

When Talk Is Not Cheap: Substantive Penance and Expressions of Intent in Rebuilding Cooperation

William P. Bottom • Kevin Gibson • Steven E. Daniels • J. Keith Murnighan

John M. Olin School of Business, Washington University, St. Louis, Missouri 63130

Department of Philosophy, Marquette University, Milwaukee, Wisconsin 53201

Western Rural Development Center, Utah State University, Logan, Utah 84322

Kellogg Graduate School of Management, Northwestern University, Evanston, Illinois 60208

bottom@olin.wustl.edu • kevin.gibson@mu.edu • sdaniels@ext.usu.edu • keithm@northwestern.edu

Abstract

Interpersonal relationships can be fragile. The mere perception of opportunistic behavior can lead to a breakdown in cooperation. Once damaged, the question then arises as to whether and how cooperation might be restored. Noncooperative game theory raises serious doubts about the possibilities, although interactional justice and impression management research have shown that verbal explanations can dampen reactions to aversive behavior. Philosophical, anthropological, and ethological research all suggest that genuine forgiveness may require something more tangible and substantive than an explanation. Thus, the current experiment investigated the effects of explanations and varying forms of substantive amends on the restoration of mutual cooperation. The results confirm that rebuilding cooperation is feasible. Apologies and simple explanations can be effective to a degree, though substantive amends have significantly more positive effects than explanations alone. In contrast to prior findings on interactional justice, acknowledgments were more effective than denials in repairing short interactions. This research demonstrates that, once breached, cooperation can be reestablished and that actions as well as explanations and apologies can augment the process in important and sometimes subtle ways.

(Relationship; Talk; Forgiveness)

To err is human, to forgive divine; to forget is stupid.

Author Unknown

The rapid proliferation of new, flatter, and more flexible forms of organization have placed an added premium on the ability to build and maintain mutually beneficial working relationships within and across organizational

boundaries (Child and McGrath 2001). Business partnerships between erstwhile competitors (e.g., Toyota and General Motors, Astra and Merck, Sony and Phillips) carry great opportunities along with clear temptations. As one senior biotechnology executive noted, formal partnerships are “only the tip of the iceberg—they exclude dozens of handshake deals and informal collaborations as well as probably hundreds of collaborations by our company’s scientists with colleagues elsewhere” (Powell et al. 1996, p. 120). Powell et al. conclude that cultivating and managing such complicated relationships is essential to innovation in R&D-intensive industries. Larson (1992) observed that interorganizational relationships generally grow from and depend on personal relationships between individuals. The ability to build, maintain, and repair cooperative interactions has become a critical management competency (Lewicki and Bunker 1996).

Commercial interactions depend on the mutual fulfillment of the individual parties’ expectations of each other and their actions. As expectations are repeatedly fulfilled by actions, accumulating rewards are likely to cement the perception of an implicit agreement between the parties, i.e., an implicit contract (Rousseau and McLean Parks 1993). The perceptual nature of these implicit agreements them an inherent fragility (e.g., Bettenhausen and Murnighan 1985). Actions that violate cooperative expectations can have serious consequences. Minor departures may foster concern. Major departures may be perceived as exploitation, generate strong emotional reactions, and can sever relationships so that future benefits are lost (Bottom et al. 1996, Murnighan 1981). Larson (1992) found that the dissolution of a partnership in an entrepreneurial alliance had even broader ramifications,

causing reputational damage that reduced the ability to conduct business with other firms.

Disappointment and anger can lead to the dissolution of relations or even to desires for vengeance (Bradfield and Aquino 1999, McCornack and Levine 1990). Should one party's actions seriously threaten the continuation of a rewarding interaction—inside, outside, or between organizational actors—one or both of the parties may seek to reestablish the connection so they can continue to receive the benefits it offers. Rebuilding cooperation may prove difficult, however: Retribution, resentment, and guilt can present immediate obstacles. In certain situations, the aggrieved party may actually be the one to seek to rebuild cooperation (Bies and Tripp 1996). In other instances, all parties may mutually seek repair. But in many cases it will be the party that caused the breach who realizes their mistake and seeks to return to the past and its joint benefits (Goffman 1971). This last situation is the domain of the current research.

While our theory considers the broader concept of relationships and their potential, our research investigates initial interactions. The context of our experiment is a potentially profitable interpersonal interaction, making it relevant to many forms of organizational interactions: between the members of the same organization, agents of different organizations (e.g., Larson's 1992 network dyads), or between cohesive groups in which the parties can mutually benefit from cooperation.

A great deal of research has addressed both the initiation of relations and the attitudes that are associated with damaged interactions. Much less has focused on the mechanisms involved in restoring cooperation. As de Waal (1989) notes, most conflicts do not simply destroy relations. Rather, they generate tension followed by either reconciliation or withdrawal. Here we investigate whether two parties can cope with this tension in a way that restores the dynamics of mutually cooperative interaction after one party has disrupted it. We investigate the response to and the effectiveness of several different forms of penance in this restoration process.

As a backdrop for the current research, we briefly review two streams of literature that have addressed damaged relationships. The first focuses on the early interactions in repeated prisoners' and social dilemmas (for a recent review see Komorita and Parks 1999). The second concerns perceptions of justice in interpersonal contexts, i.e., interactional justice (e.g., Bies and Moag 1986). Our primary focus differs from past research by investigating the effects of different forms of substantive amends, or penance, on the potential restoration of mutual cooperation. We also draw on the philosophical literature on forgiveness and ethnographic and ethological studies of ritualized social behavior to create a theoretical foundation

for our experiment. Thus, this study tests a series of propositions on the efficacy of different forms and processes of penance in restoring and rebuilding cooperation.

Cooperative Interaction

We start with a basic assumption that cooperative interactions are mutually desirable for parties that are not in direct competition. The structure of most interactions, particularly between organizational actors, includes the possibility for joint benefit or loss. In the early stages, the parties may have considerable uncertainty about each other's intentions and likely actions and, in particular, whether the interaction will be positive (Lewicki et al. 1998). People often realize that early cooperative choices can entail considerable risk: When they meet a noncooperative choice by the other party, a cooperator may suffer considerable personal and/or organizational losses. However, if neither party acts cooperatively, the chances for future benefits can quickly diminish.

This description of potentially cooperative interactions and its similarity to prisoners' dilemma games is no coincidence: The empirical research we report here uses the prisoners' dilemma for its basic structure. Game-theoretic models of repeated play in prisoners' dilemmas (Roth and Murnighan 1978, Axelrod and Hamilton 1981) determine when it makes rational sense for the parties to choose cooperatively. If (1) each party's actions are contingent on the other party's actions, (2) the interaction has a high probability of continuing, and (3) the payoffs for cooperation are sufficient, then mutual cooperation makes rational sense. In any prisoners' dilemma situation, noncooperation is also a rational choice.

When one party in an interaction makes even the smallest cooperative move and the other party notices and reciprocates, they give themselves an opportunity to increase joint gains. Mutually beneficial outcomes can quickly lead to expectations of further cooperation and a growing sense that the other is obligated to continue cooperating (Bottom et al. 1996). This pattern establishes the foundation for a cooperative psychological contract and confidence that the counterparts share similar or even identical perceptions of the situation (Rousseau and McLean Parks 1993). Localized social norms then become an explicit reality (e.g., Bettenhausen and Murnighan 1985), further cementing expectations and increasing the likelihood of future cooperation.

In dilemma experiments involving a long series of choices, the general tendency is for the parties to gravitate to one of two stable states: mutual cooperation or mutual noncooperation (Pilisuk et al. 1967). Several underlying dynamics seem to enable negotiators to shift from a stable

noncooperative state to cooperation. Individuals who believe that they are seen as competent or who feel that they can influence the system are more likely to be cooperative. Common illusory beliefs in such influence can even generate cooperation in a one-trial game (Hayashi et al. 1999). However Pilisuk et al. (1967) showed that simply providing opportunities for communication between the players did not necessarily increase cooperation, because these opportunities were frequently used to deceive.

Lindsfold (1978) proposed that cooperative interactions would only develop with the appropriate combination of verbal messages and behavioral choices. Lindsfold et al. (1983) investigated patterns of choice and communication that might encourage an otherwise competitively inclined person to act cooperatively. They found that a combination of conciliatory choice initiatives, general and specific cooperative messages, and measured retaliation could best generate cooperation. Lindsfold et al. (1986) further found that this combination could improve cooperative interactions that had been damaged by threats, insults, or challenges by one of the parties. They noted that while specific messages helped, a general note of cooperation that accompanied conciliatory acts worked best. As to actions, Hayashi (1993) has shown that a cooperate-then-exit strategy, that is, to cooperate unconditionally until a counterpart defects, is particularly effective at generating cooperation when coupled with a type of forgiveness mechanism (forgive after three rounds in this particular experiment).

The Importance of Talk

Research on interactive justice also highlights the importance of communication in building cooperation, especially when it signals intent. Goffman's (1971) analysis of intent and its effects indicates that "there is no act whose meaning is independent of reasons understood for its occurrence" and further that "the perception of an act . . . is dependent on an assessment of intent" (p. 110). Throwing someone to the ground to impede him from stepping on a just-cleaned floor is an offense. Doing so to prevent him from being shot by a sniper changes the meaning of the act completely. If the negative consequences of an action are unforeseeable, this may also mitigate the perception and experience of guilt.

Wrongdoers' verbal characterizations of their acts can shape an injured party's interpretation of the wrongdoers' intentions, which in turn can affect subsequent reactions (Mehlman and Snyder 1985). Bies et al. (1988) found that accounts identifying external causes helped blunt future conflict and reduce negative reactions to harm. Shapiro (1991) found that external explanations (suggesting that the event was unintentional) led to reductions in disapproval, perceived injustice, and the desire to punish, and

to more forgiveness. Bies and Shapiro (1987) showed that the perceived adequacy of an explanation was critical in reducing negative reactions. Goffman (1971), however, made it clear that explanations must provide sufficient substance—they must match the severity of the offense—to be believable. Shapiro et al.'s (1994) results concur, showing that explanations were less effective when the damaged party had suffered severe consequences.

These two streams of research have produced a wealth of findings relevant to the current issues. Dilemma research has used carefully controlled experiments to examine the likelihood of generating cooperative behavior under difficult circumstances. Justice research has distinguished different dimensions of the explanations given for actions and their psychological impact on aggrieved parties. Each stream suggests that communication and action both promote cooperation. The combination of the two may be most potent. The research on interactional justice further suggests that explanations that identify specific external attributions as causes are likely to be most effective.

Neither stream of research, however, has directly examined the central issues in the current project. Dilemma research has focused on establishing cooperation rather than on rebuilding it. Interactional justice research has concentrated on attitudinal and emotional reactions rather than on subsequent behavior or reconciliation. The current study draws from both literatures to investigate the consequences of disrupting cooperation and subsequent attempts to enact its restoration. To control for other factors, we investigated interactions that share a common sequence of events from initial origin to breach. The interactions begin positively and are then threatened by one party's unexpected actions. We examine the impact of two different explanations for the breach and several forms of substantive amends and assess their impact on attitudes, emotional reactions, and behavior.

Penance and Forgiveness

Rebuilding cooperation may depend, in addition to retribution, on satisfying both parties' demands of each other. If an injured party does not desire revenge, they may instead look for injurers to repent (Bradfield and Aquino 1999). Similarly, many injurers may hope that the people they have injured will forgive them. Reestablishing previous cooperation, then, may require substantive and emotive concessions by *both* parties (North 1987).

Social rituals for rebuilding damaged relationships are common among human (Boehm 1984, Koch et al. 1977, Sigmund 1999) and other highly social animal societies (de Waal 1989, Schenkel 1967). Social groups that lack

formal institutions for the enforcement of laws and the resolution of conflict use rituals to smooth social functioning and prevent dangerous escalation and violent retaliation (Arno 1976, Boehm 1984). Ritual practices typically include acknowledgment of the harmful act and the offering of a gift or some symbol of repentance. The traditional Fijian ritual known as *i soro* calls for a gift of either a bag of kava or a piece of whalebone (Arno 1976). Most rituals also entail a degree of vulnerability or submission. In Boehm's (1984) history of Montenegrin tribes, several lengthy and destructive feuds ended with a delegate of one clan crawling on hands and knees toward the rival clan while offering an apology. The humiliation of the ritual was so great that the receiving party generally rushed toward the penitent to accept their offer and put an end to the shame.

A serious obstacle to any reconciliation is when the two parties entertain different conceptions of the severity of a transgression and what constitutes fair recompense (e.g., Thompson and Loewenstein 1992). Injurers may feel that a signal of their renewed good intentions is sufficient. The injured may desire complete reparation for all that they have lost—and possibly a further penalty as well. At a minimum, they may require that the injurer suffer some substantive cost (e.g., a fine or reduction of wages or bonuses). Arno (1976) describes the case of a young Fijian who had gotten inebriated and then loudly insulted, in public, the “talking chief” of the island over his alleged incompetence. Though he had already offered the reconciliation ritual *i soro* to several others he insulted that evening, the young man informed Arno that the particulars of this case did not oblige him to do so. The chief held a different view and filed charges of drunk and disorderly conduct with the local police. The transgressor hastily offered *i soro* and the chief dropped the complaint, noting to Arno that he had been hoping the young man would take this initiative all along.

Experimental studies of impression management (e.g., Schwartz et al. 1978) also indicate that the expression of remorse following a wrongful act can mitigate punishment. Because the sincerity of any verbal apology is rarely certain, however, the payment of penance may be necessary to validate an apology. In game-theoretic terms, an apology without an objective penitential act may be (and be interpreted as) cheap talk. Cheap talk is unverifiable, costless communication from a self-interested party (Farrell and Gibbons 1989; Pillutla and Murnighan 1995). Without validation, a forgiver who trusts cheap talk risks suffering further wrong from the same party.

Forgiveness does not mean denying, condoning, or forgetting wrongdoing. It acknowledges the event by releasing negative feelings and replacing them with a sense of

goodwill toward the transgressor (Thoreson et al. 1998). Forgiveness requires effort from the person who was originally wronged and may entail at least as much risk as initial cooperation in a new interaction, but it also gives people the opportunity to rebuild cooperative relations, reduce negative affect, and increase positive affect (Witvliet et al. 2001). It should manifest itself most clearly in a willingness to consider engaging in future interactions with the transgressor.

This suggests that an apology, substantive penance, and forgiveness may all be necessary for rebuilding cooperation. If any of the three are missing, the relationship may be unbalanced, fragile, and in danger of dissolution. This conclusion differs from those suggested by interactional justice research, which concentrates on verbal explanations and accounts, and from game theory, which suggests that actions are what matters. Prior to introducing an experiment to investigate the need for verbal or substantive action (or both), we first formulate a set of more formal hypotheses.

Hypotheses

Evidence has repeatedly shown that effective explanations can dampen negative emotions and attitudes. Explanations are also likely to be helpful in the process of rebuilding cooperation. An open question addressed in the current study is whether an explanation can have more than an affective and perceptual impact and result in an actual return to cooperative behavior that involves risk or cost. We tested these ideas by investigating explanations that either acknowledged or denied intent after one party behaved noncooperatively in a previously cooperative interaction.

Dilemma research suggests that explanations alone will not be able to regenerate cooperative behavior. A harm-doer may claim lack of intent simply to avoid a negative reaction, not because it is true (e.g., Elsbach 1994, Scott and Lyman 1968). Models of cheap talk suggest that words alone should have no impact on cooperation, regardless of their intent (Farrell and Gibbons 1989).

If we expect an individual's affect to be consistent with his or her actions, then the implications of interactional justice (that explanations will have a favorable perceptual effect) and of dilemma research (that explanations will not influence behavior) present a puzzling contradiction. To deal with this inconsistency, we state our first hypothesis positively, with the understanding that if it is supported, it provides additional behavioral support for interactional justice. If it is not, it provides additional perceptual support for game theory and for dilemma research.

HYPOTHESIS 1. *Verbal explanations after a breach will have a positive effect on a harmed individual's affective reactions and on the interacting individuals' subsequent cooperative behavior.*

We can refine this hypothesis, following the interactional justice literature, by considering intent. Prior theory and research suggest that explanations evoking external attributions will be effective at mitigating the perceived damage of a negative, perceptually unjust outcome:

HYPOTHESIS 2. *An explanation for a breach of cooperation that denies intent will generate more positive affective reactions and more cooperation than an explanation that acknowledges intent.*

We have noted that the philosophical, anthropological, and ethological literature all indicate that something more than an explanation is needed to achieve forgiveness. Wilson (1988) argues that reconciliation requires penance. The most general form of penance requires that the party seeking forgiveness offer the other party a fixed payment. Reciprocity models require that penance equal the injured party's costs, i.e., the ancient logic of "a tooth for a tooth." Thus, tit-for-tat (Axelrod and Hamilton 1984) immediately penalizes but immediately forgives following a return to cooperation. Yet suffering equal costs may not provide an injured party with full psychological and emotional recompense. Boehm (1984), for instance, has shown that Montenegrin scoring rules for feuds were clear and well understood: A clan that had lost one life should take either two lives or the life of a higher-status member from the other clan.

The experience of a loss may leave a continuing residue of inequity (e.g., Shapiro et al. 1994). Whether partial retribution can overcome potentially inequitable feelings to help restore long-term cooperation is an open empirical question that we test here. A simple model of reinforcement learning or of impression management (Schlenker 1980) would suggest that quantity is critical, i.e., that more substantive penance is likely to be more effective than less. This suggests:

HYPOTHESIS 3. *Substantive offers of penance will generate more positive affective reactions and more cooperation than statements of apology without penance.*

HYPOTHESIS 4. *Larger substantive offers of penance will generate more positive affective reactions and more cooperation than smaller substantive offers.*

Contextual Factors

Because the interpretation of a message depends on its context, the parties' interaction history should interact

with the content of their messages to determine their effectiveness. Context may, for instance, affect the likelihood that a victim will believe a denial. Disbelief could provoke even harsher negative reactions than an initial negative act (Scott and Lyman 1968). Thus, explanations that lack credibility might make future cooperation even more difficult (cf., Bies and Shapiro 1987).

Ongoing relationships are multifaceted and can vary in many respects. Here we focus on one of an interaction's most basic and universal dimensions, its duration (Krackhardt 1992). As individuals begin to interact, they may feel considerable ambivalence because uncertainty about each other's intentions and actions is at its peak (Lewicki et al. 1998). A repeatedly cooperative interaction can reduce uncertainty and reassure both parties. Questions about the others' intentions (e.g., "Can I really depend on them to deliver on time?") become less salient after repeated positive experiences (e.g., "They have come through before. There's no reason to think that they won't come through now."). In contrast, the interpretation of noncooperative behavior may depend on an interaction's history. Admission of noncooperative intent early in an interaction can resolve the doubts of a victimized party, albeit negatively, while an attempt to deny intent may seem transparent and opportunistic (e.g., "Why should I believe such a lame excuse?").

Expectations of others' intentions solidify as interactions lengthen. With a cooperative history, a denial of intent may go unquestioned (e.g., "We knew they couldn't have meant to do that"). Komorita and Mechling (1967), for instance, found that a breach in cooperation was less damaging (in terms of future cooperation) after ten mutually cooperative choices than it was after four. In contrast, an admission of intent following a long history of cooperation may be so unexpected as to be perceived as a genuine violation (Morrison and Robinson 1997). Such perceptions not only incur cognitive adjustments but can also have profound emotional impact (e.g., "How could they have done that after all that we have accomplished?"). Thus, admissions of intent are likely to be much more devastating for interactions with a longer history.

This analysis provides the basis for two hypotheses:

HYPOTHESIS 5. *Following a short history of cooperative interaction, a denial of uncooperative intent will lead to more negative affective reactions and less subsequent cooperation than an acknowledgment of intent.*

HYPOTHESIS 6. *Following a longer history of cooperative interaction, a denial of uncooperative intent will lead to more positive affective reactions and more cooperation than an acknowledgment of intent.*

Procedural justice research has also demonstrated that people are more willing to accept poor outcomes when they perceive that the allocation process was fair (Folger 1993, Brockner and Wiesenfeld 1996). In contrast, negative outcomes resulting from an unfair process lead to perceptions of injustice and, in many cases, attempts at retaliation or redistribution. The justice literature also suggests that people will be more accepting of what they view as an unfair outcome if they have a chance to contribute to the process, even if the negative outcome stands (Hirschmann 1970).

In 1919, the Allied and Associated Nations convened in Paris to consider what reparations to demand of Germany for the damage it had inflicted during World War I. After extended deliberation and bargaining, the delegates ultimately decided that no fixed sum they could specify would prove satisfactory to their constituents. They devised instead a provision whereby Germany would commit to an open-ended obligation for paying reparation demands by the Allies. This provision proved to be very damaging to postwar relations and to Germany's fledgling democracy (Bottom 2001). It is conceivable, however, that some similar open form of penance might have played a constructive role had it been initiated by Germany, the culpable party, rather than imposed by the Allies. The significance of voice (Hirschman 1970) suggests that the most successful approach to substantive forgiveness may actually be an open-ended *offering* of penance. With such an offer in hand, the harmed party can specify the recompense that they desire and can exact as much penalty as they wish, completely controlling equity restoration.

Such a formula mitigates egocentric bias (Thompson and Loewenstein 1992) on the part of the wrongdoer, gives voice to the injured, and allows them to actively participate in the resolution process (Lowin 1968). Because this approach has not been systematically investigated before, we cannot predict the size of the penalty that wronged parties will seek. It is clear, however, that open offers of penance expose the offering party to significant risk. Yet with many species of primate and canine, it appears to be precisely this assumption of risk and the act of open-ended offering that is of primary importance to the forgiveness process, especially because serious physical measures of revenge are infrequently exacted (de Waal 1989). Boehm (1984) observed that receiving clans would often leave unclaimed the considerable blood money offered as part of the pacification of feuds. Instead, they would take pains to demonstrate that it was the symbolism of the offering that mattered. One of the longest recorded feuds ended with the leader of the aggrieved clan taking only one gun, the one which had

killed his brother, from a huge cache of weapons they had been offered by their enemies. If risk bearing, the opportunity for participation, and voice each have value, open offers of penance should diminish the need for a substantive penalty to equal the costs of an initial harm. This suggests the following hypotheses:

HYPOTHESIS 7. Open offers of substantive penance will generate more positive affective reactions and more cooperation than fixed substantive penance.

HYPOTHESIS 8. In response to an open offer of penance, injured parties will request less payment than the costs originally imposed by the harm.

Methods

Experimental Design

The wealth of prior data on behavior in repeated prisoners' dilemmas provides a ready benchmark to assess cooperative behavior. The dilemma has also been used as a model for many kinds of interpersonal interactions, including those between individuals working together within an organization (Kramer et al. 1996) and those between purchasing agents and sales agents for different organizations (Heide and Miner 1992).

In the present study, the dilemma required participants to make a series of (cooperative or noncooperative) choices that accumulated points and determined their chances for a bonus of \$10. Points were displayed in a 2x2 matrix with value-neutral labels for their choices. Player A chose "Up" or "Down;" Player B chose "Left" or "Right." Up-Left (the payoff for mutual cooperation) yielded 3.6 points for each player; Down-Right (the punishment for mutual defection) yielded 0.03 points for each. Down-Left gave 4 points to Player A and none to B; Up-Right yielded the reverse.

To study the rebuilding of damaged interactions we needed to create uniform cooperative conditions and then systematically introduce a breach. We primed participants' initial choices by having them read "A Tutorial on Cooperation" at the start of the experiment. This was a condensed version of the text in Murnighan (1991, pp. 13–27) describing the benefits of cooperation in repeated prisoners' dilemmas. (The full passage is available from the authors upon request.) Three pilot sessions testing instructions and instrumentation demonstrated that the priming was effective, yielding uniform initial cooperation. Participants' "opponents" were actually a program that further encouraged early cooperation by making the same choice, each round, that participants made. Thus, we created an environment where a cooperative interaction would emerge. Then we manipulated the duration of

the interaction prior to the breach, which was followed by various forms of explanations and penance.

Explanations. The first verbal statement following the breach of cooperation provided an explanation for the act that either acknowledged or denied intent. For denials, the program sent the following written message: "I didn't mean to do that; the experimenter took the wrong card." For acknowledgments, the program sent this message instead: "I must admit that I meant to do that; I was just trying to do a little better for myself."

The denial placed blame on the experimenter and denied personal responsibility. This message, however, was potentially ambiguous: Its claim could be either accepted or questioned. In contrast, the acknowledgment, "I was just trying to do better for myself," clearly accepts responsibility and admits intent. This admission, however, does not necessarily provide a good reason for the breach. It fails to excuse because pure self-interest provides neither justifiable (Mills 1995) nor acceptable grounds for breaking an implicit agreement. It is at best an example of what Konovsky and Jaster (1989) called a "weak excuse."

Penance. Across all conditions, the sequence of actions and messages following the breach and the acknowledgment or denial of intent was identical. After four additional rounds of noncooperative choices (paired with innocuous messages; see Table 1) to extinguish cooperation, the program sent each participant a verbal apology and a request to return to cooperation, i.e., "I am sorry for doing this. I think we should go back to cooperation." The penance manipulation followed this message. The four penance conditions were: (1) mere talk: "I would be willing to do this if you are;" (2) small penance: "I would be willing to play Left on the next round and let you play Down," i.e., offering one round of maximum payoffs for the participant; (3) large penance: "I would be willing to play Left on the next two rounds and let you play Down," i.e., offering two rounds of maximum payoffs; or (4) an open-ended offer of penance: "What will it take for you to cooperate again?"

Duration. We manipulated the duration of the interaction by varying the number of rounds prior to the breach. We chose 5 rounds for the short condition to give participants enough experience to formulate expectations about each other, but not necessarily firm expectations, and 15 rounds for the long condition to allow participants to establish more certainty about each other. Breaching at 15 rounds also guarded against potential fatigue. Experience in similar experiments (e.g., Kahn and Murnighan 1993) with similar populations suggested that these manipulations were reasonable.

Following the breach, a programmed message either acknowledged or denied intent. The program then made a second pair of noncooperative choices, with the following message in all conditions: "I was worried that you would choose Down this time so I chose Right." It then made a third pair of noncooperative choices, with this message: "I am just trying to protect myself."

After the apology and offer of penance, the penance was delivered in the following four rounds. Then the experimenter announced that the interaction would end after five more rounds. This changed the situation from a repeated interaction with an uncertain end to one with a certain, finite end. Noncooperative game theory predicts that rational players will neither cooperate nor expect cooperation on the last round, as there is no further chance for retribution. Given this expectation, the same logic suggests that they should not cooperate nor expect cooperation on the next-to-last round either. Applying the same logic to prior rounds creates a process of backwards induction (Stahl 1972) and the conclusion that the only rational choice on every finite round is noncooperation. Thus, the last five rounds provide a critical behavioral test for the effectiveness of different forms of penance: Any cooperation on these rounds is a strong indication of the successful rebuilding of cooperation.

Participants

A total of 225 students from the undergraduate and MBA programs at Washington University in St. Louis volunteered to participate for the opportunity to earn up to \$10 for a one-hour experiment, with a guaranteed minimum of \$3. Ages ranged from 19 to 47, averaging 22; 151 were males; 74 were females; 145 were undergraduates; 56 were MBAs; 4 were Ph.D.s. Most participants were either Caucasian or Asian American; 32 were Asian; 6 were African-American; 5 were European; 2 were South American.

Procedure

Each session included between 4 and 10 participants, always run with even numbers. Sessions were randomly assigned to the experimental conditions. Each participant sat in a cubicle, preventing observation of the actions of the experimenter or other participants. Participants read the cooperation tutorial as the experimenter read it aloud; they then read the experimental instructions.

Although all participants were actually Player As and all Player Bs were fictitious, participants were led to believe that half of the participants at their session were Player Bs. Player As chose either Up or Down; Bs chose Right or Left. Players earned experimental points according to the payoffs described previously (3.6 each for mutual cooperation; 0.03 each for mutual defection; 4.0 was

Table 1 Messages Sent by Player B, the Participants' Programmed Counterpart

Round	Short Play	Message	Long Play	Message
1	Left		Left	
2	Left	Keep it up.	Left	Keep it up.
3	Left		Left	
4	Left	Keep up the good work.	Left	Keep up the good work.
5	Left		Left	
6	Right		Left	
7	Right	[<i>Intent Manipulation</i>]	Left	This is going well.
8	Right		Left	
9	Right	I didn't think I could choose left because I thought you'd pick down.	Left	This is definitely the way to go.
10	Right		Left	
11	Right	I felt I had to protect myself.	Left	I see no reason to stop.
12	Right		Left	
13	Right	[<i>Penance Form Manipulation</i>]	Left	I'm happy to continue this way.
14	Small/large penance/bargainers choices		Left	
15	Large penance/bgr.'s choices		Left	This is smooth sailing.
16	Bargainers' Choices		Right	
17	Bargainers' Choices		Right	[<i>Intent Manipulation</i>]
18	Bargainers' Choices—Last Round		Right	
19			Right	I didn't think I could choose left because I thought you'd pick down.

20–28

Duplicates Short Interaction Rounds 10–18

the temptation to defect; 0 was the sucker's payoff). Participants were told that the points they accumulated would determine their chances of winning a \$10 payoff. Their outcome would be based on the roll of a 100-sided die, which was shown to all participants. They would win the \$10 if the roll resulted in any number between zero and the total number of points they had earned. The instructions emphasized that their chances of winning \$10 depended on their ability to accumulate points and on the roll of the die.

The instructions also indicated that the experiment would terminate with a probability of 0.07. Following every round, the 100-sided die was rolled; if a number between one and seven resulted, the experiment ended. They were told that if the interaction did not terminate for some time, the experimenter would announce the last round with considerable forewarning. In fact, the experimenter controlled the number of rounds, regardless of the roll of the die: The exact number depended on the duration and penance conditions and the end was always announced, five rounds hence (see Table 1).

The instructions and other characteristics of the interaction were constructed to encourage cooperative choices on the early rounds of play. The combination of payoffs

and the probability of continuing led to a *g* index, which computes the expected values from cooperation in a repeated prisoner's dilemma (Murnighan and Roth 1983). In this case, $g = 46.8$, indicating that mutually cooperative choices were rationally sensible and beneficial to both parties.

After even-numbered rounds participants could send a written message to the other party. At each opportunity, the program sent a message (see Table 1). Prior to the breach, each of the program's messages expressed contentment with the process and intent to continue. After Round 5 the experimenter announced that, to speed things up, participants would make choices for two rounds at a time: Participants then indicated their choices for Rounds 6 and 7. In short interactions, the program immediately defected, making two Right choices. In long interactions, the program continued to make cooperative Left choices until it defected on Rounds 16 and 17.

To ensure that cooperation had been fully extinguished prior to the penance manipulation, four rounds of non-cooperative, programmed choices followed the breach. If participants shifted to noncooperation immediately after the breach, then small penance was equivalent to the program accepting a loss equal to half of the participant's

loss; large penance was equivalent to the program accepting a loss equal to the participant's loss. In neither case would penance restore the points the participant would have received if the two players had continued to choose cooperatively. Also, if participants did not respond by choosing noncooperatively immediately, then small and large penance provided even less compensation relative to the loss.

Four rounds of single choices allowed for the delivery of all but one act of penance (an open request for five rounds of maximum payoffs). Then the experimenter announced a total of five more rounds, played one at a time. Thus, participants had nine rounds of choices after the penance manipulation, some of which included penance. Every interaction had at least five postpenance choices. After the last round, participants completed a short questionnaire assessing their attitudes, perceptions of the other player, and emotions. The experiment concluded with a thorough debriefing.

Design and Analysis

The original experiment was a 2 (denial or acknowledgment of intent) \times 2 (short or long duration) \times 4 (mere talk, small, large, or open penance) factorial design, with all variables between subjects. Participants' choices on the last five rounds (when mutual noncooperation was the only equilibrium) were the primary dependent variables. We also noted whether choices were cooperative on the last round and whether participants chose noncooperatively on all of the last five rounds.

We transcribed participants' messages and used a simple coding scheme to count their responses. Most models of the repeated prisoners' dilemma suggest that people should (and do) use contingent strategies such as tit-for-tat; thus, we counted if-then statements (e.g., "As long as you choose up, I'll choose left.") prior to the breach, after the breach but prior to penance, and after penance. We also noted the timing of messages mentioning "trust" as well as any instances of disparagement or derision (e.g., "you jerk!"), economic arguments (e.g., "the numbers are clear: up-left gives both of us the best chance of winning the \$10"), and accusations of lying.

Given our model, a direct measure of forgiveness would have been useful. However, inclusion of such an item might signal to astute participants that the experimenter knew that forgiveness would be an issue. To maintain the illusion that the other player was not programmed, we elected to rely on three questions in the postexperimental questionnaire that asked about future interactions with the other party (How much do you trust, like, and want to work with this person in the future?) and about positive and negative affect (How did you feel

during the experiment—good, distressed, aroused, surprised, pleased, angry, hostile, satisfied, astonished?). All responses were provided on five-point rating scales. A factor analysis of the questionnaire responses led to the formation of three affective/cognitive indices, described below.

Results

Phrasing of the Open Form of Penance

The open penance condition led to an unexpected outcome: Many participants simply did not respond to it, either by not responding to the offer or not requesting any penance. When we explored possible reasons for these reactions, we realized that the phrasing of the offer may have implied a shifting of responsibility. The message apologized ("I am sorry for doing this") and expressed an interest in returning to mutual cooperation, but it directed the burden of reestablishing cooperation to the victim by asking, "What will it take for you to cooperate again?"

We therefore added a new form of open penance that no longer shifted responsibility to the victim. Forty new participants (recruited exactly as we had recruited the original sample) experienced a condition in which the defector's open offer of penance accepted personal responsibility. It now read, "What can I do to get you to cooperate again?"

This new form of open penance led to a change in our data analysis strategy. We analyzed the three fixed penance conditions (mere talk, small penance, and large penance) and three open penance conditions (the original with and without requests for penance and the new form) in separate analyses. (This was a conservative strategy because an overall analysis including both fixed and open conditions would have generated many effects and inflated the apparent significance of the data.)

The open penance condition now included three levels: participants from the original open penance condition who did not respond to the open offer (Open/No Ask); those who did respond (Open/Ask); and the new participants (Responsible/Open condition). The open/ask and responsible/open conditions also created an additional dependent variable, i.e., how much penance the participants requested. We analyzed this measure as a function of intention and duration.

Control Checks and Affective/Cognitive Measures

We conducted two analyses to insure that, apart from the manipulation of the independent variables, all of the conditions were similar. Analysis of cooperative choices on the first five rounds indicates that the great majority were

cooperative (the grand mean was 4.62). The rates of initial cooperation did not vary by condition. Also, analysis of the number of rounds needed to extinguish cooperation after the breach revealed no significant effects in the analysis of fixed penance forms. There was one significant effect in the analysis of open penance, the intent by duration by penance interaction, $F(2,71) = 4.26, p < 0.05$. This effect depended on two sparse, open/ask conditions ($n = 3$ in each case) and may therefore not be reliable. Collectively, these results indicate little variation in cooperation across conditions prior to the penance manipulation.

Factor analysis of the questionnaire items led to a three-factor solution. (A maximum likelihood procedure yielded eigenvalues of 13.48, 7.09, 5.97, 0.74, and 0.37. A promax rotation yielded three factors with loadings exceeding 0.30 for each item within a factor.) The factors were *positive emotions* (good, pleased, and satisfied), *negative emotions* (distressed, angry, hostile, astonished, and surprised), and a more complex *relations* factor (trust, liking, and future interaction). Indices were created by summing the responses of each item within each factor. The scales were internally consistent (see Table 2) and were treated as separate dependent measures.

Table 2 also presents the means and intercorrelations of the attitudinal responses and the three measures of cooperation, as well as the size of the penance requested in the open penance conditions. Relations and positive emotions were significantly related to cooperative choices, as were cooperative choices after the breach but prior to penance. More notable are the significant negative correlations between the relations scale and both the penance

requested (following offers of open penance) and end-game cooperative choices. Those who asked for more penance were ultimately less cooperative and held out worse prospects for the future interaction. Those who asked for less, in contrast, were ultimately more cooperative and had a more positive outlook on the interaction.

Behavioral and Affective Findings

Once the final period had been announced, game theory suggests that rational players will choose noncooperatively on all of the last five rounds. In fact, there were many cooperative choices (the grand mean was 2.89). Most participants (173 of 224; 77.2%) made at least one cooperative choice during this period; 37.1% cooperated on the very last round.

Analysis of the fixed penance conditions indicated a significant main effect for penance on both cooperative behavior and positive emotions ($F(2,130) = 10.20, p < 0.001$, and $F(2,130) = 8.71, p < 0.001$, respectively; see Table 3). Mere talk had a positive effect but led to significantly fewer cooperative choices and less positive emotion than either small or large penance. Cooperation on the last round and the frequency of fully noncooperative endgame choices follow this same pattern. Thus, the data contradict game theory's expectations. Mere talk even led to cooperative choices on the last round (24% cooperated). All of these findings clearly support Hypothesis 1. Denials of intent, however, were not more effective than acknowledging intent, either for affective reactions or cooperative behavior. This contradicts Hypothesis 2.

These data also support Hypothesis 3, that substantive

Table 2 Descriptive Statistics and Intercorrelations of the Dependent Measures for Fixed and Open Forms of Penance

	Form					Intercorrelations						
	Max	Fixed Mean	SD	Open Mean	SD	1	2	3	4	5	6	7
1 Initial Cooperation	5	4.67	0.95	4.54	0.99	—	-0.06	0.07	0.06	0.26*	0.36**	0.28
2 Extinction Trials	6	1.95	1.63	2.19	1.26	-0.11	—	0.07	0.13	0.16	-0.03	0.04
3 Final Cooperation	5	2.89	1.83	2.88	1.85	0.06	0.19	—	-0.03	0.32**	-0.14	-0.48**
4 Expected Relations	15	9.18	3.15	11.07	2.90	-0.08	0.06	0.24**	—	-0.15	0.14	-0.52**
5 Positive Emotions	15	10.25	3.00	9.87	2.80	0.07	0.01	0.23**	0.10	—	0.02	0.20
6 Negative Emotions	25	13.8	4.35	15.42	3.92	0.14*	0.01	-0.04	0.03	0.07	—	0.24
7 Penance Requested	5	—	—	2.30	1.32	—	—	—	—	—	—	—

Note. Fixed penance condition correlations are printed below the main diagonal ($n = 142$); open penance conditions are printed above the main diagonal ($n = 83$). Coefficient alphas for the fixed and open penance conditions were 0.86 and 0.78 for positive emotions, 0.71 and 0.67 for negative emotions, and 0.88 and 0.89 for expected relations. Also, for the relations index, smaller numbers indicate a more positive relations; for emotions, larger numbers indicate more positive and more negative emotions.

* $p < 0.05$; ** $p < 0.01$.

Table 3 Mean Frequency of Endgame Cooperation, Final Round Cooperation, and Complete Endgame Noncooperation as a Function of the Different Forms of Penance

Penance Condition	Mean Cooperative Choices	Last Round Cooperations (%)	Complete Noncooperations (%)	Mean Positive Emotions [#]
<i>Fixed Form Conditions</i>				
Mere Talk	2.16 _b	12/50 (24.0%)	20/50 (40.0%)	9.32 _b
Small Penance	3.50 _a	16/44 (36.4%)	5/44 (11.4%)	10.91 _a
Large Penance	3.32 _a	22/47 (46.8%)	7/47 (14.9%)	11.03 _a
<i>Open Form Conditions</i>				
Open/No Ask	1.15 _b	6/26 (23.1%)	15/26 (57.7%)	8.85 _c
Open/Ask	3.00 _a	6/17 (35.3%)	4/17 (23.6%)	9.82 _b
Responsible/Open	3.85 _a	21/40 (52.5%)	0/40 (0.0%)	10.55 _a

Note. Means within the fixed or open conditions with common subscripts are *not* significantly different from one another using the Newman-Keuls procedure ($p < 0.05$).

[#]These means ranged from 3 to 15, with larger means signifying more positive emotions.

offers of penance would be more effective than bald apologies. Cooperation and positive affect were significantly greater in the small and large penance conditions than they were for mere talk. The data do not support Hypothesis 4, however, which predicted that larger offers of penance would be more effective than smaller offers. Instead, the data document a sharp difference between substantive and nonsubstantive penance.

Analysis of the open conditions indicated significant main effects for both final cooperative choices and positive emotions for penance ($F(1,71) = 21.73, p < 0.001$ and $F(2,71) = 3.79, p < 0.05$, respectively), as well as two interactions for final cooperative choices, between explanations and duration ($F(1,71) = 4.28, p < 0.05$) and explanations and penance ($F(2,71) = 4.00, p < 0.05$). The means for the main effects (see Table 3) indicate that the responsible/open condition led to somewhat more cooperation than the open/ask condition and both led to significantly more cooperation and positive emotion than the open/no ask condition. The effects for final round cooperation and complete endgame noncooperation also suggest stronger positive effects for responsible/open offers.

The significant interactions support the behavioral portions of Hypotheses 5 and 6: The explanations by duration interaction (see Table 4) show that acknowledging noncooperative intent was more effective in reestablishing cooperation in short interactions and that denying noncooperative intent was more effective (but not significantly) in reestablishing cooperation in longer interactions. The data also suggest that short interactions led participants who were offered open penance to be more

sensitive to expressed intentions than similar participants in longer interactions.

The explanations by penance interaction (see Table 4) indicate that open penance had its strongest behavioral effects when it accompanied a credible message that denied responsibility. As previously noted, some of the original open offers failed to generate a request for penance. When denials of intent led to no response to an open offer of penance, the likelihood of future cooperation was extremely low (a mean of 0.73 cooperative choices in the last five rounds), suggesting that these participants did not believe the denials. In contrast, participants who requested penance or those who received a responsible/open offer of penance were more likely to cooperate following a denial. Thus, it appears that believable denials in the open penance conditions were more effective in reestablishing cooperation than acknowledgments, supporting the thrust of interactional justice research and demonstrating its impact on actual behavior. Simultaneously, however, denials meant greater risk: If they were not believed, they appear to have been less effective than acknowledgments of intent.

Hypothesis 7 predicted that open offers of penance would be more effective than fixed offers. The data indicate that, overall, open offers led to a mean of 2.83 cooperative choices and that fixed offers led to a mean of 2.96 cooperative choices. There were also no significant differences in participants' affective reactions. Thus, Hypothesis 7 was not supported.

Hypothesis 8 predicted that open offers of penance would generate requests for repayment that were smaller than the costs originally imposed by the harm. The data

Table 4 Analysis of the Impact of the Open Form of Penance on Cooperative Choice

(a) Mean cooperative choices under the open form of penance as a function of explanation and duration					
Explanation	Short		Long		
Acknowledge	3.60 _a		2.48 _b		
SD	1.69		1.57		
<i>n</i>	20		21		
Deny	2.32 _b		3.2 _a		
SD	2.03		1.88		
<i>n</i>	22		20		

(b) Mean cooperative choices and simple effects analyses for the explanation by open form interaction					
Explanation	Form			<i>F</i> ¹	<i>p</i>
	Open No Ask	Open Ask	Responsible Open		
Acknowledge	2.09	2.73	3.74	5.07	0.05
SD	-2.11	-2.05	-0.73		
<i>n</i>	11	11	19		
Deny	0.73	3.5	3.95	23.8	0.001
SD	-1.49	-1.87	-0.97		
<i>n</i>	15	6	21		

Note. Means with different subscripts differ significantly ($p < 0.05$).
¹df = (2, 71).

indicate that in the original open condition the mean request was for 0.42 maximum payoffs (i.e., s/he chooses Down and the program chooses Left), including many nonresponses. In the responsible/open condition, the mean was 1.45 maximum payoffs requested. Analysis of the requests in the open/ask and responsible/open conditions yielded a main effect for duration, $F(1,49) = 5.29, p < 0.05$, and a significant interaction between intent and duration, $F(1,49) = 4.97, p < 0.05$. Short interactions led to larger average requests (2.63 maxima) than longer interactions (1.87), particularly for denials, where the mean requests for penance were 3.00 and 1.33, respectively. In contrast, acknowledging intent led to relatively consistent requests for penance (2.33 and 2.40 rounds). Given that all of the participants suffered two rounds of mutually noncooperative outcomes (payoffs of 0.03 rather than 3.6 from mutual cooperation) and each maximum outcome only provided payoffs of 4.0, Hypothesis 8 is clearly supported: They did not ask for enough to cover the costs that they had suffered.

Written Messages

Throughout the experiment, most participants used if-then statements suggesting contingent or strategic moves: of 182 participants who sent messages, 138 (76%) sent if-then statements before the breach, 135 (74%) sent if-then statements just after the breach, and 151 (83%), a slight increase, sent them after penance was offered. No systematic patterns between the conditions are discernible.

Forty-nine participants in the short interactions (50.5%) made economic arguments, compared to 35 (41.2%) in the long. Disparaging remarks were considerably more frequent in the long interactions ($n = 34; 40%$) than in the short ($n = 24; 24.7%$), as were accusations of lying ($n = 13$ in longer interactions, $n = 3$ in shorter interactions). These data suggest that breaches in longer interactions led to more emotional reactions than breaches in shorter interactions.

Participants' messages after the breach were frequently angry or provocative. In addition to illuminating their feelings, they also provide a check on the effectiveness of the explanation and duration manipulations. While many denials prompted incredulous responses (e.g., "Yeah, right. I bet he took the wrong card"), none of the acknowledged defections led to such vehement reactions.

Trust was frequently mentioned in their messages, but rarely before the breach or during the endgame. Instead, trust was most often mentioned after the breach. In the shorter interactions trust messages peaked in frequency at Round 13, just after the offer of penance. The message was generally a matter-of-fact reaction: "We had a trust. You should be more confident!" "Big loss for a little trust!" "Just as I was beginning to trust you!" In the longer interactions, the peak frequency occurred in Round 17, immediately following the breach. These comments questioned the counterpart's intelligence and morality, exemplifying the bewilderment described by Lewicki and Bunker's (1996) discussion of breakdowns in more established relationships. "How awful! We're against the other team not each other. You broke the trust factor." "What happened? Violation of trust compromises your integrity." "I expect your honesty in business. You are very foolish. If you'd maintained my trust, you'd have made a lot more money." These data are consistent with our earlier interpretation of the effects of denial and acknowledgment. In the shorter interactions, trust became important when penance was offered (i.e., it helped resolve uncertainty); in the longer interactions, trust appears to have been most important exactly when it was breached.

Both sets of messages fit Morrison and Robinson's

(1997) model of betrayal. A breach early in the interaction seemed to lead to cognitive rather than emotional reactions. A breach later in the interaction was more serious, prompting both cognitive and emotional reactions. The tenor of these messages conveys the fundamental significance of duration in the development of a cooperative interaction and its impact on the credibility of subsequent explanations (even in a short experiment).

In sum, the messages suggest that the short interactions led to more matter-of-fact and economically oriented responses and, possibly as a result, larger demands for penance. The longer interactions generated firmer expectations and more emotional reactions; they were obviously successful in generating a stronger connection between the participants.

Discussion

The results indicate that apologies and simple explanations were effective at reestablishing cooperation and increasing positive reactions (supporting Hypothesis 1) but that, overall, denials were no more effective than acknowledgments of noncooperative intent (rejecting Hypothesis 2). Substantive amends had more positive effects than explanations (supporting Hypothesis 3), but small offers of penance were as effective as larger offers (rejecting Hypothesis 4). Acknowledgments were more effective than denials in short interactions and denials were more effective than acknowledgments in longer interactions (supporting Hypothesis 5 and Hypothesis 6). Mere talk was sufficient to generate cooperation sometimes even on the last trial, a particularly stringent test because there is no opportunity for further reciprocity. Open offers of penance were not more effective than fixed offers (rejecting Hypothesis 7), although they did elicit requests for penance that were less than the injured parties' costs (supporting Hypothesis 8). This is consistent with evidence on the mild responses to similar types of offerings made in oral histories of clan-based societies (Boehm 1984) and among other species (de Waal 1989).

Responsible/open offers led to positive affect and considerable cooperation, albeit with considerable risk. These results show that explanations can influence behavior as well as perceptions and attitudes. They also document the important positive effects of substantive and open offers of penance on the reestablishment of cooperation. Once breached, cooperation can be rebuilt. Actions, explanations, and apologies all augment the process in important and sometimes subtle ways.

The form taken by different offers of penance was an important factor in this experiment, having a strong impact on all of the dependent measures. Some of the effects

were predicted, including the positive impact of fixed penance on cooperative behavior and positive emotions. Other effects were surprising. Offers to take a small fixed penalty proved just as effective in rebuilding cooperation as larger offers. Open offers required linguistic finesse to be effective: When they implied an acceptance of responsibility, they were much more effective in rebuilding cooperation.

Previous studies have shown that an explanation that denies intent can mute negative reactions (e.g., Shapiro et al. 1994). This study also documented the effectiveness of denials, but only in particular conditions: Believable denials were somewhat (but not significantly) more effective than acknowledgments. They were also more effective in longer interactions and less effective when they were not coupled with substantive penance.

On average, participants continued cooperating for just over two rounds after their counterpart had defected. On the one hand, this suggests some tolerance on their part; on the other, it gave them more zero payoffs than we had intended (an average of four) as well as two small payoffs (of 0.03, for mutual noncooperation) prior to the offer of penance. Because the fixed penance offers were only one or two rounds of maximum payoffs, they could not make up for participants' losses. Instead, penance in this experiment seemed to represent a signal of a genuine commitment toward mutual cooperation (e.g., Pilisuk et al. 1967). The fact that penance was substantive, however minimally, seems to have been critical, as it helped generate considerably more cooperation than mere talk.

Extrapolations from anthropological studies led us to predict that open offers of penance would be particularly effective. Our original manipulation, however, was only successful when it prompted a response. Subtle rewording of the offer from "What will it take?" to "What can I do?" had important effects on both cooperation and positive affect, as well as on the victims' requests for penance. It led to the highest frequencies of overall and last-round cooperation that we observed. These results reinforce the notion that forgiveness and reestablishing mutual cooperation depend on both the penitent and the victim because the participants who did respond to "What will it take?" offers were also relatively cooperative in the endgame. The fact that open offers of penance were no more effective than small, fixed offers of penance also reinforces the notion that the symbolic nature of substantive penance may be more important than its size.

The effects of mere talk were material: A simple apology led to quite a few cooperative choices on the last five rounds (24%; see Table 3) and enhanced the positive affect. These effects may be due to the fact that mutual cooperation in these interactions provided participants

with considerable potential gain. Interactions where cooperation is less profitable and riskier may well reduce the effectiveness of talk on its own. The potential generality of the current results could be augmented considerably by additional research that varies the context. For a start, this might mean investigating other sets of payoffs. Although we manipulated five different types of penance, we studied only one set of payoffs in a single experiment. Other payoff contingencies might lead to different effects.

Future research might also include direct measures of forgiveness. We did not use such measures here because we hoped to maintain the illusion of unprogrammed, real counterparts. Prior research on the correlates of forgiveness indicated that positive affect, negative affect, and future intentions could provide indirect indicators of forgiveness. The results do indicate that penance influenced positive affect and that affect was correlated in the expected directions with endgame cooperation. The affective indicators, however, do not directly track levels of endgame cooperation. This could mean that the act and state of forgiveness is sufficient but not necessary to rebuilding cooperation in interactions like these. This conclusion must be hedged, though, because it is also likely that the affect measures used here may not have been sufficiently sensitive to detect temporal variations in affect that would characterize forgiveness. Participants in this study responded to global indicators of positive and negative affect experienced over the entire set of plays. Future research should employ more direct measures of forgiveness and more temporally sensitive indicators of affect.

Other limitations of the current study reflect some of the general limitations of experimental research, i.e., short, straightforward interactions that are not based on face-to-face experience, relatively small payoffs, and knowledge that one is behaving in an experiment. More serious breaches of cooperation, however, would strain the bounds of ethical research. Thus, keeping things contained within an experimental setting allowed this study to be conducted. It also allowed us to measure a variety of cognitive and emotional reactions, as well as actual behavior.

Conclusions

Effective relationship management is critical to success in business, inside organizations, and even in politics. One of the more subtle and difficult elements of relationship management is the inevitable friction that arises and threatens ongoing cooperation. Our analysis of the dynamics of defection and penance began with a focus on

the type of explanation provided by the party that initiated noncooperative action. Past research has documented the benefits of external explanations in blunting negative perceptions. We found a similar pattern—for those who had been involved in relatively long, cooperative interactions. In short interactions, acknowledgments actually proved more effective at blunting negative reactions. The repeated dilemma framework allowed us to extend these findings beyond perceptions, showing that the proper combinations of explanation and substantive penance can contribute to reestablishing actual cooperative behavior. This research also identifies some forms of substantive penance that can be critical to this process.

These findings highlight the tight connection between the nature of an interaction and the explanations that are needed to rebuild cooperation following a breach. Defectors were most effective at restoring cooperation in short interactions when they acknowledged intent and also presented an effective, open offer of penance (i.e., one that generated a response from the participant). This strategy led to smaller requests for penalties. With breakdowns in longer interactions leading to emotional reactions, denials of intent were more effective, even the same type of denial that was so easily questioned in shorter interactions. The strength of the reactions observed here, as well as the behavioral effects, suggests that the duration of a positive interaction is very important. People do not expect or value the same type of explanation from a new acquaintance or a new colleague as they do from someone whom they have known for a long time. Denials of unexpected negative actions are more easily accepted when they come from a well-known other. In contrast, accepting blame may create better prospects for future cooperation in short interactions. By admitting a mistake, a person acknowledges their fallibility and may even generate increased attraction. Similar actions taken by organizational members following negative outcomes also appear to be tremendously effective (cf. Elsbach 1994).

From an interactional justice perspective, the current findings are counterintuitive. Previous justice research has focused on explanations that are meant to deny either intent or foreseeability. The current findings show that in particular situations, i.e., first-time short interactions, open admissions of intent may be more effective than weak attempts at denial. This opens new avenues for research on the boundary conditions of the effectiveness of admissions of guilt. In particular, future research might attempt to identify when admissions of intent will be more effective than adequate denials and exactly when, over the course of an interaction, denials become more effective than admissions.

The data further show that some form of substantive

penance greatly enhances an explanation's effectiveness. The nature of penance is critical to current and future interactions. We originally hypothesized that open offers of penance would provide the ideal mechanism for creating voice and modulating egocentric perceptions of fairness. Although they were effective when they generated a response, the frequent breakdowns in the original open condition were particularly revealing. In a difficult interpersonal atmosphere, offers of penance that seem insincere (the open/no ask condition) can further increase antagonism and accentuate a breakdown in cooperation. In contrast, simply volunteering small amounts of substantive penance was particularly effective. Experience with similar (uncontrolled) breaches in classroom prisoners' dilemma exercises (Malhotra and Murnighan 2001) suggests that the voluntary nature of penance may be critical. Requests for penance by wronged parties may be much less effective in restoring future cooperation than spontaneous offers of penance by defectors. This hypothesis is worthy of future research.

In this study, most of the requests for penance in the open condition approximated the penance paid in the fixed, small penance condition. This similarity may not be coincidental because it seems that *any* substantive penance may be sufficient to resurrect cooperation. Future research might establish whether and how much this observation generalizes because a limit (e.g., breaches that cause truly serious losses) seems likely.

This study was formulated to investigate the effects of opportunistic actions on cooperation and, more importantly, to determine whether different forms of penance might be effective in restoring mutual cooperation. The findings are clear: Small substantive offers of penance led to significant increases in future cooperation. Although the study did not directly address the issue of trust, participants often raised it. Rather than early or late in an interaction, issues of trust most often surfaced shortly after the breach. Messages about trust could indicate a genuine belief by the participants that they had an understanding with the other party. They might also be a verbal place marker for more complex messages. The current data call into question the claim that, once broken, trust cannot be repaired (Rempel et al. 1985). They also open avenues for further research to address the long-run costs of different kinds of breaches.

The data also invite new conceptualizations of the dynamics of forgiveness, penance, and reconciliation. Because conflict is an inevitable human state, violated expectations are a normal part of interpersonal interaction. This study suggests that acknowledging responsibility for a harmful act can be quite effective in starting the process of reconciliation, especially when denial might prompt

disbelief. Voluntary substantive offers of penance seem to provide particularly clear and strong foundations for rebuilding cooperation.

These results are particularly pertinent given the growing importance of managing complex interpersonal ties within and between organizations and their members. Strong and weak ties have direct influences on both an individual's organizational success and on the healthy maintenance of organizational alliances (Krackhardt 1992). Given the findings that show that the expression of interindividual conflicts can contribute to the breakdown of broader organizational ties (Larson 1992), rebuilding cooperation may often be necessary for organizational success.

The optimistic implication of our results is that reestablishing cooperation is clearly possible. Joint ventures, cross-functional collaborations, and intergroup coordination need not break down irrevocably because of conflict between individuals. Questions about the restoration of trust and the avoidance of potential breakdowns remain for future research. The current findings provide both hope and direction by identifying some of the subtleties and nuances of reestablishing cooperation. A difficult barrier to overcome may be the natural disinclination of individuals who have violated other's expectations to willingly offer and accept penance. As the current research suggests, such entreaties must also prompt positive reactions from victims. The need for harmdoers to adopt a sequence of substantive, subtle action following the breach of a positive, cooperative interaction may pose a serious obstacle to the reconciliation of cooperative interactions and the reestablishment of trust. As the importance of relational ties increases, both within and between organizations, these and a whole host of related issues will become even more critical.

Acknowledgments

The research in this paper was supported by National Science Foundation Grant SES 9210114 to the first author and by grants from the Social Sciences and Humanities Research Council of Canada and from the Dispute Resolution Research Center at the Kellogg Graduate School of Management, Northwestern University to the fourth author. This project was conceived while the authors attended the Summer Institute on Dispute Resolution at the Center for Advanced Study in the Behavioral Sciences; funds for support of the institute were provided by the Andrew W. Mellon Foundation. The authors thank Derek Lohse, Dan Tang, Christina Wan, and Cindy Grushin for research assistance. They also wish to acknowledge their gratitude for the many helpful comments received from Max Bazerman, David McPhillips, Gary Miller, Greg Oldham, Madan Pillutla, Scott Poole, Michael Roloff, Bryan Routledge, and three anonymous reviewers.

References

- Arno, A. 1976. Ritual of reconciliation and village conflict management in Fiji. *Oceania* XLVII 49–65.

- Axelrod, R., W. Hamilton. 1981. The evolution of cooperation. *Sci.* **211** 1390–1396.
- Bettenhausen, K., J. K. Murnighan. 1985. The emergence of norms in competitive decision making groups. *Admin. Sci. Quart.* **30** 350–372.
- Bies, R. J., J. S. Moag. 1986. Interactional justice: Communication criteria of fairness. R. J. Lewicki, B. H. Sheppard, and M. H. Bazerman, eds. *Research on Negotiations in Organizations*, JAI Press, Greenwich, CT, 43–55.
- , D. L. Shapiro. 1987. Interactional fairness judgments: The influence of causal accounts. *Soc. Justice Res.* **1** 199–218.
- , T. M. Tripp. 1996. Beyond distrust: Getting even and the need for revenge. R. M. Kramer and T. R. Tyler, eds. *Trust in Organizations*. Sage, Thousand Oaks, CA.
- , D. L. Shapiro, L. L. Cummings. 1988. Causal accounts and managing organizational conflict: Is it not enough to say it's not my fault. *Comm. Res.* **15** 381–399.
- Boehm, C. 1984. *Blood Revenge: The Anthropology of Feuding in Montenegro and Other Tribal Societies*. University of Kansas Press, Lawrence, KS.
- Bottom, W. P. 2001. Testing the limits of rationality: Social accounting, cognitive illusions, and the unenforceable provisions of the treaty of Versailles. Working paper, John M. Olin School of Business, Washington University, St. Louis, MO.
- , C. L. Eavey, G. J. Miller. 1996. Getting to the core: Coalitional integrity as a constraint on the power of agenda setters in decision making groups. *J. Conflict Resolution* **40** 298–319.
- Bradfield, M., K. Aquino. 1999. The effects of blame attributions and offender likableness on forgiveness and revenge in the workplace. *J. Management* **25** 607–631.
- Brockner, J., B. Weisenfeld. 1996. An integrative framework for explaining reactions to decisions: Interactive effects of outcomes and procedures. *Psych. Bull.* **120** 189–208.
- Child, J., R. G. McGrath. 2001. Organizations unfettered: Organizational form in an information-intensive environment. *Acad. Management J.* **44** 1135–1148.
- Elsbach, K. D. 1994. Managing organizational legitimacy in the California cattle industry: The construction and effectiveness of verbal accounts. *Admin. Sci. Quart.* **39** 57–88.
- Farrell, J., R. Gibbons. 1989. Cheap talk with two audiences. *Amer. Econom. Rev.* **79** 214–223.
- Folger, R. 1993. Reactions to mistreatment at work. J. K. Murnighan, ed. *Social Psychology in Organizations: Advances in Theory and Research*. Prentice-Hall, Englewood Cliffs, NJ, 161–183.
- Goffman, E. 1971. *Relations in public: Microstudies of the public order*. Basic Books, New York.
- Hayashi, N. 1993. From TIT-FOR-TAT to OUT-FOR-TAT. *Sociological Theory and Methods*. **8** 19–32.
- , E. Ostrom, J. Walker, T. Yamagishi. 1999. Reciprocity, trust, and the sense of control: A cross-societal study. *Rationality and Soc.* **11** 27–46.
- Heide, J. B., A. S. Miner. 1992. The shadow of the future: Effects of anticipated interaction and frequency of contact on buyer-seller interaction. *Acad. Management J.* **35** 265–291.
- Hirschman, A. O. 1970. *Exit, Voice, and Loyalty*. Harvard University Press, Cambridge, MA.
- Kahn, L., J. K. Murnighan. 1993. A general experiment on bargaining in demand games with outside options. *Amer. Econom. Rev.* **83** 1260–1280.
- Koch, K. F., S. Altorki, A. Arno, L. Hickson. 1977. Ritual reconciliation and the obviation of grievances: A comparative study in the ethnography of law. *Ethnology* **16** 269–283.
- Komorita, S. S., J. Mechling. 1967. Betrayal and reconciliation in a two-person game. *J. Personality and Soc. Psych.* **6** 349–353.
- , C. D. Parks. 1999. Reciprocity and cooperation in social dilemmas: Review and future directions. D. V. Budescu, I. Erev, and R. Zwick, eds. *Games and Human Behavior: Essays in Honor of Amnon Rapoport*. Lawrence Erlbaum, Mahwah, NJ, 315–330.
- Konovsky, M. A., F. Jaster. 1989. 'Blaming the victim' and other ways business men and women account for questionable behavior. *J. Bus. Ethics* **8** 391–398.
- Krackhardt, D. 1992. The strength of strong ties: The importance of Philos in organizations. R. Eccles, and N. Nohria, eds. *Networks and Organizations: Structure, Form and Action*. Harvard Business School, Cambridge, MA, 216–239.
- Kramer, R. M., M. B. Brewer, B. A. Hanna. 1996. Collective trust and collective action: The decision to trust as a social decision. R. M. Kramer and T. R. Tyler, eds. *Trust in Organizations: Frontiers of Theory and Research*. Sage, Thousand Oaks, CA, 357–389.
- Larson, A. 1992. Network dyads in entrepreneurial settings: A study of the governance of exchange processes. *Admin. Sci. Quart.* **37** 76–104.
- Lewicki, R. J., B. B. Bunker. 1996. Developing and maintaining trust in work relationships. R. M. Kramer and T. R. Tyler, eds. *Trust in Organizations: Frontiers of Theory and Research*. Sage, Thousand Oaks, CA, 114–139.
- , D. J. McAllister, R. J. Bies. 1998. Trust and distrust: New relationships and realities. *Acad. Management Rev.* **23** 438–458.
- Lindskold, S. 1978. Trust development, the GRIT proposal, and the effects of conciliatory acts on conflict and cooperation. *Psych. Bull.* **85** 772–793.
- , B. Betz, P. S. Walters. 1986. Transforming competitive or cooperative climates. *J. Conflict Resolution* **30** 99–114.
- , P. S. Walters, H. Koutsourais. 1983. Cooperators, competitors, and response to GRIT. *J. Conflict Resolution* **27** 521–532.
- Lowin, A. 1968. Participative decision making: A model, literature critique, and prescriptions for research. *Organ. Behavior and Human Performance* **3** 68–106.
- Malhotra, D., J. K. Murnighan. 2001. Tragedy in the trenches: The ineffectiveness of uninformed strategic choices in intergroup prisoners' dilemmas. Working paper, Northwestern University, Evanston, IL.
- McCormack, S. A., T. R. Levine. 1990. When lies are uncovered: Emotional and relational outcomes of discovered deception. *Comm. Monographs* **57** 119–139.
- Mehlman, R. C., C. R. Snyder. 1985. Excuse theory: A test of the self-protective role of attributions. *J. Personality and Soc. Psych.* **49** 994–1001.
- Mills, C. 1995. Politics and manipulation. *Soc. Theory and Practice* **2** 97–112.
- Morrison, E. W., S. L. Robinson. 1997. When employees feel betrayed: A model of how psychological contract violation develops. *Acad. Management Rev.* **22** 226–256.
- Murnighan, J. K. 1981. Defectors, vulnerability, and relative power:

- Some causes and effects of leaving a stable coalition. *Human Relations* **34** 589–609.
- . 1991. *The Dynamics of Bargaining Games*. Prentice Hall, Englewood Cliffs, NJ.
- , A. E. Roth. 1983. Expecting continued play in prisoner's dilemma games: A test of several models. *J. Conflict Resolution* **27** 279–300.
- North, J. 1987. Wrongdoing and forgiveness. *Philosophy* **62** 449–508.
- Pilisuk, M., J. A. Winter, R. Chapman, N. Haas. 1967. Honesty, deceit, and timing in the display of intentions. *Behavioral Sci.* **12** 205–215.
- Pillutla, M. M., J. K. Murnighan. 1995. Being fair or appearing fair: Strategic behavior in ultimatum bargaining. *Acad. Management J.* **38** 1408–1426.
- Powell, W. W., K. W. Koput, L. Smith-Doerr. 1996. Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Admin. Sci. Quart.* **41** 116–145.
- Rempel, J. K., J. G. Holmes, M. P. Zanna. 1985. Trust in close relationships. *J. Personality and Soc. Psych.* **49** 95–112.
- Roth, A. E., J. K. Murnighan. 1978. Equilibrium behavior and repeated play in prisoners' dilemma games. *J. Math. Psych.* **17** 189–198.
- Rousseau, D. M., J. McLean Parks. 1993. The contracts of individuals and organizations. L. L. Cummings and B. M. Staw, eds. *Research in Organizational Behavior*. JAI Press, Greenwich, CT, 1–43.
- Schenkel, R. 1967. Submission: Its features and functions in the wolf and dog. *Amer. Zoologist* **7** 319–329.
- Schlenker, B. R. 1980. *Impression Management*. Brooks/Cole, Monterey, CA.
- Schwartz, G., T. Kane, J. Joseph, J. T. Tedeschi. 1978. The effects of remorse on the reactions to a harm-doer. *British J. Soc. Clinical Psych.* **17** 293–297.
- Scott, M. B., S. M. Lyman. 1968. Accounts. *Amer. Sociological Rev.* **33** 46–62.
- Shapiro, D. L. 1991. The effects of explanations on negative reactions to deceit. *Admin. Sci. Quart.* **36** 614–630.
- , E. H. Buttner, B. Barry. 1994. Explanations: What factors enhance their perceived adequacy? *Organ. Behavior and Human Decision Processes* **58** 346–368.
- Sigmund, E. W. 1999. Ho'oponopono: A traditional Hawaiian practice for setting relationships right. *Trans. Anal. J.* **3–4** 373–392.
- Stahl, I. 1972. *Bargaining Theory*. Economics Research Institute, Stockholm, Sweden.
- Thompson, L., G. Loewenstein. 1992. Egocentric interpretations of fairness and interpersonal conflict. *Organ. Behavior and Human Performance* **51** 176–197.
- Thoreson, C. E., F. Luskin, A. H. S. Harris. 1998. Science and forgiveness interventions: Reflections and recommendations. E. L. Worthington, Jr., ed. *Dimensions of Forgiveness*. Templeton Foundation Press, Philadelphia, PA.
- de Waal, F. 1989. *Peacemaking Among Primates*. Harvard University Press, Cambridge, MA.
- Wilson, J. 1988. Why forgiveness requires repentance. *Philosophy* **63** 534–535.
- Witvliet, C. V., T. E. Ludwig, K. L. Vander Laan. 2001. Granting forgiveness or harboring grudges: Implications for emotion, physiology, and health. *Psych. Sci.* **12** 117–123.