s dilemmas. In addition, the payoff matrix (see Figure 1) and the instructions promoted cooperative choices. To encourage participants to take the task seriously, the experimenter announced that several participants would be randomly chosen to

	X	Y
X	\$24 \$24	\$30 \$6
Y	\$6 \$30	\$12 \$12

**Figure 1** The prisoner’s dilemma payoff matrix.

receive the actual money they won on five randomly selected trials in addition to the \$10 participation fee. Each point in the payoff matrix represented \$1. The entire interaction was conducted via computer. Prior to the first trial and after every subsequent 10 trials, participants could send any message that they wished to their counterparts as long as the messages did not contain identifying information.

As noted earlier, all participants (except those in the control condition) experienced two sequential defections by their programmed counterparts, on Trials 1 and 2, Trials 6 and 7, or Trials 11 and 12. After the two defections, the programmed counterpart cooperated, irrespective of the participant’s responses, for the remaining 30 trials (i.e., 20 trials following the defection and an additional 10 trials following the announcement of the endgame). Once the endgame was announced (i.e., last 10 trials), the number of remaining trials was made visible on the participant’s computer screen. After the last trial, participants completed a post-experiment questionnaire that assessed how trustworthy they perceived the other person to be (i.e., “How trustworthy is the other person?” “How much would you trust them in the future?” and “How honest is the other person?”) on 7-point scales and also asked about any suspicions that they might have had. Participants were then debriefed and paid.



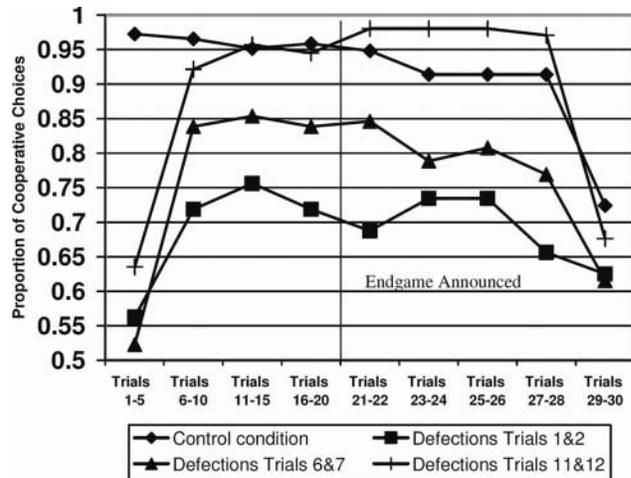
**Figure 2** The proportion of cooperative choices in the last 10 trials in Experiment 1.

### Results and Discussion

*Trust behavior.* Not surprisingly, breaches led to an immediate reduction in cooperative choices (see Figures 2 and 3) but a relatively quick return to a stable rate of cooperation. An immediate breach led to less cooperation over these subsequent trials than did later breaches, particularly a late breach (Trials 11 and 12). In the last 10 trials, people cooperated often (an average of 8.3 times), but at the very end, the previous experience of a breach led to less trust than consistent cooperation.

A one-way analysis of variance (ANOVA), with the timing of the breach as the independent variable and the proportion of cooperative choices on the last 10 trials as the dependent variable, was significant,  $F(3, 134) = 4.93, p < .01$ . Planned contrasts showed that an immediate breach ( $M = .69, SD = .41$ ) led to significantly less cooperation in the last 10 trials than did a late breach ( $M = .92, SD = .15$ ),  $t(134) = 3.54, p < .01, d = 0.75$ . Cooperation after an early breach on Trials 6 and 7 ( $M = .77, SD = .37$ ) was not significantly different from cooperation after an immediate breach,  $t(134) = 1.02, p = .31, d = 0.20$ , but was significantly less than after a late breach,  $t(134) = 2.19, p < .05, d = 0.53$ . Thus, contrary to rational models and contrast theory, immediate breaches hurt trust restoration more than late breaches did, supporting Hypothesis 1b.

Participants who experienced a late breach cooperated more but not significantly more in the last 10 trials ( $M = .92, SD = .15$ ) than participants in the no-breach control condition ( $M = .88, SD = .22$ ),  $t(134) < 1, d = 0.21$ . More pointedly, Figure 3 shows that, for the first 8 of the last 10 trials, late breach participants cooperated more than participants in any of the other conditions and approached complete cooperation. Although the no-breach control condition ( $M = .62, SD = .49$ )



**Figure 3** The proportion of cooperative choices in the 30 trials following defections in Experiment 1.

NOTE: The endgame was announced after the 20th postdefection trial. The scale of the figure changes following endgame announcement.

appeared to lead to more cooperation on the last trial than late breaches ( $M = .49, SD = .50$ ) or immediate breaches ( $M = .53, SD = .51$ ), neither difference was statistically significant,  $\chi^2(1, 80) = 1.27, p = .26, d = 0.26$ , and  $\chi^2(1, 61) = 0.50, p = .48, d = 0.18$ .

Overall, the findings clearly suggest that an immediate breach is particularly difficult to overcome and that later breaches are considerably less harmful.

*Postexperiment questionnaire.* Participants' responses to the three questions about their partners' trustworthiness were averaged into a single score ( $\alpha = .86$ ). The overall  $F$  test was significant,  $F(3, 134) = 10.86, p < .01$ . Unsurprisingly, the trustworthiness evaluation closely paralleled the behavioral data: Participants who experienced immediate breaches reported the lowest levels of trust in their counterparts. Planned contrasts showed that participants in the immediate breach condition viewed their counterparts as less trustworthy ( $M = 4.57, SD = 1.70$ ) than did those in the no-breach control condition ( $M = 6.28, SD = 0.67$ ),  $t(134) = 5.53, p < .001, d = 1.32$ ; they also reported lower trustworthiness evaluations of their partners than did participants who experienced an early breach ( $M = 5.19, SD = 1.06$ ),  $t(134) = 1.95, p = .05, d = 0.44$ , and those who experienced a late breach ( $M = 5.07, SD = 1.12$ ),  $t(134) = 1.82, p = .07, d = 0.40$ . Taken together, the behavioral and perceptual data are clear and consistent: Immediate breaches damaged interpersonal trust more than later breaches did, and later breaches had relatively short-term effects.

**TABLE 1** Results of Mediation Analyses in Experiment 1

Predictor Variable	Model 1	Model 2	Model 3	Sobel Test (Z)
Control condition	.27**	–	.10	3.06**
Defections in Trials 6 and 7	.10	–	.04	
Defections in Trials 11 and 12	.37**	–	.31**	
Trust evaluation	–	.33**	.32**	
R <sup>2</sup>	.10	.11	.18	

NOTE: Regression coefficients are standardized.

\*\* $p < .01$ .

*Mediation.* Given these consistent results, we conducted mediation analyses to examine whether lowered perceptions of trustworthiness mediated the effects of the breaches on trust, as measured by endgame cooperation (Baron & Kenny, 1986; see Table 1). We coded the timing of the breaches into three dummy variables with the immediate defection condition as the baseline (i.e., not included in the model). By using the immediate defection condition as the baseline, the model allowed us to compare endgame cooperation rates in each condition against the immediate defection condition. In Model 1, both the control condition participants and late breach participants showed more endgame cooperation than immediate breach participants, and there was no difference between the immediate and early breaches. In Model 2, trustworthiness evaluations predicted endgame cooperation significantly. Finally, including the trustworthiness evaluation in Model 3 reduced the effect of the timing of breaches: Control participants' responses were no longer significantly different from those following an immediate breach (see Sobel test in Table 1). Although including trustworthiness evaluations seemed to have reduced the impact of late breaches, the difference between late and immediate breaches remained significant.

## EXPERIMENT 2

The results of Experiment 1 demonstrated that the timing of a breach had important effects on trust restoration. Consistent with Hypothesis 1b, immediate breaches were more damaging than late breaches. When a relationship started with noncooperation, trusting behavior never achieved the levels achieved by people who never experienced a trust breach or those who had experienced just 5 or 10 cooperative trials before a breach.

The primary focus of Experiment 1 was on the behavioral consequences of a trust breach: It was designed to minimize methodological intrusions. Although we measured trust ratings at the end of the experiment, we did

not assess immediate or intermediate perceptual and emotional reactions to the breaches. Thus, Experiment 1 does not provide conclusive evidence as to why immediate breaches were so damaging.

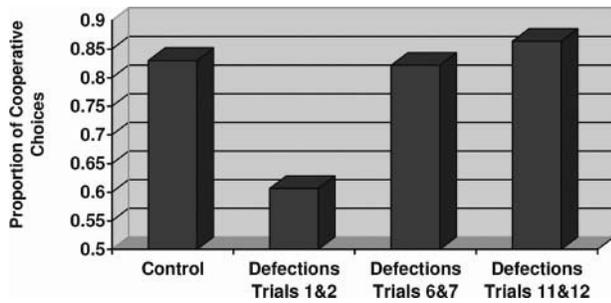
Experiment 2 further explored the damaging effects of trust breaches by measuring participants' perceptual and emotional reactions immediately following the breach (after Trials 1 and 2, Trials 6 and 7, or Trials 11 and 12) and at other, intermittent points after the breach. These additional data allow us not only to assess the reliability of the results from Experiment 1 but also to chart participants' affective and perceptual reactions over their extended interactions. In addition, because the structure of the task was held constant, the data from Experiments 1 and 2 can be combined in a meta-analysis to determine whether interesting, but nonsignificant, observations (e.g., differences among the conditions in the last 10 trials and on the last trial, i.e., ultimate trust) reach standard levels of statistical significance.

### Method

*Participants.* Participants were 108 undergraduates (64% female) from a large Midwestern university, recruited in similar fashion as in Experiment 1.

*Design and procedure.* The methods and design were essentially the same as in Experiment 1: Participants experienced no defections or two defections on Trials 1 and 2, Trials 6 and 7, or Trials 11 and 12, and the programmed counterpart always cooperated for 30 consecutive trials after the breach. Unlike Experiment 1, participants answered a short set of questions about their feelings and their perceptions of their counterparts immediately following the breach and every 10 trials thereafter.

We included three control conditions in which the programmed counterpart always cooperated. These conditions differed in the timing of the short sets of questions, which were asked at the same time as they were asked in the defection conditions, with questions starting after



**Figure 4** The proportion of cooperative choices for the last 10 trials in Experiment 2.

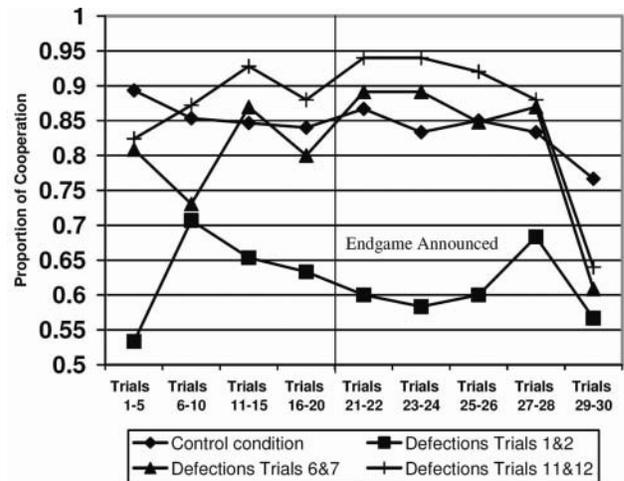
Rounds 2, 7, or 12. We observed no differences between these control conditions in the frequencies of cooperative choices, so the data were pooled for analysis.

*Perceptual and emotional measures.* To measure participants' impression of their partners, we administered Byrne's (1971) Interpersonal Judgment Scale (IJS) and an open-ended question, "What words would you use to describe the other person?" at three different times. The IJS consisted of six questions (e.g., "How much do you like and want to work with the other person on a future task?" and "How intelligent, moral, knowledgeable, and well adjusted is the other person?") with responses on 7-point scales. It was consistently reliable, with alphas ranging from .84 to .86 across the three administrations. The open-ended responses ranged from positive statements such as "trusting and cooperative" to negative descriptions such as "greedy and stupid." Three undergraduate research assistants who were blind to the hypotheses rated the open-ended responses from -3 (many negative statements) to +3 (many positives). Because interrater agreement was high (Time 1  $R_{wg} = .92$ ; Time 2  $R_{wg} = .93$ ; Time 3  $R_{wg} = .91$ ), the ratings of the participant's three responses were averaged.

We also measured participants' emotional reactions, specifically focusing on anger. Participants were asked to report their feelings of anger and irritation on 7-item Likert-type scales. Responses on this 2-item scale were consistently reliable, with alphas ranging from .81 to .85 across the three administrations of the items.

## Results

*Trust behavior.* Immediate breaches again reduced cooperative choices. Later breaches (at Trials 6 and 7 or Trials 11 and 12) led to only moderate reductions in cooperative choices (see Figures 4 and 5). People who



**Figure 5** The proportion of cooperative choices in the 30 trials following defections in Experiment 2.

NOTE: The endgame was announced after the 20th postdefection trial. The scale of the figure changes following endgame announcement.

experienced an immediate breach were less cooperative, on average, than people in all of the other conditions, throughout the experiment (see Figure 5). All three breach conditions led to less cooperation at the very end than did the no-breach control condition.

A one-way ANOVA on the proportion of cooperative choices in the last 10 trials again led to a significant effect for the timing of a breach,  $F(3, 104) = 3.47, p < .05$ . Planned contrasts showed that an immediate breach ( $M = .61, SD = .45$ ) led to significantly less endgame cooperation than did early breaches ( $M = .82, SD = .30$ ), late breaches ( $M = .86, SD = .19$ ), or the combined control conditions ( $M = .83, SD = .33$ ),  $t_s(104) > 2.29, p < .05, d_s > 0.54$ .

As in Experiment 1, cooperative choices after a late breach (Trials 11 and 12) led to more, but not significantly more, cooperation than the control conditions on the first 8 of the last 10 trials (see Figure 5). Although the control conditions seemed to have led to more ultimate trust (i.e., cooperation on the last trial), this difference was also not significant. Comparing the control condition and each of the three breach conditions on ultimate trust resulted in  $\chi^2$  tests  $< 2.50, p > .10, d_s < 0.41$ ; comparing the control with all three breach conditions combined led to  $\chi^2(1, 107) = 2.69, p = .10, d = 0.36$ .

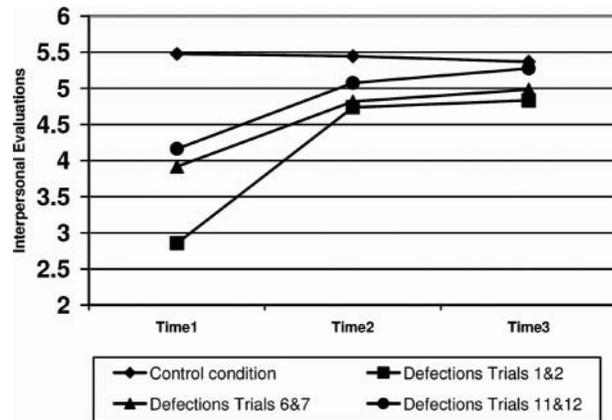
Taken together, and similar to Experiment 1, the findings again suggest that an early breach of trust is particularly difficult to overcome, and later breaches are not particularly harmful, at least until the endpoint approaches.

*Meta-analyses.* Combining the data from Experiments 1 and 2 allows us to conduct more powerful tests on the effects of the timing of a trust breach (Rosenthal, 1991), particularly whether the unpredicted effects for late cooperation following a late breach and for ultimate trust following any breach might be reliable. A first meta-analysis included all of the data in the control and late breach conditions and indicates that late breaches led to more cooperation than did the no-breach control conditions in the first 8 of the last 10 trials ( $z = 1.70$ ,  $p = .04$ ,  $d = 0.30$ ). A second meta-analysis included all of the data from both experiments and compared the controls and a combination of the breach conditions for ultimate trust. This effect was also significant ( $z = 1.74$ ,  $p = .04$ ,  $d = 0.22$ ): The experience of a breach, compared to not experiencing a breach, led to significantly less ultimate trust. This suggests that, although cooperation may be reestablished fairly quickly after a breach, complete trust may not be easy to resurrect.

*Interpersonal impressions.* Responses to the IJS and the open-ended question were significantly correlated at each of the three time points ( $r > .45$ ,  $p < .001$ , at each time). Thus, for each time point, we combined them to form a single measure of interpersonal evaluation.

Participants who experienced immediate breaches had the most negative evaluations. An ANOVA assessing the effects of the timing of the breach and the timing of participants' responses led to a significant effect for the timing of a breach,  $F(3, 104) = 15.44$ ,  $p < .001$ , and a significant Breach  $\times$  Response Time interaction,  $F(6, 208) = 12.51$ ,  $p < .001$ . Planned contrasts revealed that an immediate breach led to the most negative overall evaluations ( $M = 4.14$ ,  $SD = 0.58$ ), differing significantly from all of the other conditions: early ( $M = 4.57$ ,  $SD = 0.83$ ),  $t(104) = 2.07$ ,  $p < .05$ ,  $d = 0.60$ ; late ( $M = 4.84$ ,  $SD = 0.55$ ),  $t(104) = 3.43$ ,  $p < .01$ ,  $d = 1.24$ ; or control ( $M = 5.43$ ,  $SD = 0.95$ ),  $t(104) = 6.67$ ,  $p < .01$ ,  $d = 1.64$ .

Simple effects analyses for each time period indicated that, right after the breach (i.e., Time 1), an immediate breach led to interpersonal evaluations that were significantly more negative than each of the other conditions ( $ts > 4.10$ ,  $p < .001$ ,  $ds > 1.32$ ; see Figure 6). Although interpersonal evaluations improved over time, even after 20 rounds of cooperation (i.e., Time 3), an immediate breach still generated more negative evaluations ( $M = 4.83$ ,  $SD = 1.07$ ) than did no breach ( $M = 5.37$ ,  $SD = 1.00$ ),  $t(104) = 2.10$ ,  $p < .05$ ,  $d = 0.52$ , and marginally lower ratings than did late breaches ( $M = 5.27$ ,  $SD = 0.77$ ),  $t(104) = 1.65$ ,  $p = .10$ ,  $d = 0.47$ .



**Figure 6** Mean interpersonal evaluations in Experiment 2. NOTE: Higher interpersonal evaluation scores indicate a more favorable rating. Participants reported their impressions of their partners every 10 trials following partner breach. Time 1 was administered immediately following a breach, and Time 3 was administered immediately before the endgame was announced (i.e., 10 trials remaining).

Additional analyses investigated the possibility that these interpersonal reactions mediated the effects of the breaches on endgame cooperation (Baron & Kenny, 1986; see Table 2). We coded the timing of the breaches into three dummy variables with the immediate defection condition as the baseline (i.e., not included in the model). In Model 1, each condition led to significantly more endgame cooperation than did immediate breaches. In Model 2, interpersonal evaluations from Time 1 predicted endgame cooperation. In Model 3, including the interpersonal evaluations from Time 1 removed the previous effect for the timing of the breaches; Sobel tests were also significant in each of the three comparisons (see Table 2). Thus, interpersonal evaluations immediately following a breach completely mediated the effect of the timing of a breach on endgame cooperation.

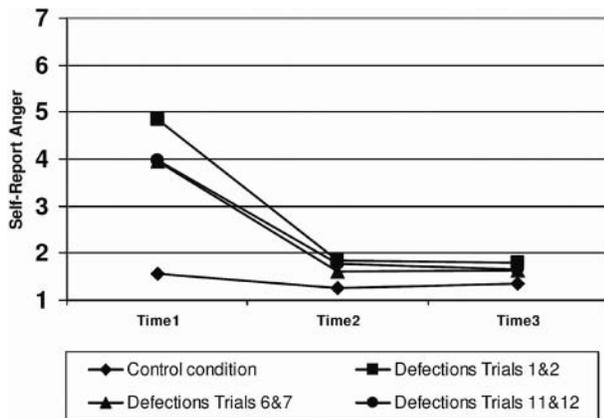
*Emotional responses.* The timing of a breach also induced anger. An ANOVA assessing the effects of the different breaches and the timing of the anger responses led to a significant effect for the timing of a breach,  $F(3, 104) = 15.26$ ,  $p < .001$ , and a significant Breach  $\times$  Response Time interaction,  $F(6, 208) = 16.69$ ,  $p < .001$ . Simple effects showed that differences in anger were more pronounced in Time 1, (i.e., immediately following a breach) and that feelings of anger tended to converge by Time 3. More specifically, at Time 1, an immediate breach produced significantly more anger ( $M = 4.83$ ,  $SD = 1.57$ ) than any of the other conditions:

**TABLE 2** Results of Mediation Analyses in Experiment 2

Predictor Variable	Model 1	Model 2	Model 3	Sobel Test (Z)
Control condition	.29*	–	–.004	2.41*
Defections in Trials 6 and 7	.25*	–	.15	2.12*
Defections in Trials 11 and 12	.31**	–	.18	2.23*
Interpersonal evaluation (Time 1)	–	.32**	.33*	
R <sup>2</sup>	.09	.10	.14	

NOTE: Regression coefficients are standardized.

\* $p < .05$ . \*\* $p < .01$ .

**Figure 7** Self-reported anger in Experiment 2.

NOTE: Higher scores indicate more anger. Participants reported their current anger 10 trials following partner breach. Time 1 was administered immediately following a breach, and Time 3 was administered immediately before the endgame was announced (i.e., 10 trials remaining).

early breach ( $M = 3.96$ ,  $SD = 1.59$ ),  $t(104) = 2.22$ ,  $p < .05$ ,  $d = 0.55$ ; late breach ( $M = 3.98$ ,  $SD = 1.44$ ),  $t(104) = 2.21$ ,  $p < .01$ ,  $d = 0.56$ ; and control ( $M = 1.57$ ,  $SD = 1.10$ ),  $t(104) = 8.88$ ,  $p < .001$ ,  $d = 2.41$  (see Figure 7). Anger at Time 1 was also correlated with endgame cooperation ( $r = -.19$ ,  $p < .06$ ); this was not the case for Time 2 ( $r = -.04$ ) or Time 3 ( $r = -.05$ ). Mediation analyses also confirmed that anger failed to mediate the effects of the timing of a breach on endgame cooperation.

### Discussion

The findings of Experiment 2 replicated and extended the findings from Experiment 1. In particular, compared to later breaches, it appears that immediate breaches are particularly harmful because of their negative impact on interpersonal evaluations. In both studies, initial missteps were tremendously costly. Even a series of 30 repeatedly cooperative choices could not

completely overcome the negative effects of an immediate breach. The results clearly favor the primacy effect (i.e., impression formation) prediction over the rational model/contrast hypothesis. *Getting off on the wrong foot* is a particularly apt saying and perhaps even an understatement.

The data also suggest that cooperative benefits were not hurt much by late breaches or even, in Experiment 2, by breaches in Trials 6 and 7. After an understandable and immediate negative reaction, cooperation rates did not differ from the no-breach controls. These findings are consistent with work by Tomlinson et al. (2004) on reactions to various types of trust breaches. They found that descriptions of a positive prior relationship with the transgressor significantly increased participants' willingness to reconcile the relationship. Our data provide behavioral evidence that as few as five cooperative prisoner's dilemma choices may be sufficient to provide a positive background that can reduce the negative effects of a trust breach.

The findings also suggest other potentially important effects. As noted, the recovery of mutual cooperation was quick. Although the programmed counterparts cooperated for 30 consecutive trials after each breach, stable levels of primarily cooperative behavior were achieved after five cooperative choices and, in some cases, fewer. Cooperation in these conditions was not markedly different from cooperation in the control conditions that were completely free of breaches. Thus, it appears that a short series of cooperative acts can restore mutually cooperative behavior. Future research warrants investigation on just how little a harm-doer must do to return to mutual cooperation. Consistent with work on procedural justice (Lind & Tyler, 1988), several studies have found that a simple apology can be amazingly effective (Bottom et al., 2002; Kim et al., 2004; Tomlinson et al., 2004). Investigating the effectiveness of low-cost, symbolic acts could be particularly fruitful.

Conversely, the meta-analytic findings on last-trial choices suggest that, although cooperation might be relatively easily restored, complete trust was not reestablished. These findings resonate with the oft-heard saying

that once broken, trust can never be completely restored. A good face can be put on a negative situation, but below the surface, the negative impact seems to persist (cf. Gibson, 2006). Moreover, the meta-analytic findings also demonstrated that late breaches led to more cooperation than the no-breach control conditions in the first 8 of the last 10 trials. Although this finding was also unexpected and not predicted, this result suggests that cooperation in ongoing relationships may be enhanced by the knowledge that the other party is capable of both long-term cooperative action and short-term competitive action. In other words, unlike the control condition in which the other party always cooperates, late breaches may present a stronger, more cautionary signal to the counterpart. Future research will need to be conducted to explore this intriguing possibility.

## GENERAL DISCUSSION

The impression formation literature has long taught people to take great care in making a positive first impression. Ambady and Rosenthal's (1993) findings, which show that the first 30 seconds of a class significantly affect teachers' global evaluations at the end of a semester, only reinforces this idea. Adding the concept of trust raises the stakes even further. Thus, the current findings help to amplify the critical nature of large acts of trust (e.g., Pillutla, Malhotra, & Murnighan, 2002; Weber et al., 2005): Acting as if there is not complete trust can put a tremendous damper, or end, to what might otherwise be a particularly beneficial relationship. The current data document not only that people are sensitive to not being trusted but also that they react most when they are not trusted at the very start of a social interaction.

The results are also consistent with previous findings on expectation violations. Prior work has argued that people enter group settings with expectations about their counterparts' behavior (Burgoon et al., 1995; Rink & Ellemers, 2006, 2007). Our findings suggest that people are especially sensitive to the violation of their trust expectations. Participants who experienced an immediate breach had negative evaluations of the other person, experienced more anger, and behaved less cooperatively following the announcement of the endgame as compared to participants who experienced a later breach or control participants who did not experience a breach.

The phenomenology of our experimental design is representative of important, initial personal interactions. When employees meet a new boss, for instance, they are particularly sensitive about how they will be treated. Thus, interactional justice—being accorded

dignity and respect—is not only important when bad outcomes occur, as is the focus in much of the research literature on justice (see Tyler, 2001, for a review), but is also particularly important at the start of an interaction.

Immediate breaches had multiple effects: They led to negative affect about the resulting outcome (distributive injustice), and they communicated disrespect (interactional injustice). After repeated cooperative choices, individuals' outcomes and their willingness to cooperate clearly increased. But the fact that their cooperative choices dropped precipitously on the last trial suggests that they may have never completely overcome the emotional damage caused by an early breach. This appears to be the case in all of the breach conditions in both experiments, not just immediate breaches. In other words, the mutual cooperation that we observed prior to the end of these interactions may not have been based on trust but on the potential for individual benefit. The abrupt drop in cooperation at the very end, in contrast, suggests that complete trust was never reestablished, even after many uninterrupted cooperative acts.

The findings from our two experiments add to a growing body of knowledge on trust restoration (e.g., Bottom et al., 2002; Kim et al., 2004; Schweitzer et al., 2006; Tomlinson et al., 2004). Whereas much of this work has examined the strategies a transgressor can employ to help restore trust (e.g., an apology vs. an offer of penance; Bottom et al., 2002), our research furthers our understanding into *how* and *why* the timing of a breach influences trust restoration. Our results suggest that immediate breaches are especially costly because they seriously damage interpersonal impressions. This was evidenced by the observation that a negative interpersonal evaluation, caused by an immediate breach, mediated the relationship between trust breaches and endgame trust. These findings indicate that interpersonal evaluations are a particularly critical component in the trust restoration process and that these evaluations are especially sensitive to the timing of a breach.

The current results complement Bottom et al.'s (2002) observation that negative emotional responses and reduced cooperation were more pronounced when participants experienced a late trust breach. Although these authors found that a late breach (on Trials 15 and 16) elicited more negative emotions and lowered cooperation than a breach experienced at Trials 5 and 6, they did not investigate immediate breaches. Thus, the current work provides important findings on the impact of an immediate breach on cooperation, trust, and anger, both over time and toward the end of an interaction.

The current findings also provide additional insights into the difficulties parties face in rebuilding trust following a breach. Although previous research indicated

that apologies can help restore trust (e.g., Bottom et al., 2002; Tomlinson et al., 2004), Schweitzer et al. (2006) were surprised to find that an apology was insufficient. A major difference between these sets of experiments was the timing of the breach: Unlike prior research, participants in the study by Schweitzer et al. experienced a breach at the very beginning of a relationship. These results further highlight the deleterious effects of early breaches and raise the question of how the timing of a trust breach moderates the efficacy of various trust restoration strategies (e.g., offering penance, apologies).

We designed our research to map the effects of the timing of a trust breach and to chart the progress of trust restoration after the breach. From a methodological standpoint, the consistency of our participants' behavior in Experiments 1 and 2, even when we interrupted them with intermittent questionnaires in Experiment 2, attests to the reliability of our results. Our research is limited, as all experiments are, by the necessary reduction of complex social relationships to short-term interactions. Although an iterated prisoner's dilemma may mirror many types of interdependent relations, cautious generalizations are clearly recommended. Ten rounds of cooperation prior to a defection, for instance, cannot possibly represent the depth of feeling that results from working cooperatively with a colleague over a week, a month, or a year. This suggests, however, that the reactions that we repeatedly observed may actually underrepresent the reactions of social and organizational actors. Future research will be needed to examine these possibilities.

## CONCLUSION

The findings from our two experiments are clear and reliable: Immediate breaches had long-run consequences that were mitigated, to some degree, by cooperative action. Any breach, however, planted a seed of distrust: A relatively quick return to cooperation looked good, on the surface, but it disappeared in the end. These findings suggest that fictional depictions that support the contrast model, even those of *Jane Eyre*, may be well placed in the realm of fiction. At the same time, the findings also suggest that there is hope—combined with costs—after a trust breach.

## NOTES

1. First, it might seem strange to call noncooperative action on the first two trials a "trust breach." Because this study was conducted in the United States, which has shown a consistent trust bias (Yamagishi & Yamagishi, 1994), this depiction becomes more reasonable. In addition, choosing to cooperate in a prisoner's dilemma game is a vulnerable act, taken with hopeful expectations; this fits our definition, and most definitions, of trusting behavior. Second, we also

included a second factor, a cooperative or an innocuous communication ("Let's cooperate" or "Hello"), prior to the first interaction. This manipulation was included to boost trust prior to the participants' first choices. It had no effect in any of our analyses. As a result, we pooled the data and do not refer to this manipulation again. (This message was not included in Experiment 2.)

2. As in prior research (i.e., Bottom, Gibson, Daniels, & Murnighan, 2002; Komorita & Mechling, 1967; Schweitzer, Hershey, & Bradlow, 2006), the programmed counterpart defected twice to ensure that participants would interpret the defections as a genuine trust breach. If a participant experienced only one defection, he or she might interpret this action as a mistake (e.g., the partner pressed the defection button accidentally). Two consecutive defections reduce the likelihood of this interpretation.

3. Like Komorita and Mechling (1967) and Bottom et al. (2002), final analyses in both experiments excluded participants who defected on Trial 1 or who defected immediately before or on the same trial as the breach. These participants were either uninfluenced by the cooperative prime, or because of the timing of their defections, they may have misinterpreted the breach as a response to their own actions rather than as their counterparts' independent actions. In essence, these participants might have interpreted the breach in a manner that was not intended. These participants—14 in Experiment 1 and 20 in Experiment 2—were evenly distributed throughout the experimental conditions; removing them did not substantially influence the results.

## REFERENCES

- Ambady, N., & Rosenthal, R. (1993). Half a minute: Predicting teacher evaluations from thin slices of behavior and physical attractiveness. *Journal of Personality and Social Psychology, 64*, 431-441.
- Anderson, N. H., & Barrios, A. A. (1961). Primacy effects in personality impression formation. *Journal of Abnormal and Social Psychology, 41*, 258-290.
- Arrow, K. (1974). *The limits of organization*. New York: Norton.
- Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology, 41*, 258-290.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.
- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior, 10*, 122-142.
- Bettenhausen, K., & Murnighan, J. K. (1985). The emergence of norms in competitive decision-making groups. *Administrative Science Quarterly, 30*, 350-372.
- Bies, R. J., Shapiro, D., & Cummings, L. L. (1988). Causal accounts and managing conflict. *Communication Research, 15*, 381-399.
- Blau, P. M. (1964). *Exchange and power in social life*. New York: Wiley.
- Bottom, W. P., Gibson, K., Daniels, S., & Murnighan, J. K. (2002). When talk is not cheap: Substantive penance and expressions of intent in the reestablishment of cooperation. *Organizational Science, 13*, 497-513.
- Bradfield, M., & Aquino, K. (1999). The effects of blame attributions and offender likableness on forgiveness and revenge in the workplace. *Journal of Management, 25*, 607-631.
- Burgoon, J. K., Le Poire, B. A., & Rosenthal, R. (1995). Effects of preinteraction expectancies and target communication on perceiver reciprocity and compensation in dyadic interaction. *Journal of Experimental Social Psychology, 31*, 287-321.
- Byrne, D. (1971). *The attraction paradigm*. New York: Academic Press.
- Camerer, C. F. (2002). *Behavioral game theory: Experiments in strategic interaction*. Princeton, NJ: Princeton University Press.
- Deutsch, M. (1958). Trust and suspicion. *Journal of Conflict Resolution, 2*, 265-279.
- Deutsch, M. (1973). *The resolution of conflict: Constructive and destructive processes*. New Haven, CT: Yale University Press.

- Fisher, R., & Brown, S. (1988). *Getting together: Building relationships as we negotiate*. London: Penguin.
- Gibson, J. L. (2006). The contributions of truth to reconciliation: Lessons from South Africa. *Journal of Conflict Resolution*, 50, 409-432.
- Grey, R. J., & Kipnis, D. (1976). Untangling the performance appraisal dilemma: The influence of perceived organizational context on evaluative processes. *Journal of Applied Psychology*, 61, 329-335.
- Holmes, J. G. (1991). Trust and the appraisal process in close relationships. In W. H. Jones (Ed.), *Advances in personal relationships: A research annual* (pp. 57-104). London: Jessica Kingsley.
- Ivanchevich, J. M. (1983). Contrast effects in performance evaluation and reward practices. *Academy of Management Journal*, 26, 465-476.
- Jones, E. E., Goethals, G. R., Kennington, G. E., & Severance, L. J. (1972). Primacy and assimilation in the attribution process: The stable entity proposition. *Journal of Personality*, 40, 250-274.
- Kelley, H. H. (1979). *Personal relationships: Their structures and processes*. Hillsdale, NJ: Lawrence Erlbaum.
- Kelley, H. H. & Stahelski, A. J. (1970). Social interaction basis of cooperators' and competitors' beliefs about others. *Journal of Personality and Social Psychology*, 16, 190-197.
- Kim, P. H., Ferrin, D. L., Cooper, C. D., & Dirks, K. T. (2004). Removing the shadow of suspicion: The effects of apology vs. denial for repairing ability- vs. integrity-based trust violations. *Journal of Applied Psychology*, 89, 104-118.
- King-Casas, B., Tomlin, D., Anen, C., Camerer, C., Quartz, S., & Montague, P. R. (2005). Getting to know you: Reputation and trust in a two-person economic exchange. *Science*, 310, 78-83.
- Komorita, S. S., & Mechling, J. (1967). Betrayal and reconciliation in a two-person game. *Journal of Personality and Social Psychology*, 6, 349-353.
- Kramer, R. M., Brewer, M. B., & Hanna, B. A. (1996). Collective trust and collective action: The decision to trust as a social decision. In R. M. Kramer & T. R. Tyler (Eds.), *Trust in organizations: Frontiers of theory and research* (pp. 90-103). Thousand Oaks, CA: Sage.
- Lewicki, R. J., & Bunker, B. B. (1996). Developing and maintaining trust in work relationships. In R. M. Kramer & T. R. Tyler (Eds.), *Trust in organizations: Frontiers of theory and research* (pp. 114-139). Thousand Oaks, CA: Sage.
- Lewicki, R. J., McAllister, D. J., & Bies, R. J. (1998). Trust and distrust: New relationships and realities. *Academy of Management Review*, 23, 438-458.
- Lind, E. A., & Tyler, T. R. (1988). *The social psychology of procedural justice*. New York: Plenum.
- Luce, R., & Raiffa, H. (1957). *Games and decisions*. New York: Wiley.
- Luhmann, N. (1979). *Trust and power*. New York: Wiley.
- Marsh, H. W., Kong, C. K., & Hau, K. T. (2000). Longitudinal multilevel models of the big-fish-little-pond effect on academic self-concept: Counterbalancing contrast and reflected glory effects in Hong Kong schools. *Journal of Personality and Social Psychology*, 78, 337-349.
- McCornack, S. A., & Levine, T. R. (1990). When lies are uncovered: Emotional and relational outcomes of discovered deception. *Communication Monographs*, 57, 119-139.
- McKnight, D. H., Cummings, L. L., & Chervany, N. L. (1998). Initial trust formation in new organizational relationships. *Academy of Management Review*, 23, 473-490.
- Murnighan, J. K. (1981). Defectors, vulnerability, and relative power: Some causes and effects of leaving a stable coalition. *Human Relations*, 34, 589-609.
- Murnighan, J. K. (1991). *The dynamics of bargaining games*. Englewood Cliffs, NJ: Prentice Hall.
- Murnighan, J. K. (1994). Game theory and organizational behavior. In B. M. Staw & L. L. Cummings (Eds.), *Research in organizational behavior* (pp. 83-123). Greenwich, CT: JAI.
- Pillutla, M., Malhotra, D., & Murnighan, J. K. (2002). Attributions of trust and the calculus of reciprocity. *Journal of Experimental Social Psychology*, 39, 448-455.
- Pruitt, D. G., & Kimmel, M. J. (1977). Twenty years of experimental gaming: Critique, synthesis, and suggestions for the future. *Annual Review of Psychology*, 28, 363-392.
- Rapaport, A., & Chammah, A. M. (1965). *Prisoner's dilemma: A study in conflict and cooperation*. Ann Arbor: University of Michigan Press.
- Rempel, J. K., Holmes, J. G., & Zanna, M. P. (1985). Trust in close relationships. *Journal of Personality and Social Psychology*, 49, 95-112.
- Rink, F., & Ellemers, N. (2006). What can you expect? The influence of gender diversity in dyads on work goal expectancies and subsequent work commitment. *Group Processes and Intergroup Relations*, 9, 577-588.
- Rink, F., & Ellemers, N. (2007). The role of expectancies in accepting task-related diversity: Do disappointment and lack of commitment stem from actual differences or violated expectations? *Personality and Social Psychology Bulletin*, 33, 842-854.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Newbury Park, CA: Sage.
- Rousseau, D. M., & McLean Parks J. (1993). The contracts of individuals and organizations. In L. L. Cummings & B. M. Staw (Eds.), *Research in organizational behavior* (pp. 1-43). Greenwich, CT: JAI.
- Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review*, 23, 393-404.
- Rubenstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica*, 50, 97-109.
- Schweitzer, M. E., Hershey, J. C., & Bradlow, E. T. (2006). Promises and lies: Restoring violated trust. *Organizational Behavior and Human Decision Processes*, 101, 1-19.
- Sherif, M., & Hovland, C. I. (1961). *Social judgment: Assimilation and contrast effects in communication and attitude change*. New Haven, CT: Yale University Press.
- Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin*, 105, 131-142.
- Solomon, R. C., & Flores, F. (2001). *Building trust: In business, politics, relationships, and life*. Oxford, UK: Oxford University Press.
- Sumer, H. C., & Knight, P. A. (1996). Assimilation and contrast effects in performance ratings: Effects of rating the previous performance on rating subsequent performance. *Journal of Applied Psychology*, 81, 436-442.
- Tomlinson, E. C., Dineen, B. R., & Lewicki, R. J. (2004). The road to reconciliation: Antecedents of victim willingness to reconcile following a broken promise. *Journal of Management*, 30, 165-187.
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, 46, 35-57.
- Tyler, T. R. (2001). Social justice. In R. Brown & S. Gaertner (Eds.), *Blackwell handbook of social psychology: Vol. 4: Intergroup processes* (pp. 344-366). London: Blackwell.
- Weber, M. J., Malhotra, M., & Murnighan, J. K. (2005). Normal acts of irrational trust: Motivated attributions and the trust development process. In R. M. Kramer & B. M. Staw (Eds.), *Research in organizational behavior* (pp. 75-101). Greenwich, CT: JAI.
- Wyer, R. S. (1969). The prediction of behavior in two-person games. *Journal of Personality and Social Psychology*, 13, 222-238.
- Yamagishi, T., Kanazawa, S., Mashima, R., & Terai, S. (2005). Separating trust from cooperation in a dynamic relationship: Prisoner's dilemma with variable dependence. *Rationality and Society*, 17, 275-308.
- Yamagishi, T., & Yamagishi, M. (1994). Trust and commitment in the United States and Japan. *Motivation and Emotion*, 1, 129-166.

Received August 4, 2007

Revision accepted June 1, 2008