
**TRUST AND DISTRUST
IN ORGANIZATIONS**

DILEMMAS AND APPROACHES

RODERICK M. KRAMER AND KAREN S. COOK
EDITORS

VOLUME VII IN THE RUSSELL SAGE FOUNDATION SERIES ON TRUST

Russell Sage Foundation • New York

Whitener, Ellen M., Susan E. Brodt, M. Audrey Korsgaard, and Jon M. Werner. 1998. "Managers as Initiators of Trust: An Exchange Relationship Framework for Understanding Managerial Trustworthy Behavior." *Academy of Management Review* 23(3): 513-30.

Williams, Michelle. 2001. "In Whom We Trust: Group Membership as an Affective Context for Trust Development." *Academy of Management Review* 26(3): 377-96.

Chapter 12

Paradoxes of Trust: Empirical and Theoretical Departures from a Traditional Model

J. KEITH MURNIGHAN, DEEPAK MALHOTRA, AND J. MARK WEBER

THE CONCEPT of exchange is central to social action (Homans 1961). In essence, anything that we might label as "social" depends on the give-and-take of more than a single individual. Many of these exchanges are informal or implicit, like the reciprocation of a smile upon meeting or the fulfillment of expectations that other drivers will conform to the driving norms of a particular locale. When exchanges become more formal or explicit, the parties typically reach agreements that regulate the transfer of objects or ideas between them, like the sale and delivery of goods in a market or the exchange of marriage vows. These kinds of exchanges are often accompanied by written documents, contracts, which stipulate many of the parameters of the exchange. However, what are often implicit in these contracts are the fundamental, sometimes value-laden expectations of the parties. In addition, in formal contracts it is literally impossible to cover all of the past, present, and future issues and concerns of the parties. Thus, every chemical and physical detail of the goods being exchanged is not stipulated in buyer-seller contracts, nor are all of the commitments that marriage partners expect of each other stated in their wedding vows.

Most people understand, at least implicitly, that contracts are limited and cannot be all-encompassing. As a result, formal exchanges often rely on a more generalized trust that extends beyond contractual obligations. This chapter begins by describing common conceptions of trust, based on early literature and everyday notions of the concept. We then

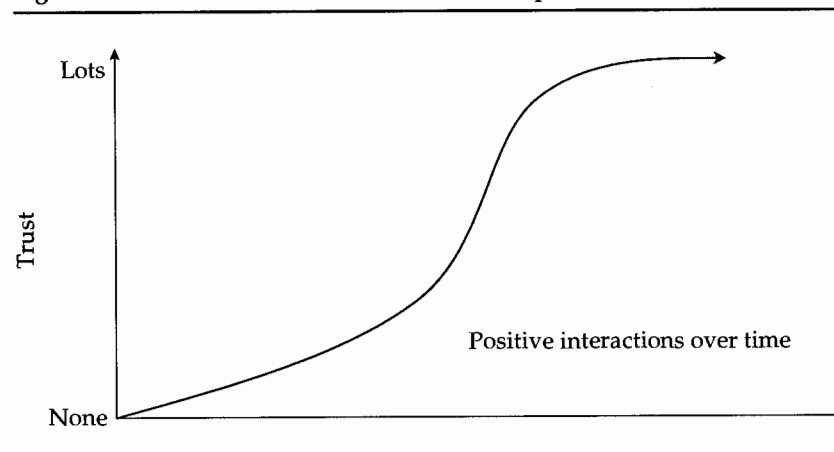
discuss various definitions of trust, choosing a recent definition as the basis for our subsequent discussion. In the remainder of the chapter, we address some converse, counterintuitive findings with regard to trust and trusting acts, some of which constitute paradoxes, in other words, sets of statements that are self-referential and contradictory, establishing a vicious cycle (Hughes and Brecht 1975). We focus most of our review on our own recent research: it shows that a rational-choice approach to trusting behavior does not provide a clear basis for behavioral predictions. We then present a preliminary model that expands current conceptualizations of trust and the processes that help establish it. We conclude by discussing a variety of theoretical and practical implications.

Common Conceptions of Trust

Most treatments of trust suggest that it builds slowly (see Kelley 1979; Rempel, Holmes, and Zanna 1985).¹ Repeated positive interactions between two (or more) parties contribute to attributions of benevolence that in turn engender trust (Mayer, Davis, and Schoorman 1995). The development of trust tends to be initially gradual but increasingly rapid, as each additional positive interaction becomes more valuable in establishing mutual trust (see figure 12.1). In essence, positive interactions form an accumulating foundation that allows both parties to have increasing confidence in each other's actions. After rising markedly, the impact of additional positive interactions becomes less important—only because trust may not be able to continue to increase exponentially. Even so, trust is likely to continue increasing, however marginally, as additional positive interactions accumulate.²

Common conceptions of trust suggest that the occurrence of a negative interaction where trust and positive outcomes were expected can be devastating. Once trust is broken, it may be lost completely and forever. According to John K. Rempel, John G. Holmes, and Mark P. Zanna (1985), not only is trust difficult to establish, it also "is doubly difficult to reestablish" (111). Figure 12.2 shows what happens when trust that has been built on long, repeated positive interactions has been violated. (There is an implicit assumption that some of these interactions entailed vulnerability and resulted from voluntary action by one party or the other.) The immediate result is the dramatic reduction or disappearance of trust, entering the realm of distrust. Now, the reestablishment of trust between the same parties becomes particularly problematic. The negative event must be cognitively resolved before trust can be reestablished, but because negative events carry considerably more weight than positive events, the task of trust reestablishment is considerably more difficult than the task of initial trust establishment.

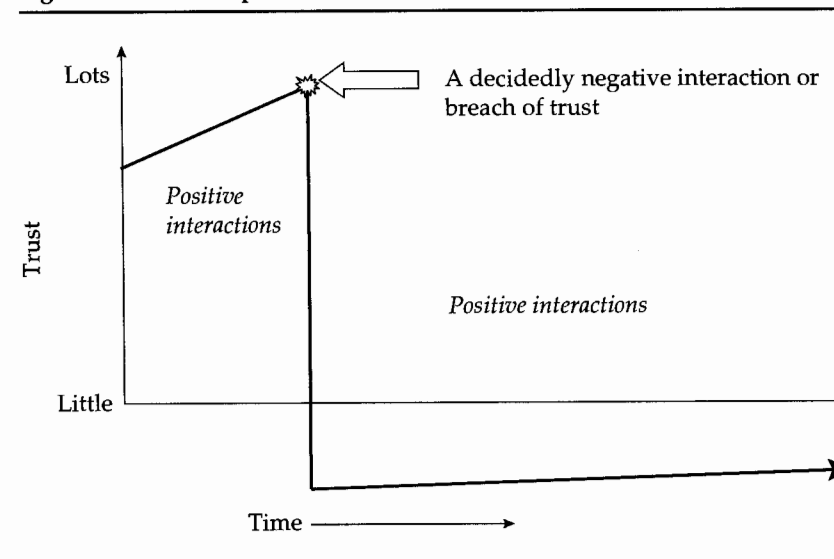
Figure 12.1 Traditional Model of Trust Development



Source: Authors' compilation.

This depiction of the trust formation process might be called "a rational-choice approach" (see Kramer 1999). In essence, it suggests that people do not take big risks in new relationships. Instead, trust builds slowly. It is almost as if people implicitly calculate the probabilities that their trust will be honored; if they have had repeated positive experi-

Figure 12.2 The Impact of a Breach of Trust



Source: Authors' compilation.

ences, they are willing to increase the stakes with greater confidence. This process matches rational models of decision making, and considerable data support this approach (Howard 1966; Rapoport and Mowshowitz 1966; Boyle and Bonacich 1970).

Although trust is obviously a critical issue in all sorts of exchanges, the literature seems to have paid only moderate attention to it until recently (see Kramer 1999). Recent progress on its definition and conceptualization has stimulated considerable research interest. Prior to turning to some of that research, we briefly summarize some of the recent conceptual discussions and present a definition of trust that helps guide our subsequent analyses.

Defining Trust

The psychological and organizational literatures have historically been populated by many different definitions of trust, and competing conceptualizations have at times been problematic. For instance, Roger C. Mayer, James H. Davis, and F. David Schoorman (1995) suggested that advances have been hindered by "a lack of clear differentiation among factors that contribute to trust, trust itself, and outcomes of trust" (711).

One of Mayer, Davis, and Schoorman's (1995) most important contributions was to focus on vulnerability as an essential element in trust. This theoretical emphasis built on Diego Gambetta's (1988) classic definition as well as earlier observations that the "willingness to take risks may be one of the few characteristics common to all trust situations" (Johnson-George and Swap 1982, 1306). More recently, in a special issue of *The Academy of Management Review* on trust (Rousseau et al. 1998), the "willingness to be vulnerable" was the most frequently cited definition of trust.

Mayer, Davis, and Schoorman (1995) were careful to point out that "trust is not taking risk per se, but rather it is a willingness to take risk" (712). They further distinguished trust from the related constructs of cooperation, confidence, and predictability. All three have a history of being confounded with trust, yet all three can quite clearly be distinguished from trust. For example, although cooperation sometimes involves risk, as in many social dilemmas, "trust is not a necessary condition for cooperation to occur, because cooperation does not necessarily put a party at risk" (713). Niklas Luhmann (1988) differentiates between confidence and trust by noting that trust requires the recognition and acceptance of risk. "If you do not consider alternatives (every morning you leave the house without a weapon!), you are in a situation of confidence. If you choose one action in preference to others in spite of the possibility of being disappointed by the action of others, you define the situation as one of trust" (102). Following Morton Deutsch (1958), Mayer, Davis, and Schoorman (1995) also note that trust is not pre-

dictability: "To equate the two is to suggest that a party who can be expected to consistently ignore the needs of others and act in a self-interested fashion is therefore trusted, because the party is predictable" (714). All three of these sometimes-associated constructs fail tests of necessity, sufficiency, or both.

After reviewing current research on trust in several disciplines, Rousseau et al. (1998) presented this definition: "Trust is a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behavior of another" (395). We use this definition as a foundation for our conceptual analysis of the dynamics of trust. At the same time, like other scholars we differentiate between trust and trusting actions. Whereas trust is a psychological construct that represents a willingness to risk, those who take trusting actions actually take those risks. Thus, trusting actions must necessarily include a potential for additional gain as well as the potential for loss. Reciprocity following a trusting action typically generates gain for the initial trustor; exploitation or even a simple nonresponse may cause a loss for the initial trustor.

As an empirical foundation for our analyses, we focus on a series of five recent research papers. Four represent our own work; the fifth comes from a decidedly different literature, and employs both different methods and a different general theoretical approach. All five, however, reach the same underlying conclusion: that observable instances of trust do not *necessarily* play out in a slowly increasing, rational fashion. Thus, these data provide the basis for a new conceptualization of the dynamics that can lead to and motivate trust.

We begin by summarizing a project on contracts and how they interact with the establishment of trust. We follow this with a summary of an ongoing project on the strategic interplay of executives and MBA students who might trust each other but often do not, to their mutual detriment. Then we turn to the discussion and description of a study that directly investigates the dissolution and potential rebuilding of trust. We follow with an investigation of the concept of reciprocity, one of the backbones of a rational-choice approach to trust. The final research project, which differs considerably from the rest, investigates the sexual choices of the members of committed gay couples in light of the threat of AIDS. Although this study offers results that are more socially jarring than the others, as part of this set of research endeavors it helps to provide the backdrop for some new theoretical ideas about the dynamics of trust.

Contracts and Trust

To varying degrees, all organizations face problems of agency, control, and uncertainty. Many of these problems stem from the simple fact that

the relationships between individuals and organizations largely represent mixed-motive interactions. On the one hand, many individuals feel a sense of organizational loyalty and identify with their organization and its goals. On the other, their own individual needs may be at odds with their organization's needs. Many organizations address these self-interest concerns by investing in costly technology to monitor employee behavior (Enzle and Anderson 1993). Companies also structure rules and incentive contracts to minimize the possibility of behavior that might conflict with organizational goals (Jensen and Meckling 1976). Employees also seek the security of employment contracts that specify their rights and responsibilities (Nye 1988). To this end, formal, binding contracts have become favored, routinized solutions to the problems of agency, control, and uncertainty in organizations and in many interpersonal domains as well. These formal agreements help reduce uncertainty (Williamson 1979), eliminate (certain kinds of) risk (Williamson 1996), enhance control (Klein 1993), and mitigate agency problems (Jensen and Meckling 1976).

Cooperation between individuals or between individuals and organizations can, however, develop without the use of contracts. When individuals trust each other, they are more likely to cooperate in strategic interactions (Mayer, Davis, and Schoorman 1995), share information in negotiations (Thompson 1991), and engage in mutually beneficial relationships (Siamwalla 1978). Trust helps to reduce uncertainty (Kollock 1994) and can lead to an increase in joint gains (Carnevale and Isen 1986). Trust can help solve agency problems (Das and Teng 1998), facilitate market processes (Arrow 1974), and lead to greater cooperation within and between organizations (Smith, Carroll, and Ashford 1995).

Contracts and trust can and do substitute for one another. Thus, a lack of trust suggests that the parties might want to create a contract to specify their rights and responsibilities (Coffrin and Cochran 1982; Sitkin and Roth 1993). Alternatively, when trust is strong, the parties may feel no need for the specifics or constraints of a contract. Instead, they may be able to fulfill a mutually beneficial agreement without having to resort to a contract's restrictions (Uzzi 1997).

Most approaches, whether slanted toward formal contracts or toward trust, acknowledge the value of both, even for the same agreement. Thus, even though contracts are important, they cannot possibly address all of the contingencies that might develop in a relationship (Parkhe 1998). This makes it necessary to cultivate trust. At the same time, it may be crucial for the parties not to underestimate the need to "get it in writing" (Nye 1988). Their task then becomes one of creating contracts while simultaneously cultivating trust. Our recent research investigates the dynamics and consequences of this process (Malhotra and Murnighan 2002).

Attributions of Trust

As we have noted, trust often depends in large part on the interaction history of the engaging parties (Boon and Holmes 1991; Lindsfold 1978; Swinth 1967; Solomon 1960). A cooperative, mutually beneficial relationship gives both parties confidence that they can trust each other.

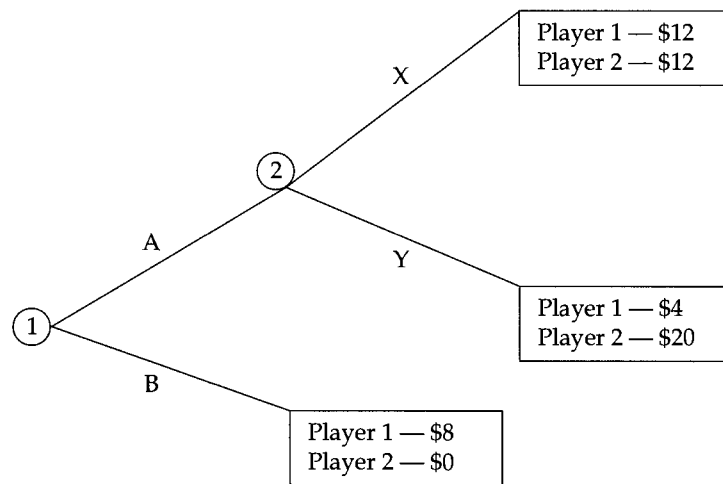
A potential difficulty with contracts is that the parties may attribute their resulting mutual cooperation to the contract, a situational attribution, rather than to each other. Contracts are likely to prompt situational attributions for two reasons. First, the situational constraint, the contract, is likely to be particularly salient. Shelley E. Taylor and Susan T. Fiske (1975, 1978) suggest that perceptually salient stimuli are more likely to be perceived as causal. Second, when there is high "consensus" (if it is thought that most people act this way), situational attributions become more likely (Kelley 1967). Contracts (and their threat of enforcement) typically generate compliance, that is, high consensus. Thus, cooperative behavior that is contractually mandated may not be seen as particularly indicative of an individual's dispositions. Also, even with a cooperative history, it is necessary that each party interpret the other's cooperation as being the result of their dispositions or motives rather than the result of other, situational factors. For instance, someone who has been obviously coerced to cooperate would not necessarily be perceived as trustworthy.

Contracts, then, may act as double-edged swords. On the one hand, they foster cooperation and help bring people together by reducing or eliminating the risk of exploitation. This is particularly true when the stakes are high and the risks of loss increase. In such interactions, contracts may be a necessary prerequisite for cooperation. On the other hand, they may make it difficult for trust to develop (Shapiro, Sheppard, and Cheraskin 1992; Sitkin and Roth 1993; Lewicki, McAllister, and Bies 1998).

Empirical Research on Contracts and Trust

Our research operationalized trust in the context of the Trust Game (Snijders 1996; Dasgupta 1988; Kreps 1990), an interaction between two people who each have two choices, made sequentially. Player 1s choose to "trust" or "not trust" player 2s. If player 1s choose not to trust, the game ends, and player 1s receive a moderate outcome and player 2s receive a small outcome (often nothing). If player 1s choose to trust, player 2s have the option of rewarding ("honoring the trust") or exploiting player 1s ("abusing the trust"). When trust is honored, both players receive a moderately high outcome. When trust is abused, player 2s maximize their personal gain to the detriment of player 1s, who receive

Figure 12.3 Structure of the Game



The figure is interpreted by reading it from left to right. Each node in the figure is a decision made by one of the players (identified by a circled number). Player 1 makes the first choice by choosing A or B. If player 1 chooses B, the game ends, since there are no more nodes after we follow the B path. The payoffs are in the boxes at the end nodes. If player 2 chooses A, then player 2 makes a choice of X or Y. The payoffs for each player are listed in the end nodes following each path.

Source: Authors' compilation.

less than if they had chosen not to trust. Figure 12.3 displays an extended form of the Trust Game.

In a first study (Malhotra and Murnighan 2002), we observed people playing the Trust Game repeatedly (three or four times) with the same partners. Cooperation between some dyads was facilitated in early rounds by allowing them to use contracts. A contract proposed by player 2s and accepted by player 1s meant that player 1s would necessarily choose to trust and player 2s would necessarily honor their trust. In their third and fourth interactions, contracts were no longer possible or no longer chosen. This provided a basis for comparing the trusting behaviors of player 1s with and without contracts. We predicted that the use of contracts would rob the parties of the opportunity to make trustworthy attributions regarding their counterparts and that the subsequent removal of contracts would have a damaging effect on trust. In other words, when player 2s could not or did not offer a contract after having done so in previous rounds, we predicted that player 1s would be less trusting than if they had not experienced contracts at all.

In this study, the removal of contracts was sometimes due to a choice by player 2s; other times contracts were removed exogenously. We predicted that player 2s' active choices would have a detrimental effect on trust. That is, trust would be least when player 2s rather than an outside force made the decision not to use a contract. Previous research indicates that people can accept a poor outcome more when they can attribute it to chance rather than to another person's choices (for example, Blount, Thomas-Hunt, and Neale 1996). Because player 1s are always better off with a contract, they may be less willing to accept the risk (to trust) when their counterpart no longer proposes one. Further, if player 2s have not proposed a contract when contracts are possible, player 1s might suspect that player 2s might exploit them.

The Experiment

All of the participants played the role of player 1 and made their choices via computer. Unknown to participants, there were no player 2s. Instead, player 2s' responses were programmed. Participants knew that one of their interactions would be randomly chosen to determine their monetary payoffs.

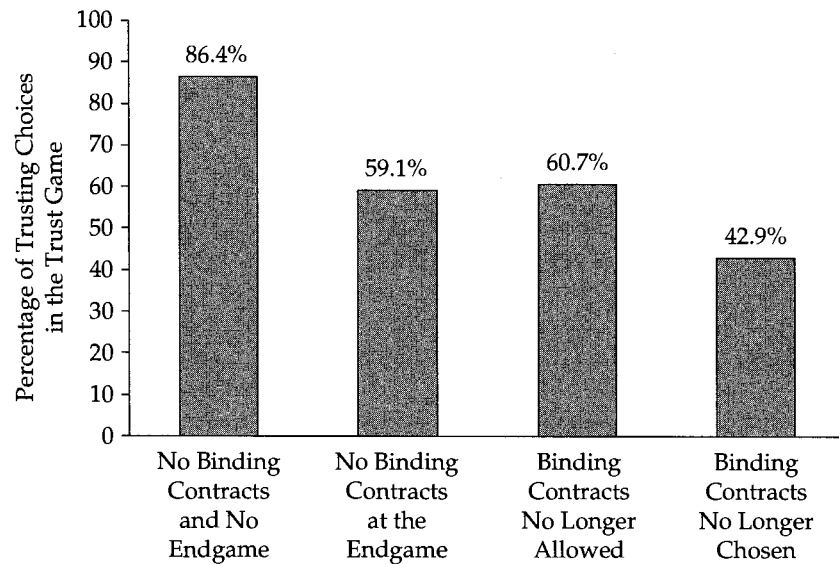
In a baseline condition, the "no-contracts" condition, participants interacted for four rounds with no mention of contracts. Player 1s simply made either "A" or "B" decisions, that is, to trust or not. Player 2s were programmed always to choose X, which was to honor trust. At the end of each round, the payoffs were announced. When the fourth round began, a message on the screen informed participants that this would be the last round. They were told that their counterpart also received this announcement.

In the other conditions, participants were told that player 2s could propose a contract. Player 2s' purported cost for proposing the contract was two dollars. Player 1s could accept or reject proposed contracts. It was in their best interest to accept. When they did, the computer automatically enforced the contract, ending the interaction and displaying a message that both players would receive twelve dollars for that interaction.

Player 2s proposed contracts in the first two rounds, but not in rounds 3 or 4. Player 1s were told that player 2s were either "not allowed" or had "chosen not" to propose a contract after the second round.

Results and Implications

In figure 12.4, player 1s chose to trust a great majority of the time (86.4 percent, or nineteen of twenty-two players) in their first three interactions, even with no mention of contracts. On the fourth and last round in the no-contracts condition, significantly fewer player 1s trusted (59.1 percent, or thirteen of twenty-two).

Figure 12.4 Data from Key Conditions in Malhotra and Murnighan (2002)

Source: Malhotra and Murnighan (2002).

Player 2s' active choices were powerful: when they chose not to propose contracts (in round 3) after having proposed contracts in the first two rounds, only 42.9 percent of the player 1s were willing to trust, significantly less than in round 3 of the no-contracts condition. In addition, there was significantly less trust after player 2s had offered contracts on rounds 1 and 2 and then chose not to (round 3) than in the first round of the no-contracts conditions. This suggests that, in this situation, a cooperative history under contracts results in less trust than no history at all.

When contracts were allowed and chosen (in the first two rounds) but then not allowed (in round 3), trust also dropped, to 60.7 percent. This was significantly less than the level of trust in the third round of the no-contracts condition (86.4 percent), suggesting that the removal of contracts, even when the decision is not in the hands of the other person, results in a drop in trust. In addition, as in the active-choice condition, there was significantly less trust after player 2s had offered contracts and then were not allowed to than in the first round of the no-contracts condition. Even without active choice, then, the use of contracts interfered with the development of trust.

These results suggest that binding contracts generate situational attributions for cooperation and can keep trust from developing. In addition, not only does trust have less chance to develop in a contractual inter-

action, but also any existing (presumptive) trust appears to diminish when contracts are employed.

Prevailing wisdom often encourages people to "get it in writing" and to "read the fine print" (see, for example, Nye 1988). This may be why the use of contracts in employment relationships is increasing (Morishima 1996). At the same time, recent research on trust consistently highlights the benefits of, and need for, cultivating trust (Coleman 1990; Fukuyama 1995; Kramer and Tyler 1996; Mayer, Davis, and Schoorman 1995; McAllister 1995; Putnam 1993; Misztal 1996; Seligman 1997; Sitkin and Roth 1993). Some authors have suggested that contracts might set the foundation for mutual trust (Coffrin and Cochran 1982; Lorenz 1999). The current findings provide a stark contrast, suggesting that one common prescription (contracts are helpful) might undermine the other (trust is essential). In essence, the use of binding contracts may make it difficult to cultivate trust. In addition, trust was less likely with a cooperative interaction history that was mandated by the presence of binding contracts than in cases with no interaction history at all. This suggests that contracts kept the interacting parties from seeing each other's cooperative behaviors as indicative of trustworthiness. As Gambetta (1988, 219) notes, "[I]f other people's actions were heavily constrained, the role of trust in governing our decisions would be proportionately smaller, for the more limited people's freedom, the more restricted the field of actions in which we are required to guess *ex ante* the probability of their performing them."

This study explored the relationship between explicit, formal arrangements and implicit, informal understandings. These two forms of agreement characterize many human relationships, particularly within organizations. The irony is that the relationship between the two is far from clean or simple. Whereas formal agreements such as contracts help to reduce risk and enhance the likelihood of cooperative interaction, they can work against the development of informal understanding and mutual trust. The converse may also be true: people with strong trust bonds may shy away from contracts, even though these may reduce their risks.

This is our first paradox: *Long-term cooperation may require the simultaneous use of contracts and trust, but the use of one can undermine the other.*

Trust Development in Strategic Interactions

In the contracts project just described, we focused on how binding contracts affected the trustors' perceptions of their counterparts. We now consider how a trusted party's own actions might affect their perceptions of the trustors. We investigated this issue in a recent study (Malhotra and Murnighan 2003) involving an iterated prisoner's dilemma framed as the gas station game (Murnighan 1991).

Table 12.1 The Profit Table in the Gas Station Game

		Team 2	
		Maintain Price (Cooperate)	Cut Price (Compete)
Team 1	Maintain price (cooperate)	Team 1: \$1,200 Team 2: \$1,200	Team 1: \$400 Team 2: \$1,600
	Cut price (compete)	Team 1: \$1,600 Team 2: \$400	Team 1: \$800 Team 2: \$800

Source: Authors' compilation.

In this exercise, two groups represent owners of competing gas stations. The goal of each group is to maximize profits for their station. In each round ("each week") participants make a simple, strategic choice: they cut their price or keep it the same. The groups make their decisions simultaneously, and the other's choice is not known before one's own decision is made. The payoff for each round is determined by the two groups' choices (depicted in table 12.1). The payoffs fit the definition of a prisoner's dilemma (Rapoport and Chammah 1965): if both groups keep their prices unchanged, both make a moderately high profit (fictitious amounts of \$1,200 each); if one cuts and the other doesn't, the station with the lower price receives the maximum payoff (\$1,600) and the status quo station receives the minimum payoff (\$400); and if both groups cut their prices, they both make a moderately low profit (\$800 each). In a one-shot interaction, the strictly dominant strategy is to cut: regardless of the other group's choice, cutters get \$400 more than they would have received from remaining constant. Since both groups are in the same position, the unique Nash Equilibrium prediction (where neither party has any motivation to change their behavior) is that they will both cut and each will receive \$800 each round. Cutting prices may be individually rational (in the short term), but it is collectively suboptimal: both groups can do better by leaving their prices alone. This choice, however, entails obvious risks. Without communication, trust, or the prospects of future interaction, both groups may find it difficult to avoid cutting.

Our research investigated the effects of these three factors—communication, trust, and expectations of future interaction—in multiple rounds (seven) of gas station games with the opportunity to communicate starting with the third round. This context gave the parties an opportunity to develop trust. The fact that the first two rounds of their interaction did not allow them to communicate meant that we could also evaluate the dynamics of trust development when some of the parties had initially cooperated and others had not. Thus, we investigated the development of trust in dyads that had begun their interactions coop-

eratively, competitively, or with a combination of cooperation and competition. Because of the repeated nature of the interactions, we were also able to track the effectiveness of cooperative and competitive strategies.

The Interactions

The participants in this study were two hundred groups of three, four, or five executives, M.B.A. students, or executive M.B.A.s. Each session included from six to twelve groups, yielding three to six interacting pairs, all in the same room. Participants did not know how long their interaction would last. At the beginning of each round, groups privately made their decision to cut or maintain prices. Once all groups had made their choices, each group's profits were displayed on the board for all to see. These payoffs were visible throughout the seven rounds of their interactions. During rounds 1 and 2 the groups were not allowed to communicate with each other. After round 2, the experimenter announced that the groups could send representative(s) to the other group; they could say anything they wished. In addition, the gas stations were "not in the United States and informal conversations between competitors were neither illegal nor unethical." Groups were still required to make their decisions privately.

After round 4, the experimenter announced that the interaction would end after three more rounds. Communication continued to be possible on each of the last five rounds. The participants were reminded of the approaching endpoint prior to rounds 5, 6, and 7.

Results and Discussion

Of the two hundred groups in this study, exactly half chose to cut their prices on the first round and half chose to maintain the same price. This suggests that the initial choice was not easily made. In fact, strong strategic arguments can be made for either choice (Murnighan 1991).

Thirty pairs (sixty groups) mutually cooperated (both groups maintained prices) on round 1. Not surprisingly, they tended to continue cooperating at a relatively high rate (twenty-eight pairs mutually cooperated on round 2; sixteen mutually cooperated on the last round) and accumulated the greatest total profits. They also tended to enjoy the interaction more than other groups did.

We were interested in finding out which combination of first-round choices would lead to the greatest difficulty in fostering trust and longer-term cooperation: pairs in which one party initially competed and the other cooperated, or those in which both parties initially competed? Forty pairs included one group cooperating (maintaining prices) and one group competing (cutting prices) on round 1; thirty pairs included both groups cutting. Competitive choices are signals that a group either expects the

other to act competitively or hopes to take advantage of the other's possible cooperation, or both. If initial competition predicts future trust (or lack of it), then pairs of mutual competitors should develop the worst relationship and perform poorest. In other words, pairs in which both groups are initially distrustful (and perhaps untrustworthy) are likely to have the greatest difficulty in building trust. Alternatively, because pairs in which one group cooperates and the other competes immediately post different profits (\$1,600 for one and \$400 for the other), asymmetric choices may instigate feelings and perceptions of inequity, especially for the initial cooperators. Moreover, because initial competitors are unlikely to share perspectives, asymmetric first-round choices may also generate communication and coordination difficulties after round 2.

The results showed that groups that made different initial choices had more trouble building a positive relationship over the next six rounds than mutually competitive groups. Their overall joint gains (across seven rounds), however, did not differ. The reason for this is that dyads in which groups made different initial choices had lower joint gains than mutual competitors when communication first became available (rounds 3 and 4), but higher joint gains as the endgame approached (rounds 6 and 7). Asymmetric outcomes in round 1 made trusting each other difficult, even when communication was possible; mutual competition made subsequent cooperation easier. This pattern reversed, however, when the endgame was announced, as more initial mutual competitors reverted to the competitive Nash Equilibrium prediction. Thus, the "absolute" level of initial distrust seemed to do more to reduce uncertainty than it did to create bad feelings.

In the few cases ($n = 7$; 17.5 percent) when initial competitors and their cooperative counterparts both changed their choices (to cooperative and competitive, respectively) on round 2, profitability increased markedly. This development was particularly rare, however, because most initial competitors thought of themselves as winners: they had just received the maximum payoff of \$1,600 and were unwilling to either apologize for their actions or provide payback. Instead, they often thought of their first-round choice as smart. Thus, when the groups could communicate with each other, they were on decidedly different wavelengths: initial cooperators claimed that they had been cheated and demanded reparations; initial competitors suggested that past outcomes were history and were no longer relevant. Round 1 competitors often indicated that, since no promises were ever made (because there was no communication in rounds 1 or 2), no one could have been cheated. Their cooperative counterparts, however, felt that their trust had already been violated. Communication, rather than being an opportunity to move toward joint gain, then, often contributed to mutual distrust that became increasingly difficult to overcome.

The data indicate that in these dyads, the first-round competitors would actually have profited more over the seven rounds if they had provided immediate payback to their cooperative counterparts. The data show that "substantive penance" could have had positive effects in either round 2 or round 3. After round 3, however, payback was both less likely and less effective (taking a "hit" now cost more than the subsequent gains from mutual cooperation). Had the interaction lasted longer than seven rounds, this kind of penance would likely have been even more effective for initial competitors. In general, the longer the shadow of the future, the more important it becomes to repair a relationship, even if it is costly in the short run.

Interestingly, the worst mistake, in terms of profits, for an initial cooperator who was paired with an initial competitor was to cooperate again on round 2. Instead, a tough stance (after having been the "sucker" in round 1) was associated with higher subsequent profits. This suggests that, in the current context, tit-for-tat (starting cooperatively and then choosing your counterpart's previous choice) was an excellent strategy after the initial round had been lost. In general, however, over all of the groups, cooperating in round 1 was no more effective than competing (average outcomes of \$7,285 and \$7,230). For interactions that are longer than seven rounds, however, cooperating in round 1 will likely provide greater payoffs than competing in round 1, making tit-for-tat, which prescribes a cooperative first move, even more effective.

The results of this study suggest a number of implications for contexts in which early decisions are made with imperfect information and later decisions depend in part on the results of prior decisions. First, a party's own choices are likely to influence their attributions about the trustworthiness and morality of others. Even when we choose non-cooperatively, we are likely to see ourselves as more trusting (and trustworthy) than others see us (for example, Tenbrunsel 1998). When we choose cooperatively, we are likely to consider the choice situation as a trustworthiness test that we have passed and that our counterparts might not. Just as pointedly, our perceptions of ourselves and of others are not likely to match the perceptions of those who do not choose as we have chosen. Second, mutual distrust may be a source of useful certainty that allows for relatively accurate attributions of each other's behaviors. These attributions seem to reflect convergent expectations that stem from symmetric initial strategies. Finally, it is beneficial for initial losers to take a strong stand in round 2; it is similarly beneficial, and more difficult, for initial winners to cooperate in round 2.

Thus, this study generated our second and third paradoxical results. First: *Noncooperative behavior is damaging to joint outcomes, and it is more damaging to joint outcomes if one party is cooperative when the other party is non-cooperative.* Second: *To repair a damaged relationship, it may be impor-*

tant for those who were initially trustworthy and those who were initially untrustworthy to switch roles.

The next section builds on issues of substantive penance, which we have touched on only briefly, by investigating the process of the breakdown and potential reestablishment of trust.

The Power of Penance

Organizations are dense with interdependencies and exposures to risk. Because trust provides a basis for cooperative action in organizations (Kahn et al. 1964) maintaining trust in relationships is crucial. The empirical literature, however, provides little in the way of systematic evidence on the dynamics of reestablishing trust once it has been violated (a recent paper by Kim et al. [forthcoming] is a notable exception). Some of our own recent research (Bottom et al. 2002) also began to address this issue. This research examined the dynamics and effectiveness of different kinds of penance on the restoration of cooperation in a repeated prisoner's dilemma.

Once a string of cooperative actions has occurred, expectations of continued cooperation increase for both parties. When one party changes course and chooses noncooperatively, the conditions surrounding that choice (expressed and perceived intentions), the benefits from reverting to cooperation, and the length of their prior cooperative interaction can all contribute to the potential for the relationship to be resurrected—or not. The philosophical literature on forgiveness (for example, North 1987) suggests that an apology, substantive amends, and forgiveness may all be required to return a relationship to a trusting, cooperative state. This experiment investigated all of these issues.

The experiment's procedures paired individuals with a programmed counterpart. Participants did not know their counterparts were programmed, but reported believing that they were interacting with another person. The program reinforced (and successfully stimulated) early, cooperative action. Later, the programmed counterpart breached the participant's trust by choosing competitively after sending messages promoting cooperation. The breach was followed by a sequence of mutual competition and then by initiatives (by the program) to reestablish cooperation.

The study manipulated three variables: intent, penance, and the duration of the parties' interaction prior to the breach of trust. The manipulation of intent occurred immediately after the breach. The programmed participant either denied intent ("I didn't mean to do that; the experimenter took the wrong card"), or acknowledged it ("I must admit that I meant to do that; I was just trying to do a little better for myself"). The duration of their interaction was manipulated by providing partici-

pants with a series of five or fifteen choices (almost always mutually cooperative) prior to the breach. After the breach, four rounds of non-cooperative play ensued to extinguish any residual cooperation. Then, the programmed counterpart began a sequence of actions that included the penance manipulation, first sending all participants an apology and a request for a return to the original cooperative behavioral norm. The apology was "I am sorry for doing this. I think we should go back to cooperation." Offers of penance were added to this message. The different versions were (1) mere talk: "I would be willing to do this if you are"; (2) small penance: "I would be willing to lose on the next round and let you win"; (3) large penance: "I would be willing to lose on the next two rounds and let you win"; or (4) an open-ended offer of penance: "What will it take for you to cooperate again?", which solicited penance requests. A follow-up condition was also run that presented the open-ended penance offer more personally: "What can I do to get you to cooperate again?"

After all of the acts of penance had occurred, participants were notified that the end of the interaction would come in five more rounds of choices. Game theory predicts that once a definite endpoint is known, a purely rational participant will choose strictly competitively (Luce and Raiffa 1957). Consequently, the last five rounds gave the participants an opportunity to choose cooperatively in conditions that entailed risk, in other words, they needed to trust their counterparts to choose cooperatively, especially in the last round. Participants' choices in the last five rounds determined the effectiveness of the different forms of penance, along with the dynamics of intent and the duration of the interaction on subsequent trusting choices.

The results showed that substantive offers of penance both large and small led to more cooperative choices in the final rounds of play than did mere talk. However, mere talk—sometimes called "cheap talk" (Farrell and Gibbons 1989)—led to many cooperative choices in the last five rounds, which runs counter to game theory's predictions. An apology alone was not as effective as an apology plus substantive penance, but it was considerably better than nothing.

Early breaches (after five rounds) typically led to cognitive rather than emotional reactions. Later breaches (after fifteen rounds) prompted many more questions about the counterpart's actions, prompting both cognitive and emotional reactions, for example, "What the hell!?"

The duration of the interactions prior to the breach and the acknowledgment or denial of intent also interacted to influence later cooperative choices in the open-ended penance conditions. Acknowledging rather than denying noncooperative intent was more effective at generating future cooperation in short interactions but the reverse was true for long interactions. It seems that a longer period of trustworthy behavior on

another's part may incline people to be more open to denials of intent, as much as to say, "We knew they couldn't have meant to do that." Perhaps it is easier to believe that you have not misjudged a person than to assume you have been wrong about them for a long time. An examination of the penance requests in the open-ended penance conditions also revealed that when faced with denials, people in shorter interactions made larger penance requests than people in longer interactions. In contrast, an admission of intent from a long-time trustworthy partner could be experienced as a particularly profound violation, and might therefore be better to deny. Conversely, during the early, uncertain period in which people feel particularly vulnerable to new partners, denials of intent apparently seemed implausible. These results suggest a fourth paradox: *As trust in a relationship increases, it is simultaneously less likely that the parties will exploit each other and more likely that they will get away with exploiting each other.*

Open-ended offers of penance that acknowledged the wrongdoer's responsibility for harm were most effective. Yet requests for penance were similar in size to the offers of small (fixed) penance, which could not fully compensate the participants for the economic opportunities that they lost due to the breach of trust. This suggests that it was not the size of penance that mattered as much as its voluntary nature. In some sense, the ultimate value of penance may have been symbolic rather than economic.

Several pictures emerge from this study. The most effective explanations for a breach of trust in reestablishing cooperation depended in part on the duration of the relationship prior to the breach. Breakdowns in longer relationships had more emotional implications, while those in shorter relationships led to a more calculated, economic approach to penance by the wronged party. Sincerity in offers of penance was critical, but the size of a penitent's offer was less significant in reestablishing trust than its voluntary nature. In addition, these data suggest that, unlike the expectations of a rational-choice approach, the reestablishment of some level of trust was possible with an apology and no substantive penance (even though substantive penance was more effective).

Reciprocity and the Perception of Trust

The rational choice approach suggests that trust development is an iteratively reciprocating process. Thus, consideration of the trusted parties' perceptions is necessary to understand and predict how trust development will evolve in a given relationship. Reciprocity models (Cialdini 1993; Gouldner 1960) suggest that trusted parties who interpret trustors' choices to accept vulnerability as completely voluntary may respond by being trustworthy. Further, higher levels of perceived trust are likely to trigger even greater trustworthiness. Thus, as individuals choose to

become more and more vulnerable, the target of their trust may feel more and more obligated to reciprocate and honor their trust—if, that is, trusted parties perceive the trustor's actions as trustors do, that is, as risky or potentially costly.

Trusting interactions can be conceptualized as non-zero-sum games, in other words, mutual trust can increase both parties' payoffs, often in both economic and in social terms. When someone trusts, the trusted parties' benefits typically increase, and if a trusted party reciprocates by being trustworthy, the trustor's benefits also increase. Thus, if we lend people money when they are in need, we make them better off. When they return what they owe, we gain, possibly monetarily but certainly socially, making the prospects for future "win-win" interactions more likely. The continued exchange of positive interpersonal acts can build trust and expand the possibilities for both parties to obtain individual and joint benefits. All this fits the rational-choice approach to trust.

An essential component of reciprocity is the perception of being trusted. If the parties who received our loan do not perceive our actions as acts of trust, they may be less likely to reciprocate. In many situations, people make positive attributions of their own actions (Taylor and Brown 1988) and devalue the contributions or concessions of others (Stillinger et al. 1990). Problems surface when the parties' perceptions of their interaction are incompatible (Bettenhausen and Murnighan 1985). Because shared understandings are necessary for effective cooperation and trust development, the variation in the perceptions of the trusting and trusted party can be critical.

We investigated these kinds of processes in two experiments (Pillutla, Malhotra, and Murnighan 2003). In particular, we manipulated the amount of trust that trustors exhibited and measured how much the trusted parties reciprocated. This allowed us to map the pattern of trust and reciprocity in a situation that involved an easily measured economic outcome (money).

Trust and Reciprocity

Joyce Berg, John Dickhaut, and Kevin McCabe (1995) developed a version of the Trust Game that differs from the version developed by Chris Snijders (1996), Partha Dasgupta (1988), and David M. Kreps (1990), described earlier. This version is particularly appropriate for investigating reciprocity. In Berg, Dickhaut, and McCabe's (1995) initial experiment, participants received ten dollars for participating. They were then directed to one of two rooms. Participants in room A (player 1s) were told that they could offer as much of their ten dollars as they wished to one of the people in room B (where the player 2s waited). They knew that the recipients would receive three times the amount that

they sent. Thus, if a player 1 sent five dollars, the player 2 would receive fifteen dollars (in addition to the ten dollars he or she would receive at the outset). Player 2s would then freely decide how much to return to player 1s. Everyone acted anonymously.

Since this one-shot situation provided no opportunity for reputation building, player 2s had no incentive to return any money. Game theoretic analysis, then, suggests that player 1s should send no money. The unique Nash Equilibrium prediction for this game is that player 1s will not send any money and, if player 2s receive any money, they will return none of it to player 1s. This outcome is far from collectively efficient, since player 1s who send their entire ten dollars can create an additional twenty dollars of joint value.

By sending money to player 2s, player 1s signal that they trust player 2s to return enough to make their trusting action individually as well as jointly profitable. This choice, however, is a unilateral gamble: player 1s cannot ensure that player 2s will act benevolently. Thus, the structure of this game perfectly fits, in a simple, easily observable way, our working definition of trusting action.

Berg, Dickhaut, and McCabe's (1995) results suggest that player 1s often trusted and player 2s often reciprocated player 1s' trust, but reciprocity was not universal, either in frequency or size. In particular, player 2s rarely returned enough to make the two parties' outcomes equal. Berg, Dickhaut, and McCabe (1995, 137) concluded that "some of the subjects who did not reciprocate may not have interpreted room A behavior as initiating a trust." This suggests that trusting actions may only be reciprocated if player 2s view them as trusting.

An important strategic question for player 1s is how much they should send for their act to be seen as trusting. Berg, Dickhaut, and McCabe (1995) provide only a partial answer to this question because their data are restricted to the actual amounts that player 1s sent. We explored this issue further by explicitly manipulating the amounts that player 1s sent. Thus, we presented player 2s with a variety of different amounts. We also varied player 1s' initial endowments. Manipulating these factors allowed us to provide a clear picture of a broad range of the consequences of different degrees of trusting acts.

A simple reciprocity model suggests that the higher the amount sent, the higher the return should be (Cialdini 1993; Gouldner 1960). That is, as trust increases, so should reciprocity. From player 2's perspective, however, player 1s have many choices. For instance, sending small amounts may prompt player 2s to ask, "Why didn't they send more?" Player 2s might answer this question by deciding that player 1s are not particularly trusting. Small offers might also be viewed as insulting, leading to negative attributions and consequently little reciprocation (compare Pillutla and Murnighan 1996).

In contrast, player 2s might view large amounts as favors that create an obligation. Receiving a large amount from player 1 may make it difficult to justify keeping the lion's share of the gains. When more has been received, it becomes more difficult for player 2 to deny that player 1's act is an act of trust. Many self-serving and other interpolations are possible following the transferral of small amounts. Few are possible for very large amounts. Not reciprocating would indicate that they themselves are not fair, an identity attribution that people generally seek to avoid (Greenberg 1990; Murnighan, Oesch, and Pillutla 2001). The desire to think of themselves as fair, or trustworthy, then, might compel player 2s to reciprocate sizable offers more than they reciprocate less than sizable offers.

The Experiments

Participants were all player 2s; we experimentally manipulated player 1s' endowments and offers. Player 2s each received offers from one player 1 (experiment 2) or from 16 different player 1s (experiment 1). Player 2s knew that player 1s started with either ten or twenty dollars; their offers were two, three, four, five, six, nine, ten, twelve, eighteen, or twenty. Player 2s were given no initial endowment. As in Berg et al. (1995), player 2s received three times the amount that player 1s sent. Participants in experiment 1 knew that one of their decisions would be randomly chosen to determine their actual monetary payoff. After being informed of the amount that a player 1 had sent, player 2s indicated how much they would return. All decisions were anonymous.

The data across the two experiments were consistent and reliable. The more player 2s received, the more they returned. Also, except for one condition, player 2s returned less for comparable offers (for the same amount of money offered) when player 1s had a twenty- rather than a ten-dollar endowment. Thus, the proportion that player 1s sent was clearly important. In almost all of the conditions, if it was possible, player 2s' modal returns equalized their outcomes and player 1s' outcomes. For instance, when player 1s sent nine of ten dollars, the most frequent player 2 response was a return of thirteen dollars, leaving both players with fourteen. This equalizing pattern held even for the most lucrative condition: when player 1s sent twenty, more player 2s returned thirty than any other response.

These data suggest that reciprocity increases as trust increases. On average, however, player 1s were better off only when they had sent all of their ten-dollar endowment or almost all, eighteen or twenty dollars, of their twenty-dollar endowment. In trust terms, player 1s only benefited from trusting player 2s when they fully (or almost fully, in the twenty-dollar case) trusted. In contrast, partial trusting acts led, on average, to small monetary losses for player 1s.

This result may be explained, at least partially, by player 2s' equality motives (see Messick 1993). For example, when player 1s send five dollars of a twenty-dollar endowment and player 2s desire equality, they should return nothing. This means that they neither return as much as was sent nor do they reward player 1s for freely sending them money. However, even when player 2s received enough to allow them to equalize their outcomes and reward player 1s—say, when player 1s sent twelve dollars of twenty, player 2s could return fourteen to give each of them twenty-two dollars—their average returns gave player 1s final outcomes that were less than their initial endowment. In this case, player 2s' average return was ten dollars and forty-five cents, leaving player 1s with eighteen dollars and forty-five cents and player 2s with twenty-five dollars and fifty-five cents.

The data clearly indicate that player 1s do better by not engaging in trusting acts rather than by trusting partially. However, when they maximize their trusting acts, the data suggest that, on average, they come out slightly ahead (by forty-eight cents when they send their ten-dollar endowment and by thirty-three cents when they send their twenty-dollar endowment). At the same time, by offering the entire amount they risk receiving nothing: in both endowment conditions, sending everything led to over 20 percent of player 2s keeping the entire amount, leaving full trustors with nothing.

Postexperiment questionnaires indicated that player 2s felt obligated by receiving money; they also perceived, in general, that sending money was a smart, trusting act. Both sets of feelings—of obligation and the combination of perceptions of intelligence and a trusting act—mediated the amount sent–amount returned relationship. Thus, player 2s' attributions may have been as important as player 1s' acts in determining player 1s' ultimate monetary outcomes.

This study suggests that many player 2s constructed negative attributions of partial trustors and that, unlike sending nothing, which entails no risk, sending everything may be jointly profitable but was tremendously risky for trustors. Trusting fully is effective, on average, because it seems to compel most trusted parties to feel obligated and reciprocate. The fifth paradox, then, is clear: *Trust, by definition, entails risk and vulnerability, which is important to manage, but trustors are punished for hedging, and may need to choose between large acts of trust, or none at all.*

Risking as Part of Loving

At the beginning of this paper, we described common conceptions of trust as fitting many of the parameters of a rational-choice approach. We also presented an S-curve to depict the normal trust development

process (figure 12.1). As our research has shown, in several contexts, the trust development process is neither neat nor strictly rational. Instead, degrees of risk and perceptions of interdependence can lead to a variety of trust development patterns. These patterns also appear in a recent field study that presents a striking perspective on the paradoxes of trust.

Paul R. Appleby, Lynn C. Miller, and Sadina Rothspan (1999) surveyed forty-six long-term homosexual male couples on their reasons for engaging in safe or unsafe sexual behavior. They note that stable relationships do not, at least in the context studied, make unprotected sex safe among gay men. For example, 73 percent of gay men in stable relationships reported having sex with someone besides their partner at least once during their relationship (Peplau and Cochran 1988). With estimates that HIV infection among gay men in urban populations range anywhere from 21 percent to 51 percent (Valdiserri et al. 1988), promiscuity becomes a serious source of concern.

In Appleby, Miller, and Rothspan (1999), 21 percent of the couples had been completely monogamous. Thus, a majority of the survey respondents admitted to having sex outside of their committed relationship. Only 30 percent of those who reported having sex outside the relationship, however, claimed to have consistently used condoms. The risks that unprotected sex had for their committed relationships may also have been apparent to their partners, as 84 percent reported knowing about their partners' outside sexual activities.³ Surprisingly, and tragically, those couples in which one or both partners engaged in extra-relationship sexual contact were no more safe in their sexual behavior within their primary relationships (by using condoms) than were couples who did not engage in any extra-relationship sexual contact.

Clearly, many of the men in this study chose to be vulnerable to a potentially life-threatening degree. Appleby, Miller, and Rothspan's survey probed the rationales for the choices they made in their primary relationships. Individuals who engaged in risky (unprotected) sex were more likely to use love, trust, and commitment as explanations for their behavior than those who engaged in safer sex. In addition, individuals who were "more dependent upon their relationships and who desired a stable and lasting relationship practiced riskier sex" (1999, 81). This is consistent with an earlier finding that people who feel that their partner is someone "special" engage in more risky sexual behavior (Kelly et al. 1991). Respondents consistently indicated that making requests for safer sex conveyed a negative message to their partners.

Beyond being profoundly troubling, these data suggest that people often take risks to build trust in relationships rather than building trust so that they can take risks. Our definition of trusting incorporates a willingness to assume risk as an essential element, and the rational-choice

approach to trust suggests that trust will grow gradually, will become established after many positive interactions, and that greater risk taking will follow. The Appleby, Miller, and Rothspan (1999) data suggest the exact converse, that people take (large) risks as a way to build or maintain trust. In other words, to establish mutual trust, individuals must act as if they are trusting. If they do not, then trust may not develop at all. The sixth paradox that this suggests cuts to the heart of traditional conceptions of trust development: *While building trust is necessary to accommodate large risk taking, large risk taking may be necessary to build trust.*

This conclusion strengthens the conclusion drawn from our reciprocity study (Pillutla, Malhotra, and Murnighan 2003): to build trust, individuals may have to take more risks than are actually warranted, at least in terms of rational choice. The Appleby, Miller, and Rothspan (1999) findings point out that, for relationships that are highly valued or are perceived as being of great future importance, the risks that are necessary to establish a person as both trusting and trustworthy may be significant. There is another related, and striking, similarity between the Appleby, Miller, and Rothspan data and our reciprocity study. In the reciprocity study we found that the greatest returns accrued to those who demonstrated absolute trust in their counterparts; partial trusting actually had negative effects. There are not many more powerful examples of absolute trusting acts than sexually active men in a contemporary urban gay community choosing to practice unsafe sex. Furthermore, the participants in the Appleby, Miller, and Rothspan (1999) study seem to intuit what our data demonstrated: acts that might be viewed as "partial" trust (for example, sex with a condom) can actually undermine trust.

Whether these results apply to other kinds of relationships, for example, between two organizations, becomes an intriguing research question. Small, dependent firms that envision considerable long-term profitability, for instance, may be willing to engage in considerable risk taking (for example, operating without a contract) to establish themselves with new partners. Further, a suggestion that would reduce the "trust-related" risk, say through the formalization of informal agreements in contractual form, may be experienced as negative or suspicious. This may be particularly true when one of the two parties values a developing relationship more than the other. In such situations, we might expect to see the pattern we frequently observed in our prisoner's dilemma research (Malhotra and Murnighan 2003): the more dependent party choosing cooperatively and the less dependent party choosing competitively. Whether the inequities that follow lead to the same difficulties that we observed in our research is yet another intriguing empirical question.

Discussion and Conclusions

This chapter began with the notion that common conceptions of the trust development process might not be so simple or so rational as traditional rational-choice theories might lead us to expect. After defining trust and trusting actions and considering these common conceptions, we reviewed a series of research projects that clearly call into question the assumptions underlying what we, and others, have called a rational-choice approach to trust.

The first project, on the impact of contracts, concentrated on the actions and thoughts of trustors. It showed that contractually mandated interactions interfered with the trust establishment process. Providing an opportunity to use binding contracts reduced the likelihood that two parties would be able to easily develop trust. Attribution theory suggests that the parties gave credit to the contracts rather than to their counterparts for their early cooperation and these attributions made it more difficult for them to cooperate with and trust each other later.

The second project focused on the strategic interplay of pairs of parties who could create mutual benefits for each other if they interacted effectively. Communication opportunities and repeated interactions allowed them to develop trust and a mutually beneficial cooperative relationship. Pairs of groups that were lucky enough to experience mutually cooperative choices initially were particularly effective over the course of their interactions. Pairs of groups that experienced mutually competitive choices initially were less effective but were able to work from a common basis of understanding. Their communications were less troubled and they were often able to build a cooperative relationship, if only for a short time. As the end of their interaction approached, their ability to cooperate diminished markedly. Thus, their early competitive choices were harbingers of their later choices, and both parties seemed to expect this. When one of two groups cooperated initially and its counterpart competed, they had a difficult time establishing trust. Only in the rare instances when both groups reversed their initial choices early in their interaction, thereby cleaning the slate and reestablishing equivalent outcomes, were they able to reap the benefits of cooperation. In essence, the strategic interplay of the parties in this potentially competitive context made the establishment of trust a difficult task and one that often required counterintuitive strategies by those involved.

The third project pursued the topic of strategic interactions and trust even further by investigating the potential antecedents of trust reestablishment following a breach. Unlike many conceptions of the breakdown of trust, this experiment showed that some level of trust could be reestablished without the exchange of anything more than a sincere

apology. In addition, offers of substantive penance were more effective than apologies alone, but the size of these offers was not particularly important. Thus, penance appeared to be more important symbolically than substantively. Reestablishing trust, however, does appear to require active intervention and is far from automatic.

Project number four shifted the focus from the trustor and the two parties' strategies and interactions to the perceptions of the trusted party. It investigated the likelihood of reciprocity when an initial trustor made trusting choices that varied widely in degree. The results suggest that actions that communicate partial rather than complete trust lead to negative attributions and less reciprocity by the trusted parties. Many trusted parties acted to equalize their financial outcomes and those of their trustors but on average, trustors lost resources if they did anything other than trust completely or not at all. Linear models of reciprocity do not fit these data. Instead, attributional processes appear to be causal forces in the choice to potentially reciprocate trusting actions.

The final project reviewed here concerned the sexual choices of gay men in committed homosexual relationships. Most of the respondents in the Appleby, Miller, and Rothspan (1999) survey reported engaging in sexual encounters outside their primary relationship, and many of those were unsafe. Yet they also reported engaging in unsafe sex within their relationships, even though most were aware of their own or their partners' past extrarelationship activities. Response patterns suggested that participants may have been motivated to put themselves at risk to signal love, trust (or trustworthiness), and commitment, and to avoid signaling distrust (or a lack of trustworthiness). These data also do not fit the rational-choice approach to trust.

Reconsidering Definitions

Early in this paper we adopted Rousseau et al.'s (1998) definition of trust: "a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behavior of another." A rational approach to the establishment of positive expectations would require small, positive exchanges first and larger exchanges later, rather than the reverse. In many ways, this matches Charles E. Osgood's (1962) GRIT approach, which he proposed as a means for establishing peaceful relations in competitive contexts. GRIT—graduated reciprocation in tension reduction—is a process that begins with a unilateral, cooperative act that entails minimal risk (much like the first cooperative choice in tit-for-tat). When the other party reciprocates, larger, riskier cooperative acts follow until both sides trust each other enough to make serious concessions.

Needless to say, the risks of unsafe sex in the gay community, or anywhere else for that matter, are not minimal. One event can ultimately

mean an individual's untimely death. Similarly, participants in the reciprocity study demonstrated the inefficacy of small and moderate acts of trust. Moreover, all five studies demonstrate the need to better understand the attributional processes that underlie trust (and reciprocity) decisions, the effects of perceived dependence on trust, and the differing perspectives of trustors and trusted parties. These issues suggest that the rational-choice approach to trust is not universally applicable. Instead, the data we have summarized here suggest the need for an alternative model. In the next section, we outline the beginnings of such a model. We elaborate on it more completely in another paper (Weber, Malhotra, and Murnighan, under review).

The Attributional Basis of Trust

The data that we have summarized in this paper consistently and repeatedly focus our attention on the attributional analyses of interacting parties who must decide whether they can trust each other. All of their analyses, in turn, are intimately connected to considerations of their counterparts' intentions—their trustworthiness. We think that an attributional model is needed to complement the rational-choice approach to trust. In essence, it would suggest that, to develop trust, individuals attempt to (and feel that they need to) influence others' attributions of their actions. Not only do initial trustors want their trust honored, but they also want to become the recipients of trust—they want to be trusted. To do so, they must not only judge their counterparts to be trustworthy, but they must also influence them to think that they themselves are trustworthy. To create such attributions, they implicitly invoke reciprocity and act (much sooner than the rational model would predict) as if they truly trust these relevant others. Thus, if Sandy wants to build trust with Chris, Sandy may act as if she has considerable trust in Chris. Sandy's hope is that Chris will think that Sandy is very trusting, as Sandy has obviously taken an action that is particularly risky. Sandy hopes that this risky act will communicate to Chris that Sandy can be trusted and that Chris will then take a similar (or greater) risk. In the process, Sandy may overestimate Chris's desire to augment their relationship and build trust. Thus, Sandy may expect that Chris will look at Sandy's action as a golden opportunity to create a stronger bond between them and will reciprocate in kind. But if Chris has more modest intentions, the reciprocation that Sandy hopes for may not be forthcoming.

This process differs from the rational-choice approach in many ways. First, it explicitly differentiates the perceptions of the two parties, with greater risks being taken by trustors earlier in their interaction. In addition, the trusted party's responses can vary widely. In Pillutla, Malhotra, and Murnighan (2003), for instance, the modal response was to equalize out-

comes. This strategy honored and rewarded trust. But many recipients of the other party's trust did not honor it at all and kept all of the money (this was the second most frequent response). Others simply returned an amount that was equal to the initial trustors' offer (providing them with no benefit) and some reacted in ways that seemed to mix these three responses. A motivated attributions model of trust would suggest that initial trustors who take considerable risks may suffer considerably from their choices, sometimes very quickly. The paradox here is that because partial trustors are judged harshly, greater risks are necessary to increase the chance that the trusted party will reciprocate and that trust will develop. But the greater risks are serious gambles that can result in considerable net losses for trusting initiators.

The rational-choice approach to trust suggests that the parties should only risk more when they have a foundation of previous, positive experiences. Risky choices should only increase in magnitude following sufficient reciprocity. The rational choice approach also implies, however, that trustors will consciously evaluate their risks before they take them. Even after repeated positive interactions, then, rational choice suggests that the parties will suspect counterparts who could be less than honorable and will therefore temper their risk taking.

An attributional model would suggest that dependent actors who see considerable benefit from a stronger relationship may take these kinds of disproportionate risks, even without an elaborated, positive history. Their vision of the future may dominate, interfere with, and overcome any of their rational calculations. And if their initial trusting act is honored, attributional euphoria may make subsequent rational analysis even less likely. The paradox here is that trust may be based more on the expectations of one's own benefits than on rational calculations or clear consideration of the other party's needs and likely responses. Historical evidence and its potential predictability can be overwhelmed by expectations of potential benefits.

It is interesting to note that if dependent initial trustors do not take a big risk and the other party realizes that they have held back, significant reciprocity is not likely, and a weak foundation of moderately positive acts may limit the likelihood of encouraging cooperative actions by either party. One contributor to these kinds of relationship problems is that each party is likely to overestimate its own contribution and underestimate the other's contribution to the trust development process (Taylor and Brown 1988). In the extreme, this will result in mutually perverse attributions regarding the other's trustworthiness (Kramer 1998). Then neither party is likely to step in and make the contributions that are necessary to create the basis for a positive, cooperative spiral. Instead, if both parties are likely to feel that they are already giving more to the relationship (by being more trusting, accepting more risks, and so forth)

and the inequity that they perceive is likely to increase, positive attributions of the other party become increasingly less likely.

A further complicating factor is that an individual's own actions can influence his or her conclusions about another's intentions. This was certainly the case among the participants in the gas station game (Malhotra and Murnighan 2003), where groups that made an initially competitive choice often expressed an unwillingness to trust their initially cooperative counterparts. Similarly, Ann Tenbrunsel (1998) found that when stakes are high, misrepresentation became more likely in negotiations and the misrepresenters often assumed that their counterparts had been less than truthful as well, even when they had no real evidence for drawing this conclusion.

An attributional model of trust would build on the foundation of paradoxes and counterintuitive empirical results identified here. At its core, it would suggest that people often take large risks to develop trust and may not have the option of developing trust gradually because recipients of trusting acts judge partial trustors harshly. This makes it difficult for interacting parties to use (or suggest) contracts or to accept small doses of initial vulnerability. While this may not dissuade particularly dependent parties from engaging in (risky) collaborations, other partnerships may be less likely. The ultimate paradox here is that the very thing that trust is supposed to facilitate, namely, cooperation, may be threatened by the processes that underlie trust development. With that said, a silver lining in this review is that the reestablishment of trust, while by no means easy, might not be as improbable as originally suspected. More research on rebuilding trust and on the effects of communication and expectations of future interaction is clearly necessary, and perhaps critical, with regard to this issue. In particular, this kind of research holds the promise of discovering encouraging news about the initial trust-building process. Our final observation is that our attributional conceptualization of the trust development process suggests that people can, at times, effect more rapid development of trust by straying from the risk-averse behavioral prescriptions of a rational incremental model. Given people's apparent propensity for acting "irrationally" in the social domain, this is good news, indeed.

This project benefited from the assistance and support of several organizations and many people. First, we gratefully acknowledge the financial assistance of the Russell Sage Foundation and the Dispute Resolution Research Center at the Kellogg Graduate School of Management at Northwestern University. We also thank Rod Kramer, Karen Cook, and all of the participants in the conference at Stanford for their constructive comments on earlier versions of this manuscript.

Notes

1. Debra Meyerson, Karl E. Weick, and Roderick M. Kramer's (1996) article on swift trust may be an exception that proves the rule.
2. This discussion is admittedly general and does not address the possibility that people might trust each other in one domain but not in another. Expanding our discussion to deal with these conceptual issues goes far beyond the scope of this paper. We can visualize a model that is domain-specific and can accommodate a variety of different domains that, in the most trusting of interactions, might expand the "total" trust to phenomenal amounts. (With appropriate scaling, these conceptualizations can be kept to a manageable size.)
3. However, it is not clear from the data when partners became aware of the extrarelationship sexual contacts.

References

- Appleby, Paul R., Lynn C. Miller, and Sadina Rothspan. 1999. "The Paradox of Trust for Male Couples: When Risking Is a Part of Loving." *Personal Relationships* 6(1): 81-93.
- Arrow, Kenneth. 1974. *The Limits of Organizations*. New York: Norton.
- Berg, Joyce, John Dickhaut, and Kevin McCabe. 1995. "Trust, Reciprocity, and Social History." *Games and Economic Behavior* 10(1): 122-42.
- Bettenhausen, Kenneth, and J. Keith Murnighan. 1985. "The Emergence of Norms in Competitive Decision-Making Groups." *Administrative Science Quarterly* 30(3): 350-72.
- Blount, Sally, Melissa C. Thomas-Hunt, and Margaret A. Neale. 1996. "The Price Is Right—Or Is It? A Reference Point Model of Two-Party Price Negotiations." *Organizational Behavior and Human Decision Processes* 68(1): 1-12.
- Boon, Susan D., and John G. Holmes. 1991. "The Dynamics of Interpersonal Trust: Resolving Uncertainty in the Face of Risk." In *Cooperation and Prosocial Behavior*, edited by Robert A. Hinde, and Jo Groebel. New York: Cambridge University Press.
- Bottom, William, Steven Daniels, Kevin S. Gibson, and J. Keith Murnighan. 2002. "When Talk Is Not Cheap: Substantive Penance and Expressions of Intent in Rebuilding Cooperation." *Organization Science* 13(5): 497-513.
- Boyle, Richard, and Phillip Bonacich. 1970. "The Development of Trust and Mistrust in Mixed-Motive Games." *Sociometry* 33(2): 123-39.
- Carnevale, Peter J., and Alice Isen. 1986. "The Influence of Positive Affect and Visual Access on the Discovery of Integrative Solutions in Bilateral Negotiations." *Organizational Behavior and Human Decision Processes* 37(1): 1-13.
- Cialdini, Robert B. 1993. *Influence: Science and Practice*. New York: HarperCollins.
- Coffrin, Clyde, and Gavin Cochran. 1982. "Management Contracts: Against/For." *Credit Union Magazine* 48(5): 52-54.
- Coleman, James. 1990. *Foundations of Social Theory*. Cambridge, Mass.: Harvard University Press.
- Das, T. K., and Bing-Sheng Teng. 1998. "Between Trust and Control: Developing Confidence in Partner Cooperation in Alliances." *Academy of Management Review* 23(3): 491-512.
- Dasgupta, Partha. 1988. "Trust as Commodity." In *Trust: Making and Breaking Cooperative Relations*, edited by Diego Gambetta. Oxford: Blackwell.
- Deutsch, Morton. 1958. "Trust and Suspicion." *Journal of Conflict Resolution* 4(2): 265-79.
- Enzle, Michael E., and Sharon C. Anderson. 1993. "Surveillant Intentions and Intrinsic Motivation." *Journal of Personality and Social Psychology* 64(2): 257-66.
- Farrell, Joseph, and Robert Gibbons. 1989. "Cheap Talk Can Matter in Bargaining." *Journal of Economic Theory* 48(1): 221-37.
- Fukuyama, Francis. 1995. *Trust: The Social Virtues and the Creation of Prosperity*. New York: Free Press.
- Gambetta, Diego. 1988. "Can We Trust Trust?" In *Trust: Making and Breaking Cooperative Relationships*, edited by Diego Gambetta. Cambridge, Mass.: Blackwell.
- Gouldner, Alvin W. 1960. "The Norm of Reciprocity: A Preliminary Statement." *American Sociological Review* 25(1): 161-78.
- Greenberg, Jerald. 1990. "Looking Fair Versus Being Fair: Managing Impressions of Organizational Justice." In *Research in Organizational Behavior*, edited by L. L. Cummings and B. Staw. Volume 12. Greenwich, Conn.: JAI Press.
- Homans, George C. 1961. *Social Behavior: Its Elementary Forms*. New York: Harcourt Brace and World.
- Howard, Neil. 1966. "The Theory of Meta-Games." *General Systems* 11(1): 167-86.
- Hughes, Patrick, and George Brecht. 1975. *Vicious Circles and Infinity: A Panoply of Paradoxes*. New York: Penguin.
- Jensen, Michael C., and William H. Meckling. 1976. "Theory of the Firm: Managerial Behavior, Agency Costs, and Ownership Structure." *Journal of Financial Economics* 3(2): 305-60.
- Johnson-George, Cynthia, and Walter Swap. 1982. "Measurement of Specific Interpersonal Trust: Construction and Validation of a Scale to Assess Trust in a Specific Other." *Journal of Personality and Social Psychology* 43(6): 1306-17.
- Kahn, Robert L., Donald M. Wolfe, Robert P. Quinn, J. D. Snoek, and Robert A. Rosenthal. 1964. *Organizational Stress: Studies in Role Conflict and Ambiguity*. New York: McGraw-Hill.
- Kelley, Harold H. 1967. "Attribution Theory in Social Psychology." In *Nebraska Symposium on Motivation*, edited by David Levine. Volume 15. Lincoln: University of Nebraska Press.
- . 1979. *Personal Relationships: Their Structures and Process*. Hillsdale, N.J.: Erlbaum.
- Kelly, Jeffrey A., Seth C. Kalichman, M. R. Kauth, H. G. Kilgore, H. V. Hood, P. E. Compos, S. M. Rao, T. L. Brasfield, and J. S. St. Laurence. 1991. "Situational Factors Associated with AIDS Risk Behavior Lapses and Coping Strategies Used by Gay Men Who Successfully Avoid Lapses." *American Journal of Public Health* 81(10): 1335-38.
- Kim, Peter H., Donald L. Ferrin, Cary C. Cooper, and Kurt T. Dirks. Forthcoming. "Removing the Shadow of Suspicion: The Effects of Apology Versus Denial for Repairing Competency-Based Versus Integrity-Based Trust Violations." *Journal of Applied Psychology*.

- Klein, Benjamin. 1993. "Contracts and Incentives: The Role of Contract Terms in Assuring Performance." In *Contract Economics*, edited by Lars Werin, and Hans Wijkander. Cambridge, Mass.: Blackwell.
- Kollock, Peter. 1994. "The Emergence of Exchange Structures: An Experimental Study of Uncertainty, Commitment and Trust." *American Journal of Sociology* 100(2): 313-45.
- Kramer, Roderick M. 1998. "Paranoid Cognition in Social Systems: Thinking and Acting in the Shadow of Doubt." *Personality and Social Psychology Review* 2(2): 251-75.
- . 1999. "Trust and Distrust in Organizations: Emerging Perspectives, Enduring Questions." *Annual Review of Psychology* 50(1): 569-98.
- Kramer, Roderick M., and Tom R. Tyler, eds. 1996. *Trust in Organizations*. Thousand Oaks, Calif.: Sage.
- Kreps, David M. 1990. "Corporate Culture and Economic Theory." In *Perspectives on Positive Political Economy*, edited by James E. Alt and Kenneth A. Shepsle. Cambridge: Cambridge University Press.
- Lewicki, Roy J., Daniel J. McAllister, and Robert J. Bies. 1998. "Trust and Distrust: New Relationships and Realities." *Academy of Management Review* 23(3): 438-512.
- Lindskold, Sven. 1978. "Trust Development, the GRIT Proposal, and the Effects of Conciliatory Acts on Conflict and Cooperation." *Psychological Bulletin* 85(4): 772-93.
- Lorenz, Edward. 1999. Trust, Contract and Economic Cooperation. *Cambridge Journal of Economics* 23(3): 301-15.
- Luce, R. Duncan, and Howard Raiffa. 1957. *Games and Decisions: Introduction and Critical Survey*. New York: Wiley.
- Luhmann, Niklas. 1988. "Familiarity, Confidence, Trust: Problems and Alternatives." In *Trust: Making and Breaking Cooperative Relations*, edited by Diego Gambetta. Oxford: Blackwell.
- Malhotra, Deepak, and J. Keith Murnighan. 2002. "The Effects of Formal and Informal Contracts on Interpersonal Trust." *Administrative Science Quarterly* 47(3): 534-59.
- . 2003. "Strategies, Trust, and Consequences in Repeated, Finite Prisoner's Dilemma Games." Unpublished paper. Draft under revision, Kellogg School of Management, Northwestern University.
- Mayer, Roger C., James H. Davis, and F. David Schoorman. 1995. "An Integrative Model of Organizational Trust." *Academy of Management Review* 20(3): 709-34.
- McAllister, Daniel. J. 1995. "Affect- and Cognition-Based Trust as Foundations for Interpersonal Cooperation in Organizations." *Academy of Management Journal* 38(1): 24-59.
- Messick, David M. 1993. "Equality as a Decision Heuristic." In *Psychological Perspectives on Justice: Theory and Applications*, edited by Barbara A. Mellers and Jonathan Baron. Cambridge: Cambridge University Press.
- Meyerson, Debra, Karl E. Weick, and Roderick M. Kramer. 1996. "Swift Trust in Temporary Groups." In *Trust in Organizations*, edited by Roderick M. Kramer and Tom R. Tyler. Thousand Oaks, Calif.: Sage.
- Misztal, Barbara A. 1996. *Trust in Modern Societies*. Cambridge, Mass.: Blackwell.
- Morishima, M. 1996. "Renegotiating Psychological Contracts: Japanese Style." *Journal of Organization Behavior* 3(1): 139-58.
- Murnighan, J. Keith. 1991. *The Dynamics of Bargaining Games*. Englewood Cliffs, N.J.: Prentice Hall.
- Murnighan, J. Keith, John Oesch, and Madan M. Pillutla. 2001. "Player Types and Self-Impression Management in Dictatorship Games: Two Experiments." *Games and Economic Behavior* 37(3): 388-414.
- North, Joanna. 1987. "Wrongdoing and Forgiveness." *Philosophy* 62(3): 449-508.
- Nye, David. 1988. "Trust Is a Well-Drawn Employment Contract." *Across the Board* 25(10): 32-41.
- Osgood, Charles E. 1962. *An Alternative to War or Surrender*. Urbana: University of Illinois Press.
- Parkhe, Arvind. 1998. "Understanding Trust in International Alliances." *Journal of World Business* 33(3): 219-40.
- Peplau, Letitia A., and Susan D. Cochran. 1988. "Value Orientations in the Intimate Relationships of Gay Men." In *Gay Relationships*, edited by John De Cecco. New York: Harrington Park Press.
- Pillutla, Madan M., Deepak Malhotra, and J. Keith Murnighan. 2003. "Attributions of Trust and the Calculus of Reciprocity." *Journal of Experimental Social Psychology* 39(3): 448-55.
- Pillutla, Madan M., and J. Keith Murnighan. 1996. "Unfairness, Anger, and Spite: Emotional Rejections of Ultimatum Offers." *Organizational Behavior and Human Decision Processes* 68(2): 208-24.
- Putnam, Robert D. 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, N.J.: Princeton University Press.
- Rapoport, Anatol, and Albert N. Chammah. 1965. *Prisoner's Dilemma: A Study in Conflict and Cooperation*. Ann Arbor: University of Michigan Press.
- Rapoport, Amnon, and Abbe Mowshowitz. 1966. "Experimental Studies of Stochastic Models for the Prisoner's Dilemma." *Behavioral Science* 6: 444-58.
- Rempel, John K., John G. Holmes, and Mark P. Zanna. 1985. "Trust in Close Relationships." *Journal of Personality and Social Psychology* 49(1): 95-112.
- Rousseau, Denise M., Sim B. Sitkin, Ronald S. Burt, and Colin Camerer. 1998. "Not So Different After All: A Cross-Discipline View of Trust." *Academy of Management Review* 23(3): 393-404.
- Seligman, Adam B. 1997. *The Problem of Trust*. Princeton, N.J.: Princeton University Press.
- Shapiro, Debra, Blair H. Sheppard, and Lisa Cheraskin. 1992. "Business on a Handshake." *Negotiation Journal* 8(4): 365-77.
- Siamwalla, Ammar. 1978. "Farmers and Middlemen: Aspects of Agricultural Marketing in Thailand." *Economic Bulletin for Asia and the Pacific* (June): 38-50.
- Sitkin, Sim B. 1995. "On the Positive Effect of Legalization on Trust." In *Research on Negotiation in Organizations*, edited by Robert J. Bies, Roy J. Lewicki, and Blair H. Sheppard. Volume 5. Greenwich, Conn.: JAI Publishing.
- Sitkin, Sim B., and Nancy L. Roth. 1993. "Explaining the Limited Effectiveness of Legalistic 'Remedies' for Trust/Distrust." *Organization Science* 4(3): 367-92.
- Smith, Ken, Stephen Carroll, and Susan Ashford. 1995. "Intra- and Inter-organizational Cooperation: Toward a Research Agenda." *Academy of Management Journal* 38(1): 7-23.

- Snijders, Chris. 1996. *Trust and Commitments*. Groningen: Interuniversity Center for Social Science Theory and Methodology.
- Solomon, Leonard. 1960. "The Influence of Some Types of Power Relationships and Game Strategies Upon the Development of Interpersonal Trust." *Journal of Abnormal Social Psychology* 61(2): 23–30.
- Stillinger, Constance A., Michael Epelbaum, Dacher Keltner, and Lee Ross. 1990. "The 'Reactive Devaluation' Barrier to Conflict Resolution." Unpublished paper (mimeographed). Palo Alto, Calif.: Stanford University.
- Swinth, Robert L. 1967. "The Establishment of the Trust Relationship." *Journal of Conflict Resolution* 11(3): 335–44.
- Taylor, Shelley E., and Jonathan D. Brown. 1988. "Illusion and Well-Being: A Social-Psychological Perspective on Mental Health." *Psychological Bulletin* 103(2): 193–210.
- Taylor, Shelley E., and Susan T. Fiske. 1975. "Point-of-View and Perceptions of Causality." *Journal of Personality and Social Psychology* 32(3): 439–45.
- . 1978. "Salience, Attention, and Attribution: Top of the Head Phenomena." In *Advances in Experimental Social Psychology*, edited by Leonard Berkowitz. Volume 11. New York: Academic Press.
- Tenbrunsel, Ann E. 1998. "Misrepresentation and Expectations of Misrepresentation in an Ethical Dilemma: The Role of Incentives and Temptation." *Academy of Management Journal* 41(3): 330–39.
- Thompson, Leigh. 1991. "Information Exchange in Negotiation." *Journal of Experimental Social Psychology* 27(2): 161–79.
- Uzzi, Brian. 1997. "Social Structure and Competition in Interfirm Networks: The Paradox of Embeddedness." *Administrative Science Quarterly* 42(1): 35–67.
- Valdiserri, Ronald O., David W. Lyder, Laura C. Leviton, Catherine M. Callahan, Lawrence A. Kingsley, and Charles R. Rinaldo. 1988. "Variables Influencing Condom Use in a Cohort of Gay and Bisexual Men." *American Journal of Public Health* 78(7): 801–5.
- Weber, J. Mark, Deepak Malhotra, and J. Keith Murnighan. Under review. "Normal Acts of Irrational Trust and the Attributional Process of Trust Development."
- Williamson, Oliver E. 1979. "Transaction Cost Economics: The Governance of Contractual Relations." *Journal of Law and Economics* 22(0): 3–61.
- . 1996. *The Mechanisms of Governance*. New York: Oxford University Press.

Chapter 13

Untangling the Knot of Trust and Betrayal

SANDRA L. ROBINSON, KURT T. DIRKS, AND HAKAN OZCELİK

TRUST IS critical to organizational effectiveness. Trust enhances cooperation, improves communication, facilitates citizenship behaviors, in addition to improving group and organizational performance (Davis et al. 2000; Dirks 1999; O'Reilly and Roberts 1976; Podsakoff et al. 1990). Despite the importance of trust, however, current organizational environments often challenge the trust that employees bestow on organizations. Indeed, as trends toward downsizing, restructuring, and temporary employment continue, perceptions of unfair treatment (Brockner, Tyler, and Cooper-Schneider 1992), broken contracts (Robinson 1996), and experiences of betrayal will remain a part of the organizational landscape. And although it is an accepted assumption that these organizational changes and actions increase the frequency of breach and betrayal, we know very little about how existing trust in the employment relationship may impact those experiences of betrayal when they do occur.

The purpose of this chapter is to explore the potential relationship between trust and betrayal, with a special focus on how prior trust influences one's experience of breach and betrayal. In doing so, we put forth a theory suggesting that prior trust may either mitigate or enhance the negative effects of betrayal, but that this relationship is dependent upon other conditions in the betrayal context.

Exploration of this question is important for a number of reasons. First, the current literature is unclear about the role of trust in the betrayal experience. As we will explicate, employees' prior trust in management may either intensify the experience of betrayal or inoculate