THE REPAIR OF TRUST: A DYNAMIC BILATERAL PERSPECTIVE AND MULTILEVEL CONCEPTUALIZATION

PETER H. KIM
University of Southern California

KURT T. DIRKS
Washington University in St. Louis

CECILY D. COOPER
University of Miami

We investigate the repair of trust by examining the cognitive and interpersonal processes through which people resolve differences in their interpersonal beliefs. We begin by discussing the phenomenon of trust, the ease with which trust can be violated, and the challenge of trust repair. We then draw from an array of literature to develop a multilevel conceptualization of how trust repair may be pursued. Finally, we integrate these insights to identify three overarching implications for research.

We all rely on beliefs about ourselves, those around us, and the likely course of our interactions to navigate our social worlds. Of these beliefs, perhaps few are as central as those guiding our decisions to trust. Trust, which we define as a psychological state comprising the intention to accept vulnerability based on positive expectations of the intentions or behavior of another (e.g., Rousseau, Sitkin, Burt, & Camerer, 1998), has been found to offer numerous benefits for individuals, groups, and organizations (see Dirks & Ferrin, 2002, for a meta-analytic review). However, these positive expectations often can be violated, resulting in reduced trust and the need to repair trust.

Research across a variety of literature in organizational behavior has focused on the ways in which trust violations occur and the host of problems that result from these transgressions. Relevant inquiries include efforts to understand the nature of trust violations and the dynamics of distrust within and between organizations (Elangovan & Shapiro, 1998; Lewicki, McAllister, & Bies, 1998; Robinson, 1996; Zaheer, Lofstrom, & George, 2002), violations of psychological contracts (Morrison & Robinson, 1997; Robinson &

Rousseau, 1994), revenge and other deviant behaviors that stem from broken relationships (Aquino, Tripp, & Bies, 2001; Bies & Tripp, 1996), and noncooperative behavior in mixed-motive contexts (Bottom, Gibson, Daniels, & Murnighan, 2002). These studies have identified significant economic, emotional, and social costs involved in such situations. Thus, a clear implication of this work is the need to understand how trust might be repaired after such violations have occurred.

Scientific studies, however, have only begun to give the matter of trust repair much attention. A small but growing body of work has directly examined this issue (Ferrin, Kim, Cooper, & Dirks, 2007; Kim, Dirks, Cooper, & Ferrin, 2006; Kim, Ferrin, Cooper, & Dirks, 2004; Nakayachi & Watabe, 2005; Schweitzer, Hershey, & Bradlow, 2006; Sitkin & Roth, 1993). And other research has addressed a number of closely related topics, such as the interpersonal and structural factors that promote forgiveness (Aquino, Tripp, & Bies, 2006; McCullough, Worthington, & Rachal, 1997), responses mistrusted parties can provide to facilitate reconciliation and the restoration of cooperation (Bottom et al., 2002; Tomlinson, Dineen, & Lewicki, 2004), and the use of verbal accounts to mitigate the negative consequences of a violation (Ohbuchi, Kameda, & Agarie, 1989; Riordan, Marlin, & Kellogg, 1983; Sigal, Hsu, Foodim, & Betman, 1988). Despite this initial

progress, this body of research continues to suffer from three critical limitations that, if not addressed, are likely to hinder further advancements in the field.

First, although this literature has identified a number of different strategies for repairing trust, there remains a notable lack of conceptual coherence to the existing research. For example, studies have investigated an assortment of tactics that can be used following a violation, including apologies (e.g., Kim et al., 2006; Tomlinson et al., 2004), denials (e.g., Kim et al., 2006; Sigal et al., 1988), promises (Schweitzer et al., 2006), excuses (Shapiro, 1991; Tomlinson et al., 2004), reparations (Bottom et al., 2002), legalistic remedies (Sitkin & Roth, 1993), hostage posting (Nakayachi & Watabe, 2005), and even no response at all (Ferrin et al., 2007). However, each of these studies has focused on just a single or handful of specific tactics, with no comprehensive theoretical account of how various tactics relate to one another or where other as yet unexamined tactics might belong.

Second, virtually all of the studies in the trust repair literature have focused on the actions that the mistrusted party (i.e., the trustee) might take to repair trust while portraying the trustor as a relatively passive observer. For example, Tomlinson et al. (2004) studied a number of trust repair tactics trustees might initiate while portraying trustors as recipients who simply evaluate these trust repair efforts. By doing so, this study, as well as the others, largely failed to recognize that the trustor often plays an active role in the trust repair process. Not only is the trustor's willingness to accept the mistrusted party's efforts of great importance for determining the likelihood of trust repair (Lewicki & Bunker, 1996) but so too are the actions trustors themselves undertake to influence this outcome. Thus, some conceptual basis is needed to address the complex issue of how these various trustee and trustor efforts interact with one another to determine whether, and the extent to which, the repair of trust occurs.

Third, perhaps in large part because of the aforementioned limitations, the trust repair literature lacks consensus regarding what kinds of efforts are effective at repairing trust, and it has, at times, even reached contradictory conclusions. For example, some studies have suggested that trust may be repaired more successfully if mistrusted parties identify, acknowledge, and assume responsibility for the offense (Lewicki & Bunker, 1996; Ohbuchi et al., 1989), whereas others have suggested that an apology will fail to ameliorate the negative consequences of a trust violation because it involves an acknowledgment of guilt (Riordan et al., 1983; Schlenker, 1980; Sigal et al., 1988). Indeed, this situation may be likened to John Godfrey Saxe's poem about six blind men trying to describe an elephant, who end up arriving at different conclusions and failing to understand the whole because they came into contact with different parts ("Though each was partly in the right, and all were in the wrong!").

The purpose of this paper is to address these limitations by establishing a research program (Lakatos, 1978) on trust repair that considers the cognitive and interpersonal processes through which people may resolve differences in their interpersonal beliefs. We begin by discussing the phenomenon of trust, the ease with which trust can be violated, and the particular challenge of trust repair. Next, we discuss the central ideas that represent the hard core of this research program. We then draw on these foundations to develop more specific elements of our theory, including a progressive bilateral model of how the repair of trust may occur. Finally, we integrate these insights to consider their implications for future research.

TRUST FORMATION, VIOLATION, AND REPAIR

Trust is a complex and multifaceted construct. Scholars have discussed both "trusting intentions"—that is, a willingness to make oneself vulnerable to another in the presence of risk—and "trusting beliefs"—that is, the perceived trust-relevant qualities of the trustee, such as competence, integrity, or benevolence (McKnight, Cummings, & Chervany, 1998). Building on these notions, as well as supporting evidence by Kim et al. (2004), we presume that trust is composed of both trusting beliefs and trusting intentions and that the latter are influenced via the former. Trust violations, in turn, concern incidents that lower these trusting beliefs in and trusting intentions toward a trustee (e.g., owing to trustors' initial assumptions about their trustees' guilt, degree of responsibility, future behavior, and type of transgression). Finally, trust repair concerns improving the trusting beliefs and trusting intentions that have been lowered by the trust violation (typically, by addressing

these initial trustor assumptions). For clarity, throughout this paper we use the term *trustor* to refer to the individual whose trust has been violated (since he or she is in the position of evaluating the mistrusted party) and the term *trustee* to refer to the mistrusted party (since he or she is the target of the trustor's trusting beliefs and trusting intentions).

Dynamics of Trust

It would be one thing if trustworthiness were a tangible resource that could be expended on certain occasions and replenished on others. If that were the case, the repair of trust could simply entail assessing the magnitude of the violation via an objective analysis and meting out the appropriate restitution. Evidence reveals, however, that our assessments of trustworthiness are far from systematic; they are susceptible to a host of social cognitive factors.

It is often assumed, for example, that trust in others develops gradually over time (e.g., Lewicki & Wiethoff, 2000). But recent observations suggest that individuals can exhibit surprisingly high levels of trust even without a history of interaction (McKnight et al., 1998; Meyerson, Weick, & Kramer, 1996; Weber, Malhotra, & Murnighan, 2005). These researchers explain that this high initial trust can arise for a variety of reasons, including an individual's disposition to trust; feelings of dependence; a belief that impersonal structures such as regulations and laws support one's likelihood of success in a given situation; and rapid, cognitive cues arising from group membership, reputations, and stereotypes. These same researchers also note, however, that such trust can be quite fragile because of the tentative and assumption-based nature of these antecedents.

People can often behave in ways that violate trust, such as by exploiting dependencies or by neglecting to fulfill expectations. Yet research suggests that trust can be damaged even when individuals have not committed such infractions. Numerous studies have shown that people can be quite willing to believe unsubstantiated allegations of such acts, despite the difficulties of determining whether the allegations are accurate (e.g., Bell & Loftus, 1989; Penrod & Cutler, 1995; Ross, Ceci, Dunning, & Toglia, 1994). Evidence from these and other studies (e.g., Kim et al., 2004) has also shown that trust can be vio-

lated even with those who have not been directly harmed by the transgression. Moreover, it appears that when such violations occur, the onus is frequently placed on the accused party to remove the shadow of suspicion (Hendry, Shaffer, & Peacock, 1989).

Trust Repair

Those who wish to repair trust, however, typically confront a range of complications. First, since trustworthiness is ultimately in the eye of the beholder, there may be times when trust has been violated but the trustee does not realize that a violation has occurred (e.g., when the trustee violates trust unknowingly or the trustor is not visibly harmed by the transgression). In such cases the trust violation will not necessarily be followed by a trust repair attempt; the latter may require that the trustee be informed of this need by the trustor (or third party). Second, trust repair requires more than simply compensating for the negative expectations arising from a trust violation by improving others (e.g., responding to damaged beliefs about the trustee's integrity by bolstering beliefs about the trustee's competence; Baumeister & Jones, 1978). Such an approach may sometimes encourage trustors to engage in trust-relevant behaviors, but it does so by augmenting nonviolated expectations (i.e., trust building) rather than by addressing the salient negative expectations that have arisen from the violation (i.e., trust repair). Similarly, although parties may initiate a number of legalistic remedies (e.g., policies, procedures, contracts, monitoring) to promote trustworthy behavior (e.g., Nakayachi & Watabe, 2005; Sitkin, 1995; Sitkin & Roth, 1993), such constraint-oriented remedies may not necessarily repair trust itself. Rather than addressing the underlying trustworthiness of the trustee, these remedies seem more specifically concerned with enabling desired outcomes by reducing a trustor's vulnerability or risk (i.e., by allowing trustors to act as if trust existed by limiting the ability of even untrustworthy trustees to commit a transgression; Das & Teng, 1998).[1]

---

[1] In fact, evidence reveals that the implementation of such constraints may even hinder trust, since trustors cannot determine whether the trustee's subsequent behavior is due to the trustee's trustworthiness or the imposed constraint (Malhotra & Murnighan, 2002).

Trust repair, therefore, cannot simply focus on reestablishing seemingly trusting behaviors, since such behavior can arise for a variety of reasons (Kee & Knox, 1970; Lewis & Weigert, 1985: 977). Instead, these considerations reveal that the trust repair process ultimately involves the interaction of both the trustor and trustee as they attempt to resolve discrepancies in their beliefs. Theorizing in this domain must, accordingly, move beyond the nature of the violation, its cost to the trustor, or the trustee's responses in isolation to consider the interplay of these and other factors that can shape how we construe our social worlds. Thankfully, these complexities of social construal have been the object of research in other fields (e.g., the literature on identity negotiation, on the fundamental attribution error, and on dispositional attribution), and this research has established a number of solid foundations on which we can build. The following sections, therefore, will draw on this and other literature to develop a research program that can account for the dynamics of trust repair, specify a bilateral model of how the repair of trust may occur, and, finally, integrate these insights into three overarching implications for future research.

RESEARCH PROGRAM

Our paper is based on the philosophical view that the appropriate unit of scientific growth is best characterized as a continuous progression of historically related theories, or "research program" (Lakatos, 1978). This continuity arises from the nurturing and articulation of a "hard core" of leading ideas, which gives the research program its impetus and originality and which cannot be abandoned without abandoning the research program altogether. Rather, the hard core is shielded from falsification by a "protective belt" of auxiliary theories and hypotheses that are derived from the hard core but can be adjusted, either because of empirical developments or because of natural next steps in the refinement of these ideas.

The research program, furthermore, provides guidance on how to alter this protective belt by deriving a set of problem-solving techniques from the program's inner logic (i.e., its "positive heuristic"). Thus, a research program may be buffered from disconfirmation through adjustments to its protective belt as long as each suc-

cessive link in the resultant chain of theoretical models arises from the program's positive heuristic, rather than ad hoc, and offers greater explanatory power than the one preceding it.[2] We adopt this approach because it permits a conceptualization of trust repair that can grow along with its nascent literature, allows us to distinguish the vital elements of our portrayal (i.e., the hard core) from those we expect to face the brunt of empirical testing and updating (i.e., the protective belt), and sets a rigorous standard for subsequent theory development.

The Hard Core

The hard core of our research program comprises three basic assumptions about the repair of trust following a violation. They concern (1) the notion that trustees want to be considered trustworthy, (2) the inclination of trustors to believe that greater trust in their trustees is not deserved, and (3) the efforts of trustors and trustees to resolve these discrepant beliefs.

Trustees. Our analysis of trust repair begins with the assumption that trustees want to be considered trustworthy and, hence, will advocate the belief, following a violation, that they should be trusted. This assumption is based on two compatible rationales. First, this advocacy may reflect trustees' personal beliefs that trust is truly warranted—for example, if the violation was the result of a groundless allegation or if trustees simply construe the transgression differently owing to their personal knowledge of the incident, self-serving attributions, positive illusions, or other considerations (e.g., Riess, Rosenfeld, Melburg, & Tedeschi, 1981; Taylor & Brown, 1988). Second, this advocacy may arise when trustees recognize that being perceived as

---

[2] A classic example of such a research program can be found in Newton's gravitational theory. This theory posited a hard core containing three laws of dynamics and his law of gravitation, as well as a protective belt of theories and hypotheses that began with a theoretical model for a planetary system with a fixed pointlike sun and one single pointlike planet. Newton then considered a system with more planets, then worked out the case where the sun and planets were not mass points but mass balls, and then followed this by starting work on the implications of spinning balls and their wobbles, admitting the role of interplanetary forces. He then started work on perturbations and then considered the implications of planets that are bulging rather than round (Lakatos, 1978: 50).

trustworthy may allow them to obtain desired resources from trustors, such as esteem, information, assistance, and so on (Dirks & Skarlicki, 2009). Indeed, the latter consideration suggests that there may be cases where trustees do not actually believe they are trustworthy but seek to repair trust anyway (e.g., the case of unwarranted trust repair undertaken by a "con man"). However, even here our analysis assumes that trustees will act as if they believe that greater trust is deserved.

Trustors. In contrast, we assume trustors are predisposed to believe that greater trust in their trustees is not deserved. More specifically, once trust has been violated, conscious concerns about harm from further transgressions, as well as less conscious influences such as the confirmation bias (Nickerson, 1998) and aspects of our evolutionary hardwiring that deter us from tolerating transgressors (de Quervain et al., 2004), should encourage trustors to maintain that mistrust unless they are given an adequate reason to update it. Trustees may attempt to offer that reason with a trust repair attempt, but it is ultimately the trustors who set the threshold (either consciously or unconsciously) for how adequate a response is required. Indeed, even in cases where trustors want to repair their trust in the trustee themselves (e.g., when trustors are sufficiently motivated to preserve the relationship), they must strive to do so in ways that are sufficient to overcome the resistance threshold that they have (again, either consciously or unconsciously) set on their own.

Negotiation of trustworthiness. Finally, our analysis assumes that these competing efforts will create a dynamic through which trustees and trustors attempt to resolve their discrepant beliefs about the trustees' trustworthiness. The identity negotiation literature, which concerns individuals' attempts to resolve discrepancies between how they are viewed by others and how they view themselves (Swann, 1987), provides a foundation for understanding these interactions. This research indicates that parties' competing beliefs can be resolved in different ways—there are times when targets are able to act in ways that disconfirm perceivers' beliefs (Podsakoff & Farh, 1989), occasions when targets fall prey to the beliefs of perceivers and confirm them through their actions (Jussim, Soffin, Brown, Ley, & Kohlhepp, 1992), and other outcomes that fall somewhere in between. This lit-erature has, furthermore, shown that which of these outcomes occurs depends on the relative strength of these competing party beliefs (e.g., Eisenstadt & Leippe, 1994; Kim, Diekmann, & Tenbrunsel, 2003; Nease, Mudgett, & Quinones, 1999). Thus, by depicting the trust repair process as a negotiation of identity between trustors and trustees—specifically, the resolution of a disagreement over whether the trustees should be trusted—we can gauge trust repair by assessing the relative success of their efforts. In particular, these notions entail that trust repair depends on the strength of the trustees' efforts to promote the belief that they should be trusted, outweighing the opposing efforts of the trustors (see Figure 1). Otherwise, a number of alternative outcomes may arise, as we discuss later in the paper.

The Protective Belt

The problem of trust repair underscored by these hard core assumptions thus becomes how trustees may bolster the strength of their efforts to be seen as trustworthy relative to the strength of their trustors' resistance. To solve it, we require a more specific portrayal of the ways in which trust repair can occur. Our research program offers such a portrayal through a bilateral model of trust repair. This model is noteworthy in that each of its levels represents a distinct yet cumulative stage in the elaboration of a theory in a manner that is both consistent with our hard core assumptions and capable of explaining more and more of the empirical world. In this way it exhibits the defining characteristics of a "progressive theoretical [and empirical] problemshift," the satisfaction of which provides a meaningful basis for determining that this research program is worthwhile (Lakatos, 1978: 48–49).

Our bilateral model of trust repair involves several interrelated principles. First, we presume that trustors' and trustees' disagreement about whether the trustees should be trusted can be resolved on multiple levels. These levels represent a logically derived sequence of questions about whether a trustee can be trusted following a violation. Specifically: (1) Is the trustee innocent or guilty of committing the transgression? (2) If the trustee is guilty of the transgression, should this be attributed to the situation or to the person? (3) If the transgression is attributed at least in part

FIGURE 1
Potential Outcomes of Negotiation Efforts by a Trustor and Trustee
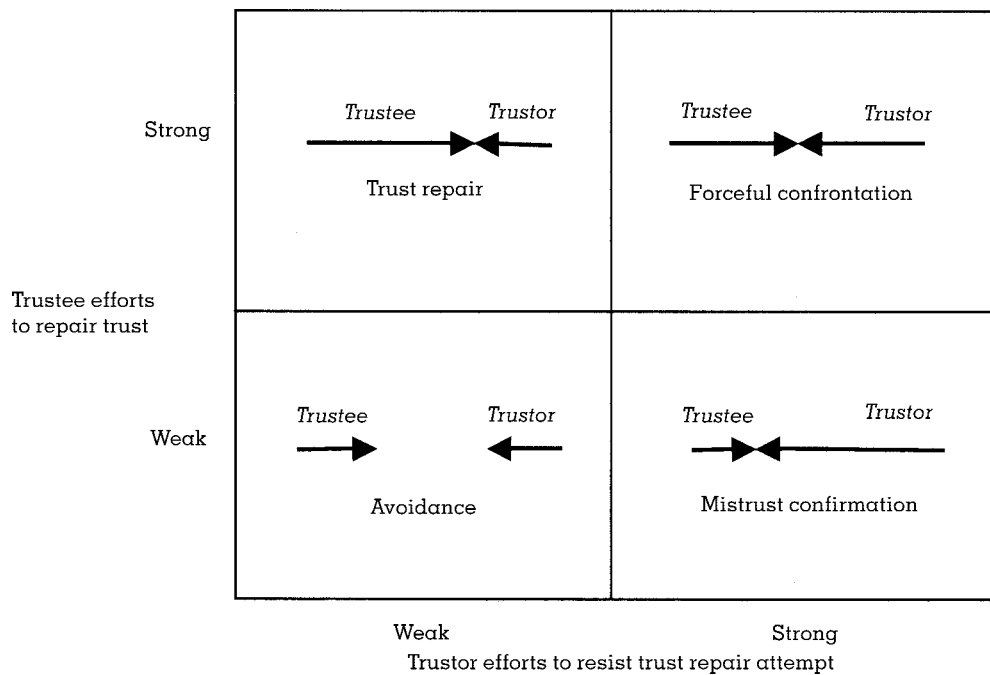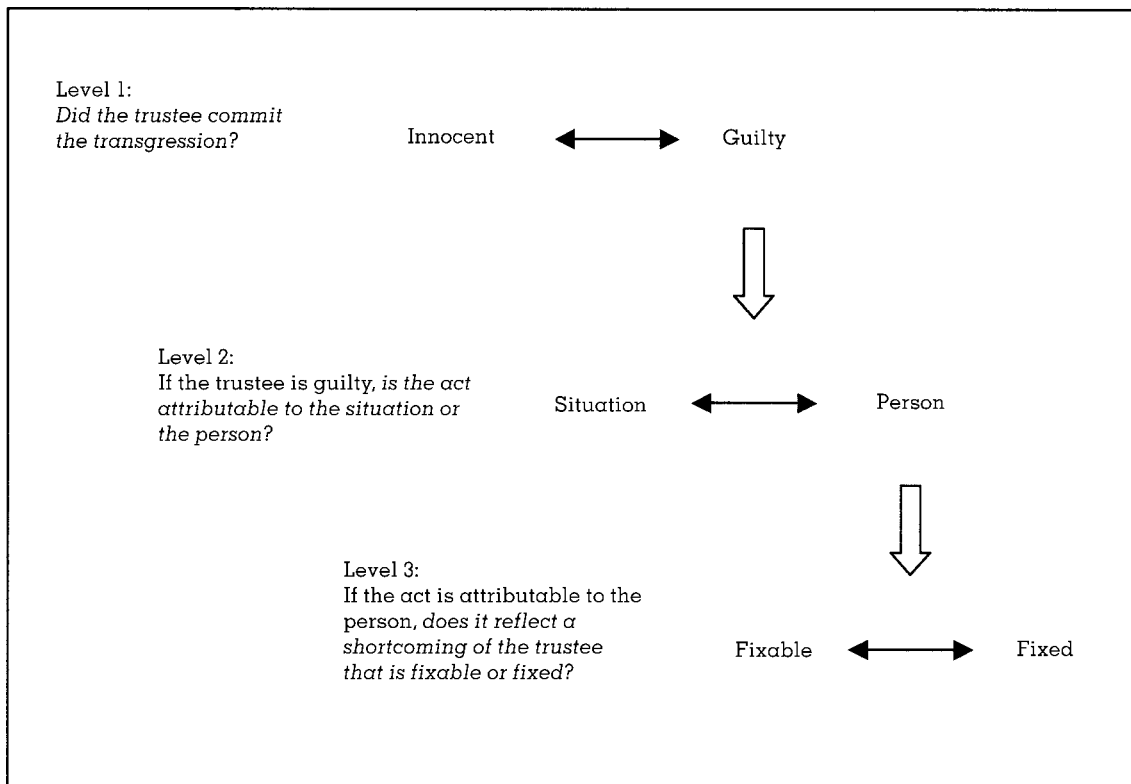

to the person, is the personal shortcoming fixable or is it an enduring characteristic of the trustee? As will become apparent later in the paper, these levels roughly correspond to different segments of the literature, which have explored particular strategies for repairing relationships following a transgression. We refer to them as *levels* because they represent a cascading series of questions that can be ordered by their scope, with each successive level offering a less comprehensive means through which trust may be repaired—for example, Level 1 would be the broadest in scope because it concerns whether the trustee should be completely exonerated from the alleged transgression (see Figure 2).

Second, we expect that, all else being equal, trustees will prefer the broadest level of trust repair possible (e.g., trustees will prefer to be found completely innocent [via Level 1] rather than partially guilty [via Level 2]). This principle represents a corollary to the hard core assumption that trustees want to be considered trustworthy. It derives from the notion that broaderlevel solutions will prove less onerous for a trustee than those that are narrower, not only because of the psychic benefits of being more completely exonerated but also because broader-

level solutions are likely to impose fewer subsequent burdens on the trustee. Success in convincing trustors of one's innocence, for example, will require little in the way of subsequent remedial action, given that the alleged transgression will have been shown to be false, whereas success in convincing trustors that the transgression was at least partly attributable to the situation will entail somehow addressing these concerns (e.g., by mitigating situational influences through new policies or procedures and/or increasing one's resistance to those pressures). Moreover, both of these outcomes should prove less taxing for the trustee than convincing trustors that although the trustee is fully responsible for the transgression, this deficiency can be fixed, since the latter may require some of the most fundamental of changes to the trustee's nature (e.g., with concerted efforts to address these deep-rooted limitations and to show that they have been corrected).

Third, again all else being equal, we expect that trustors will resist broader levels of trust repair more than those that are narrower. This principle represents a corollary to the hard core assumption that trustors will be predisposed to believe that greater trust in their trustees is not deserved. It derives from the notion that

FIGURE 2
Bilateral Model of Trust Repair

Level 1:
Did the trustee commit the transgression? Innocent ⟷ Guilty

Level 2:
If the trustee is guilty, *is the act attributable to the situation or the person?* Situation ⟷ Person

Level 3:
If the act is attributable to the person, *does it reflect a shortcoming of the trustee that is fixable or fixed?* Fixable ⟷ Fixed

narrower-level solutions are likely to provide greater reassurance to trustors than those that are broader, not only by offering greater validation of trustors' belief that their trustees were untrustworthy (e.g., that at least some fault lies with the trustees [Level 2], as opposed to none at all [Level 1]) but also because of the greater subsequent burdens that narrower levels of trust repair will levy on their trustees (as described above). Thus, we presume that trustors will resist certain types of repair efforts more than others and that it is through this "differential resistance" that trustors may exert as significant an influence on the repair of trust as their trustees.

Our fourth principle concerns the resolution of these competing tendencies. As a corollary to the hard core assumption that the repair of trust depends on the relative strength of trustees' and trustors' efforts, we expect that the specific nature of these efforts will reflect the level at which trust repair is pursued. In particular, for Level 1, the degree of trust repair will depend on the extent to which trustees' efforts to promote

the belief that they are innocent of the transgression outweigh their trustors' efforts to uphold the belief that the trustees are guilty. For Level 2, the degree of trust repair will depend on the extent to which trustees' efforts to promote the belief that the transgression should be attributed to the situation outweigh their trustors' efforts to uphold the belief that the transgression should be attributed to the trustees. And for Level 3, the degree of trust repair will depend on the extent to which trustees' efforts to promote the belief that they will redeem themselves in the future outweigh their trustors' efforts to uphold the belief that such redemption will not occur.

Finally, the identity negotiation perspective that helps ground the bilateral trust repair model underscores the notion that trustors' and trustees' competing beliefs will be resolved through a dynamic process that can unfold not only within but also across levels. For example, although trustees may prefer to be seen as innocent, they may come to discover that their trustors adamantly believe they are guilty. In

such cases trustees may be better off shifting their efforts to a narrower-scope level of trust repair (by claiming that they are only partially responsible or that although they are fully responsible, the deficiency can be fixed) so that at least some trust repair occurs, despite the added liabilities the narrower-scope solution entails. Thus, trustees may engage in one or more "strategic retreats" whereby broader trust repair efforts are abandoned, either immediately or after a broader-scope attempt has failed, in favor of narrower-scope efforts that may be more likely to succeed.[3] And, by the same token, trustors may alter the nature of their own resistance (e.g., by more adamantly opposing trustees' claims of innocence than trustees' efforts to attribute some of the blame to the situation), and perhaps even signal where that resistance will be lower (e.g., by asking trustees to explain why they were untrustworthy), to affect the types of repair attempts their trustees pursue.

These principles highlight the potential for an iterative call and response of endeavors, with trustees and trustors each attempting to promote their competing beliefs, gauging the effectiveness of their efforts, and then engaging in additional efforts to influence this outcome further. Indeed, the motivation to initiate such additional efforts may be quite common, given the opposing efforts of trustees and trustors to shift such beliefs in their favor as they attempt to converge on a mutually shared appraisal of the trustees. Thus, for all these reasons, this multilevel conceptualization highlights the dynamic bilateral process through which the repair of trust can occur. Below, we elaborate the model by describing each of its levels and the ways in which trustees and trustors may affect them.

Level 1: Innocent ↔ Guilty

Given that a trust violation is based on the premise that a trustee has committed some form of transgression, perhaps the most comprehensive way in which trust can be repaired is by affecting the extent to which this premise is ultimately deemed to be true (i.e., on a continuum ranging from definitely guilty to definitely innocent). Alleged transgressors who seek trust repair can challenge trustors' beliefs that the trustees had committed the transgression and, to the extent that they are successful, completely vanquish the notion that this lack of trust was deserved. Trustors, in contrast, are likely to resist such innocence-claiming efforts in favor of maintaining their belief in the trustees' guilt, thereby impeding such trust repair.

Studies have revealed a number of ways in which trustees and trustors may pursue these opposing efforts. One of the most obvious means of doing so is to provide tangible evidence concerning a trustee's innocence or guilt regarding the alleged transgression. Kim et al. (2004) examined this notion as part of a larger inquiry and found, in a simulated job interview at an accounting firm, that candidates who had been accused of misfiling a tax return with a prior employer were indeed considered more trustworthy—and were more likely to be hired—after an investigation revealed they were innocent, rather than guilty, of this alleged violation.

Research has also revealed that a trustor's beliefs about a trustee's guilt can be influenced through less tangible means, such as the provision of verbal denials. Sigal et al. (1988), for example, asked participants to watch a videotape of a simulated debate in which one political candidate was accused of sexual or financial misconduct by the other, and they found that the accused party received more votes and was considered to be more honest, ethical, and trustworthy when that candidate denied culpability rather than apologized for the misconduct. Similarly, Riordan et al. (1983) used fabricated reports that a fictitious senator had taken a bribe and found that subsequent evaluations of the senator were less negative when the senator denied rather than admitted responsibility for the transgression.

A trustee's verbal denial can be problematic, however, when it is inconsistent with subsequent evidence regarding the trustee's innocence/guilt. Indeed, Kim et al. (2004) observed that information about guilt after a denial indicates that the denial was a lie and, thus, fosters lower levels of trust than if the trustee had of-

---

[3] A well-known example can be found in the shifting response of former President Bill Clinton following the accusation that he had an affair with a White House intern. He initially denied guilt (Level 1): "I did not have sexual relations with that woman, Miss Lewinski." But when this offering became untenable, he shifted to another tactic (Level 2): "I did have a relationship with Miss Lewinski that was not appropriate. . . . I can only tell you I was motivated by many factors."

fered an apology. This finding highlights the potential for different trust repair efforts to interact, and even to interfere, with one another, as well as the resultant need to assess their combined effects on parties' competing beliefs to determine whether such efforts will succeed.

This depiction also suggests that trustors are not simply passive bystanders but, rather, are actively involved in this negotiation process. Trustors may therefore be able to impose their own beliefs in such a way that trustees abandon their trust repair efforts or even come to believe that this lack of trust is truly deserved. Research on the social psychology of false confessions, for example, has shown that although many of us may find it hard to believe that anyone would confess to a crime he or she did not commit (Kassin & Wrightsman, 1980, 1981), there are numerous instances in which such confessions have been obtained (Bedau & Radelet, 1987; Rattner, 1988). Thus, scholars have begun to investigate how interrogators, who earnestly believe in the suspect's guilt, manage to obtain such an end (Kassin & McNall, 1991). Although systematic evidence regarding such techniques is scarce, these inquiries highlight at least the possibility that trustors' beliefs about their trustees' guilt can affect trustees' beliefs about themselves.

Therefore, we can observe that both trustors and trustees can initiate tactics to affect the likelihood of trust repair. Each of these tactics can, furthermore, be understood to operate by affecting the strength of the trustors' and/or trustees' beliefs about the trustees' innocence/ guilt. Thus far, research pertinent to this strategy has largely focused on the provision of tangible evidence and verbal denials. However, the aforementioned principles that underlie this approach also suggest that these specific tactics are likely to represent just a few of the techniques that can be employed. Indeed, trustees and trustors may each engage in a range of as yet unexamined efforts to strengthen their own beliefs and/or to weaken the beliefs of the other. Level 1 (innocent ↔ guilty) thus concerns whether the net effect of such efforts strengthens the beliefs of trustees and/or trustors, and it highlights how the degree of trust repair may depend on the extent to which trustees' efforts to promote the belief that they are innocent of the transgression outweigh trustors' efforts to uphold the belief that the trustees are guilty.

Yet even if it is clear to all parties that a trustee is guilty, trust may still be repaired through other means. To do so, however, the bilateral trust repair model suggests that the trustee must narrow the scope of his or her repair efforts in a way that concedes this matter of guilt but attempts to limit its impact on how the trustee is viewed. Specifically, in the following section we consider a whole class of efforts designed not to challenge trustees' responsibility for a transgression completely (as described above) but, rather, to acknowledge having committed it while also deflecting as much of the blame as possible to the situation.

Level 2: Situation ↔ Person

Researchers have observed that when considering the underlying cause of an action, observers attempt to subtract the effect of the situation and attribute what remains to the actor (Kelley, 1973). The logic that drives this "discounting principle" is that dispositional and situational forces operate in a hydraulic fashion so that as situational forces grow stronger, the role of the actor grows weaker (McClure, 1998). A problem, however, is that when making such assessments, observers generally fail to fully appreciate the potential influence of situational forces and, hence, err on the side of making dispositional attributions—a phenomenon known as the fundamental attribution error (e.g., Jones & Harris, 1967; Nisbett, Caputo, Legant, & Maracek, 1973).

The targets of such observations, on the other hand, generally possess a much greater appreciation of the constraints imposed by the situation than observers (e.g., Jones & Nisbett, 1972; Nisbett et al., 1973). In some cases such situational attributions arise because these external forces are more salient to actors than to observers (i.e., an informational mechanism). In other cases, particularly after a negative event, actors may be inclined, either consciously or unconsciously, to adopt such situational attributions in order to preserve their self-esteem—a motivational mechanism (Snyder & Higgins, 1988; Snyder, Higgins, & Stucky, 1983). But regardless of the underlying mechanism, this research highlights an inherent difference in how actors (e.g., trustees) and observers (e.g., trustors) are likely to interpret behavior, and they show that an actor's culpability can depend on the degree to

which the actor is able to convince the observer that the behavior was induced by the situation (i.e., on a continuum ranging from entirely the actor to entirely the situation).

One way trustees might seek to mitigate their blame, and thus facilitate trust repair, is by explaining that their actions were at least partly caused by external forces (i.e., by providing an excuse). Indeed, evidence from several studies indicates that excuses can often prove beneficial for those who have committed a wrong (Snyder & Higgins, 1988) because they reduce the perceived responsibility of the transgressor (Riordan et al., 1983). Crant and Bateman (1993), for example, discovered that when supervisors in a large accounting firm read scenarios that described an unsuccessful audit, they blamed the subordinate less if an external causal account was offered than if such an account was not offered. Wood and Mitchell (1981) found that when nurse managers read scenarios depicting their subordinates erring in patient care, they assigned less responsibility and punished less severely when they were given accounts of external causes for the poor performance than when they were given an apology. Moreover, Weiner, Amirkhan, Folkes, and Verette (1987) found that external attributions can reduce anger in the victim of a transgression and thereby reduce tension in potentially inflammatory situations.

Alternatively, trustees may attempt to reframe the situation in a way that convinces trustors to reassess the magnitude or nature of the transgression itself (i.e., by providing some form of justification). Justifications not only involve the acceptance of responsibility but also point out that the act in question was appropriate owing to the nature of the situation (Scott & Lyman, 1968). Thus, in most cases justifications refer the trustor to a situational factor, such as a norm, that the trustor may not have recognized but that when taken into consideration reflects a positive or appropriate motive. Nesdale, Rule, and McAra (1975), for example, found that observers were more likely to approve of a harmful action when the action was explained to be the result of a good rather than bad motive. Similarly, Shapiro (1991) found that those who were deceived by their partners were less punitive and less likely to retaliate if altruistic rather than selfish reasons were given for the deception. Moreover, Elsbach's (1994) study of organizational accounts in the California cattle industry observed

that efforts to justify controversial actions by referencing normative and socially endorsed organizational practices (e.g., federally approved guidelines) provided an effective means of managing organizational legitimacy.

A thorough assessment of Level 2 (situation ↔ person) also moves us beyond the traditional focus on a trustee's role in shifting responsibility to the situation by underscoring a competing force that has received much less attention. Specifically, when trustees attempt to highlight the influence of the situation with an excuse or justification, trustors may often be inclined to resist such efforts. Research in the field of evolutionary psychology suggests that people are generally predisposed to infer that some agent, rather than the situation, is the cause of salient events (Boyer, 2001: 144–145). Kim and colleagues (2006) have also noted that trustors may discount mistrusted parties' efforts to blame situational influences. Moreover, as with the potential for "mistrust confirmation" described by Level 1 (innocent ↔ guilty), excuses may sometimes backfire. For example, some evidence suggests that when individuals offer an excuse for a failure but the excuse is rejected, they are likely to exhibit performance deficits in the future and in this way to confirm with their subsequent actions the negative feedback they received (Abramson, Seligman, & Teasdale, 1978; Mikulincer & Nizan, 1988).

These considerations underscore the notion that "every account is a manifestation of the underlying negotiation of identities" (Scott & Lyman, 1968: 59). Trustees may use excuses, justifications, and perhaps other as yet unexamined measures to convince trustors (and perhaps even themselves) that the transgression should be attributed to the situation rather than to the trustees' inherent lack of trustworthiness, whereas trustors may implement a range of measures (virtually none of which have received systematic research attention) to oppose such claims and attribute the transgression to trustees' untrustworthy dispositions. Level 2 (situation ↔ person) therefore reveals how the degree of trust repair may depend on the extent to which trustees' efforts to promote the belief that the transgression should be blamed on the situation outweigh trustors' efforts to uphold the belief that the transgression should be blamed on the trustees.

Deflecting blame from a trustee's disposition to the situation represents an opportunity to re-

pair trust even when the trustee has clearly committed the violation and, hence, finds it infeasible to directly oppose the underlying basis for the mistrust. With both Levels 1 (guilty ↔ innocent) and 2 (situation ↔ person), however, the overriding goal is to reduce the mistrusted party's perceived culpability in some way. But this general approach to repairing trust has been challenged by a separate set of studies suggesting that trust repair is more likely to occur if trustees assume *more* blame (e.g., with an apology or an internal attribution) rather than less (e.g., with a denial or external attribution). The bilateral trust repair model resolves this apparent contradiction by suggesting that these other studies represent a trust repair strategy that is even narrower in scope than the prior two levels. In particular, it suggests that even if it proves infeasible to oppose all culpability or even some of it, trust may still be repaired if trustees can confine blame to the past and make the case that the problems will be corrected in the future.

Level 3: Fixable ↔ Fixed

The key elaboration posed by this level is that although all trust repair requires the trustee to be considered more trustworthy in the future, this may be accomplished through measures that are either "retroactively prophylactic" (i.e., minimizing the initial taint of untrustworthiness) or "subsequently curative" (i.e., fixing the problem after it has been found) in nature. Retroactively prophylactic measures, such as those described by Levels 1 and 2, have the potential to repair trust more comprehensively than those that are substantively curative. The former are designed to reduce concerns about the future by reducing trustees' culpability for the past, whereas the latter are only designed to target concerns about the future. Nevertheless, the kinds of subsequently curative methods that we describe may still repair trust effectively to the extent that the question of whether these flaws will persist in the future (i.e., on a continuum ranging from entirely fixable to entirely fixed) represents trustors' paramount concern.

Discussions within the trust repair literature regarding these more future-oriented concerns have suggested that they may be alleviated if mistrusted parties identify, acknowledge, and assume some responsibility for the trust-

damaging events (Lewicki & Bunker, 1996). One way of doing so is by offering an apology, which is a statement that acknowledges both responsibility and regret for a trust violation (Ferrin et al., 2007; Kim et al., 2006; Kim et al., 2004). Although an apology acknowledges guilt, which alone should reduce trust, its concomitant expression of regret signals an intention to avoid similar violations in the future, which should reduce trustors' concerns about continued vulnerability and should thereby improve trust. Thus, we can observe that the effectiveness of an apology as a response to a trust violation depends on the notion that this response's benefits (due to potential redemption) will outweigh its costs (due to the confirmation of guilt).

This reasoning has been supported by several empirical studies. Findings indicate that victims of psychological harm generally have more favorable impressions of the perpetrator, experience more positive affect, and are more likely to refrain from severe aggression toward the perpetrator when the culprit apologizes for the wrongdoing than when the culprit does not (Ohbuchi et al., 1989). Experimental studies of impression management reveal that the expression of remorse following a transgression can reduce punishment (e.g., Schwartz, Kane, Joseph, & Tedeschi, 1978). Moreover, research on social dilemmas demonstrates that, at least in short-term interactions, an apology can more effectively reestablish cooperation after an opportunistic act than can a denial (Bottom et al., 2002).

Additionally, even when an apology is offered, trustees may convey it in a manner that assumes more blame or less, and at least some evidence reveals that those who assume full blame with an internal attribution are seen as more likely to correct their shortcomings in the future (i.e., to achieve redemption) than those who attempt to mitigate their blame with an external attribution. Tomlinson et al. (2004) found that victims of a broken promise were more willing to reconcile a professional relationship when the violator offered an internal rather than an external attribution. Likewise, Hodgins and Liebeskind (2003) reported that victims exhibited more positive evaluations and expectations of future relationships when transgressors assumed more rather than less responsibility for the act in question. Furthermore, Schlenker, Pontari, and Christopher (2001) noted

that excuse makers risk being seen as deceptive, self-absorbed, and ineffectual.

Moving beyond the implications of verbal responses, studies suggest that individuals may implement a variety of substantive measures to address concerns about future transgressions. Bottom et al. (2002) proposed and found supporting evidence for the notion that *voluntarily* paying a financial penalty following a transgression will substantiate the trustee's expression of remorse for his or her behavior and will underscore a commitment to avoid similar transgressions in the future. Similarly, Nakayachi and Watabe (2005) found across several contexts that the *voluntary* implementation of a monitoring system and sanctions helped to restore trust following a transgression. They suggested that these voluntary behaviors (i.e., not simply the imposition of a monitoring system and sanctions) signal positive intentions by the trustee and, thus, reduce expectations of subsequent transgressions.

Finally, evidence reveals that the effectiveness of all of these future-oriented responses will depend on how the transgression has been framed—specifically, as a matter of competence or integrity. Competence (i.e., the extent to which one possesses the technical and interpersonal skills required for a job) and integrity (i.e., the extent to which one adheres to a set of principles that a perceiver finds acceptable) deserve particular attention when considering the challenges of trust repair, for at least three reasons. First, numerous researchers have observed that competence and integrity represent two of the most important qualities for determining trustworthiness (Barber, 1983; Butler & Cantrell, 1984; Schindler & Thomas, 1993). Second, research has shown that these dimensions offer important bases on which individuals evaluate a variety of targets, including leaders (Pancer, Brown, & Barr, 1999), job candidates (Cook & Elmer, 1999), and potential collaborators (Kee & Knox, 1970). Third, evidence suggests that there may be some inherent differences in the way people assess positive versus negative information about competence versus integrity that may affect how they resolve differences in their interpersonal perceptions (see Snyder & Stukas, 1999, for a review). In particular, research suggests that although individuals tend to weigh positive information about competence more heavily than negative information about competence,

they tend to weigh negative information about integrity more heavily than positive information about integrity (Reeder & Brewer, 1979).

These differences in how people assess positive versus negative information about competence and integrity offer a critical foundation for understanding why trustees may be more or less capable of repairing trust through claims of future redemption. In particular, the notion that people tend to weigh positive information about competence more heavily than negative information about competence suggests that when trust violations concern matters of competence, trustees' signals that their limitations will be addressed (i.e., positive competence information) may be sufficient to allay concerns about their guilt (i.e., negative competence information). In contrast, the notion that people weigh negative information about integrity more heavily than positive information about integrity suggests that when trust violations concern matters of integrity, concerns about trustees' prior guilt (i.e., negative integrity information) will outweigh trustees' signals that those limitations will be addressed (i.e., positive integrity information). Thus, recent studies have shown that apologies, efforts to assume full blame with an internal attribution, and voluntary substantive actions such as penance and regulation (all of which convey positive information that these problems will be fixed) tend to repair trust more effectively when the transgression is attributed to matters of competence rather than integrity (Dirks, Kim, Cooper, & Ferrin, 2005; Ferrin et al., 2007; Kim et al., 2006; Kim et al., 2004).

These results were also observed despite the fact that within each study the actual trust violation was the same; each was simply framed as a competence- or integrity-related matter. This observation highlights the notion that the underlying bases for trust violations can be ambiguous and possibly even involve elements of both competence and integrity. Therefore, trustees and trustors may each attempt to influence this framing (i.e., with trustees advocating attributions of competence and trustors advocating attributions of integrity) to support their competing views about whether greater trust in the trustee is deserved.

Overall, then, these findings highlight an opportunity to repair trust through responses that directly address concerns about the extent to which the underlying problems may be cor-

rected in the future. Trustees may apologize and may do so by making an internal (dispositional) rather than external (situational) attribution, may voluntarily institute substantive measures to address concerns about future transgressions, may strengthen these efforts by framing the transgression as a matter of competence, or may engage in other (as yet unexamined) endeavors to convince trustors that they will be trustworthy in the future, whereas trustors may attempt to frame the transgression as a matter of integrity or may implement a range of other as yet neglected measures to resist such trustee efforts and maintain lower expectations of trustee redemption. Level 3 (fixable ↔ fixed) therefore reveals how the degree of trust repair may depend on the extent to which trustees' efforts to promote the belief that they will redeem themselves for the future outweigh their trustors' efforts to uphold the belief that such redemption will not occur.

Thus, we can observe that each of these three levels represents a distinct means of trust repair. Each level identifies a discrete set of opposing forces and provides the basis for a distinct stream of empirical research, yet each also builds cumulatively on the others. In this way these levels may be likened to the discrete yet complementary elements of some of the most prominent conceptualizations in our field (e.g., the outcome value and outcome alternatives dimensions described in power-dependence theory [Emerson, 1962; Kim & Fragale, 2005; Kim, Pinkley, & Fragale, 2005b] or the integrative and distributive dimensions that provide the conceptual underpinnings of negotiation research [Kim et al., 2003; Lax & Sebinius, 1986; Neale & Northcraft, 1991]). As with those other areas of inquiry, the proposed multilevel conceptualization can provide the foundation for extensive elaborative research. Thus, to guide such efforts, we identify three overarching theoretical implications that can serve as guiding principles—"positive heuristics" in the terminology of Lakatos (1978)—for pursuing this future work.

Theoretical Implications

Consolidation and broadening of research on trust repair tactics. One implication of the present analysis is that it provides a basis not only for understanding the implications of previously studied trust repair tactics but also for assessing the potential implications of tactics that have yet to be explored. These opportunities arise from recognizing that these tactics are important not in and of themselves but, rather, because of how they affect the relative strength of parties' competing beliefs with regard to the three levels of the bilateral trust repair model. Hence, this framework offers the opportunity to consolidate a wide array of research, account for how previously studied tactics would relate to one another (i.e., by classifying them according to the level at which they would operate and whether they would bolster the relative strength of the trustee's or trustor's beliefs), and then explore how a much broader range (and myriad variations) of such initiatives may achieve similar ends.

What would be the impact on beliefs about innocence or guilt, for example, of trustees or trustors soliciting the support of neutral third parties via direct appeals, publicity, or rumor? How would variations of different types of apologies (e.g., the different types of apologies described by Schlenker & Darby, 1981) influence the relative strength of beliefs about whether the trustees' personal shortcomings would be addressed? Could, as Schweitzer et al. (2006) have speculated, such tactical variations help explain why apologies have been found to repair trust in some studies but not their own? And how might such effects also depend on the expression of specific emotions? These sorts of inquiries could contribute not only by revealing these potential techniques (e.g., verbal, tangible, cognitive, emotional, individual, collective) but also by assessing their psychological implications (i.e., the relative impact of each technique on the focal belief and how this may depend on the way it is pursued) and, even further, by assessing their additive and interactive effects (e.g., the effect of an apology on its own compared to an apology plus substantive repentance).

Indeed, by guiding these efforts, we hope our analysis will also begin to rectify the conspicuous lack of attention the trust repair literature has paid to the tactics trustors might employ. As is the case with trustee tactics, the three levels of the bilateral trust repair model provide the basis for understanding what issues need to be addressed and allow for a broad range of trustor influences on these considerations. Furthermore, we may help guide the process of identifying these trustor influences by drawing on in-

sights from the expectancy literature. This literature is based on the notion that we can create our own social reality by influencing the behaviors we observe in others such that we may induce them to confirm even our erroneous expectations if we behave in a manner consistent with those beliefs—a phenomenon commonly known as the self-fulfilling prophecy (Jones, 1986). Hence, studies in this domain have identified at least three mechanisms that may inform research on trustor tactics.

First, research indicates that perceiver expectancies may be confirmed when perceivers act in ways that reduce targets' opportunities to exhibit disconfirming behavior (Rothbart, 1981). This suggests that trustors may impede trust repair by constraining their trustees in a manner that would restrict trustees' ability to offer exculpating evidence (Level 1: innocent ↔ guilty), highlight the significance of situational influences (Level 2: situation ↔ person), or demonstrate how their trust-inhibiting limitations would be fixed (Level 3: fixable ↔ fixed)—for example, by denying trustees the chance to respond to an allegation, giving trustees insufficient time to explain their side of the story, or cutting off all subsequent contact with trustees.

Second, studies have shown that perceivers may encode or interpret targets' behavior in a way that confirms perceivers' expectations (Darley & Fazio, 1980; Strenta & Kleck, 1984). Hence, trustors may maintain their beliefs in their trustees' untrustworthiness by construing trustees' repair efforts in a way that would discount information offered to support trustees' innocence (Level 1: innocent ↔ guilty), the influence of situational forces (Level 2: situation ↔ person), or the possibility that trustees' limitations would be fixed (Level 3: fixable ↔ fixed)—for example, by ignoring this information, questioning its significance, or identifying countervailing information.

Third, evidence reveals that perceiver expectations are more likely to be confirmed when there are greater anticipated costs and/or fewer anticipated benefits from modifying perceivers' beliefs (Zanna & Pack, 1975). This finding is consistent with research on false confessions (as mentioned with regard to Level 1: innocent ↔ guilty) in which scholars observed that suspects were more likely to confess to a crime they did not commit when they were given offers of sympathy, tolerance, face-saving excuses, and moral justifications for those alleged offenses

(Kassin & McNall, 1991). And, more broadly, it suggests that trustors may affect trust repair by altering the balance of costs and benefits trustees associate with attempts to repair trust at any given level of the bilateral trust repair model—for example, by expressing stronger emotions to bolster the power of their accusation, by signaling the kinds of responses they would be more or less willing to accept, or by threatening greater repercussions if certain points of contention are not confirmed.

Beyond individual tactics: Relational determinants of success. Second, by emphasizing the active roles that trustees and trustors play in the trust repair process, the negotiation of identity that underlies each level of the bilateral trust repair model can reconcile why even the exact same trust repair effort may repair trust in some cases but not others. Closer scrutiny of the identity negotiation perspective suggests that even if trustees initiate trust repair tactics that exert a strong influence on a given level of trust repair, their ultimate implications for trust will depend on the tactics trustors implement as well. Indeed, this notion may be portrayed via the same 2 (trustee influence: strong versus weak) × 2 (trustor influence: strong versus weak) matrix, regardless of the level of trust repair pursued (see Figure 1).

This depiction reveals that whereas powerful efforts by the trustee (e.g., a fervent denial) may achieve some degree of trust repair if the trustor's efforts are weak (e.g., minimal or no response), these same trustee efforts may do little to repair trust if the trustor's efforts are strong (e.g., a strident accusation), and instead may produce some form of "forceful confrontation." Alternatively, whereas weak trust repair efforts by the trustee (e.g., a half-hearted denial) and strong opposing efforts by the trustor (e.g., providing evidence of guilt) may result in a form of "mistrust confirmation," whereby trustees come to believe that they are untrustworthy and even confirm this expectation with their behavior, weak efforts by both the trustee and trustor (e.g., with neither attempting to assert or substantiate their beliefs) may result in the persistence of mistrust through a lower-energy stalemate than that described by forceful confrontation, an outcome more akin to "avoidance."

Such interactions highlight the importance of considering the active role played by the trustor, in addition to the trustee, in determining the

likelihood of trust repair (and, by doing so, may help explain inconsistencies in the efficacy of any given trust repair tactic). They also reveal the potential value of investigating factors that could affect the strength of each party's efforts (e.g., whether trustors who are motivated to trust their trustees [Weber et al., 2005] would exert weaker efforts to resist trustees' repair efforts and become more likely to embrace them than trustors who lack such motivations). Furthermore, they illustrate the need for research to move beyond the basic question of whether trust has been repaired to give the three alternative outcomes more meaningful attention.

Beyond a silo approach: Multilevel considerations. Finally, this multilevel analysis reveals how a thorough account of trust repair must do more than simply focus on one or another of these levels (i.e., the implicit approach of prior research), since each provides an incomplete portrayal of what may ultimately transpire. Instead, the bilateral trust repair model highlights the need to investigate the various conditions under which these different levels should be pursued, and the need to provide an account of the dynamic process of shifting across levels. On a practical basis, this recognition may be crucial to the extent that at least some of each level's associated trust repair efforts can affect, or even preclude, the use of others.[4] Moreover, on a conceptual basis, it allows us to assess how the interpersonal context may influence the relative effectiveness of these three levels for trust repair (and, hence, explain why even if a trustee implements a highly influential tactic, and this tactic is stronger than that of the trustor, it may still fail to repair trust if the tactic operates on a level that, owing to these contextual influences, becomes less significant than others). Given the potential range of such contextual influences, their full account falls beyond the scope of the

_____

[4] Level 1 (innocent ↔ guilty) and Level 2 (situation ↔ person), for example, represent mutually exclusive options in the sense that a trustee cannot logically deny having committed a transgression and also explain that situational forces induced that commission. Level 3 (fixable ↔ fixed) can operate in an additive manner with Level 2 (situation ↔ person)—for example, the trustee can explain that situational forces encouraged the transgression while apologizing for the act. Finally, Level 3 (fixable ↔ fixed) is incompatible with Level 1 (innocent ↔ guilty) in the sense that a trustee cannot logically deny having committed the transgression and also apologize for committing it.

present work. Thus, below we attempt simply to illustrate such possibilities with an example of how power, which has been identified as a key determinant of the negotiation of identities (Snyder & Stukas, 1999), may affect these decisions.

Power has been broadly defined as the probability that a person can carry out his or her own will despite resistance (Weber, 1947). Research on power has identified a wide range of factors that may affect its construal (see Kim et al., 2005, for a review), including the value of benefits from a given relationship and the value of alternatives to that relationship (Emerson, 1962), as well as a party's expertise, legitimacy, likability, ability to reward or punish (French & Raven, 1959), and even network centrality (Brass, 1992). However, with regard to the negotiation of identity underlying trust repair, this illustrative analysis is concerned with how, after being determined by such features of the interpersonal context, the power of one party may affect the other and through these means influence how the repair of trust may most effectively be pursued.

One such implication is that as a party's power in a relationship grows, so too should its ability to affect the relationship partner's perceptions. Less powerful parties have been observed to be more dependent on the benefits from their relationships than their more powerful counterparts and, hence, more responsive to their partners' cues—presumably, to minimize negative outcomes from those with power over their fates (Geiss, 1993). Thus, targets have been found to confirm their perceivers' perceptions when these perceivers had the power to control targets' outcomes, whereas when targets had power to control perceivers' outcomes, such confirmation did not occur (Copeland, 1994). These considerations, therefore, highlight the potential influence of power on trust repair by suggesting that the extent to which trustees can directly challenge trustors' beliefs about the trustees' guilt (i.e., through Level 1: innocent ↔ guilty) should increase as their power relative to the power of their trustors grows.

The fundamental nature of power as the probability that a person can carry out his or her own will despite resistance, however, also suggests that those possessing it may face difficulty convincing others that their transgressions were somehow induced by the situation. Evidence reveals that people are generally inclined to ascribe control and responsibility for important

but causally indeterminant outcomes to those in positions of power (Meindl, Ehrlich, & Dukerich, 1985). Hence, perceivers have been found to place greater blame on their targets, rather than the situation, when targets' power is high rather than low (Overbeck, Tiedens, & Brion, 2006). Thus, these findings highlight the importance of power for efforts to repair trust through the mitigation of trustees' blame by suggesting that the ability of trustees to do so (i.e., through Level 2: situation ↔ person) should decrease as their power relative to that of their trustors grows.

Moreover, with such increases in trustee power, their transgressions may increasingly be seen by trustors to arise from a lack of integrity rather than a lack of competence. Lord Acton's well-known adage that "power tends to corrupt and absolute power corrupts absolutely" underscores the widely held suspicion that the powerful are inclined to act in self-serving ways. And though such tendencies are not without exception (Chen, Lee-Chai, & Bargh, 2001), a substantial body of evidence has supported the propensity of powerholders to further enrich themselves at the expense of others. Individuals with a relative advantage in power have been found, for example, to extract a greater share of benefits from their interactions than their less powerful counterparts, to devalue the ability and worth of the less powerful, and even to view the less powerful as objects of manipulation (Georgesen & Harris, 1998; Kim, 1997; Kim & Fragale, 2005; Kipnis, 1972). Thus, to the extent that trustors are aware of such inclinations, they should be particularly disposed to attribute the transgressions of high-power trustees to a lack of integrity (Fragale, Rosen, Xu, & Merideth, 2009), as opposed to a lack of competence, and such attributions of low integrity should be particularly difficult to overcome with signals that their integrity will improve in the future (Ferrin et al., 2007; Kim et al., 2006; Kim et al., 2004). These notions, accordingly, highlight the importance of power for trustees' efforts to repair trust by suggesting that their limitations will be corrected in the future (i.e., through Level 3: fixable ↔ fixed), since the efficacy of this approach should decrease as their power relative to that of their trustors grows.

Thus, we can observe that the implications of power for the repair of trust are not necessarily beneficial. Whereas the viability of trust repair through Level 1 (innocent ↔ guilty) should in-crease, its viability through Level 2 (situation ↔ person) and Level 3 (fixable ↔ fixed) should decrease as the power of trustees relative to that of their trustors grows. We might therefore expect that to the extent trustees possess greater power than their trustors, the trustees should repair trust more successfully by directly challenging trustors' belief that the trustees are guilty of the transgression than through any of the other levels of trust repair (i.e., by making situational attributions or by claiming that these problems will be fixed). This reasoning also implies, perhaps counterintuitively, that since trustees possess *less* power relative to that of their trustors, they should have more success at repairing trust via Levels 2 and 3 but not Level 1. Such possibilities, along with many others that have yet to be explored, highlight the need to consider the interpersonal context and its likely effects on all three levels before we can determine how the repair of trust would most likely occur.

CONCLUSION

The purpose of this paper was to elucidate a research program on trust repair by examining the cognitive and interpersonal processes through which people attribute dispositions to others. Although the matter of trust repair has drawn increasing scientific attention, the field has lacked an adequate theoretical account of why certain repair efforts are more likely to succeed than others. Thus, we have sought to consolidate these prior efforts, reconcile their apparent discrepancies, and provide a rigorous foundation on which future research could be based. Through these endeavors we have developed a multilevel model of the dynamic bilateral process through which the repair of trust may be pursued and then identified several overarching theoretical implications to guide future research.

The cumulative nature of these levels should not be overlooked. Although each offers additional explanatory power, this is achieved by building onto, rather than challenging, the key insights of the others. Indeed, every one of these levels can provide the basis for a distinct set of research questions. Level 1's depiction of trust repair as a function of efforts by trustors and trustees to directly oppose their counterpart's beliefs about the trustees' guilt/innocence sug-

gests a range of opportunities to investigate what may affect the relative strength of these competing forces. Level 2's elaboration concerning how the repair of trust might unfold not only through direct opposition but also through deflection (i.e., of the blame that might otherwise be placed on a trustee) highlights an entirely distinct set of inquiries researchers might pursue to explore what would facilitate or impair such deflection efforts. Finally, Level 3's intertemporal elaboration, which reveals how trust may be repaired by acknowledging one's blame but then claiming that these problems can be corrected, reveals the trade-offs involved in many trust repair responses (i.e., due to their focus on one's prior blame versus potential redemption) and, hence, raises a third set of questions that researchers might explore.

Moreover, we can observe that this multilevel portrayal offers the basis for addressing several critical limitations in this field (as described in the introduction). First, regarding the lack of conceptual coherence in the trust repair literature, this depiction allows us to consolidate a wide array of research, account for how previously studied tactics would relate to one another (i.e., by classifying them according to the level at which they would operate and whether they would bolster the relative strength of the trustee's or trustor's beliefs; see Figure 2), and then proceed to explore how a much broader range (and myriad variations) of such initiatives may achieve such ends. Second, in response to this nascent literature's tendency to treat trustors simply as passive observers, the present conceptualization highlights the active roles trustors can play in the trust repair process, reveals how little we know about the specific tactics these trustors might implement, and underscores how the efforts of both trustees and trustors must be considered to determine whether (and the extent to which) the repair of trust would occur (see Figure 1). Finally, this analysis provides the basis for resolving the troublesome inconsistencies that have begun to appear in the literature by revealing how such contradictions can be reconciled by (1) considering the myriad ways in which these parties might implement a given tactic (see "Consolidation and broadening of research on trust repair tactics"), (2) revealing how the exact same tactic can either succeed or fail depending on the strength of the trustor's efforts (see Figure 1), and (3) illus-

trating the potential for the interpersonal context to affect the relative importance of the model's levels for trust repair (i.e., such that the implementation of an otherwise influential tactic may fail to repair trust if it operates on a level that, because of these contextual influences, becomes less viable than others).

Limitations and Future Directions

Having proposed this research program, we must also acknowledge that it remains far from complete. First, although attributions of competence and integrity have been identified as two of the most important dimensions of trust, researchers have identified benevolence as another important antecedent (e.g., McAllister, 1995). Therefore, it would be useful to consider whether attributing a trust violation to a lack of benevolence rather than a lack of competence or integrity (as discussed with regard to Level 3: fixable ↔ fixed) would produce meaningful differences for trust repair.[5] With regard to this issue, one study suggests that when people evaluate matters of benevolence, they do not exhibit the kinds of asymmetries we discussed with regard to matters of competence or integrity and, hence, neither weigh negative information about benevolence as heavily as negative information about integrity nor weigh positive information about benevolence as heavily as positive information about competence (Trafimow & Trafimow, 1999). If so, we might expect that benevolence attributions would fall somewhere between those of competence and integrity with regard to determining whether the trust-inhibiting qualities of the trustee were ultimately deemed to be fixable versus fixed.

Second, it is important to recognize that most of the trust repair studies reported in this paper

---

[5] The neglect of benevolence attributions in the trust repair literature may be given a number of speculative explanations, including (1) the incipient nature of this literature (e.g., even research on the implications of competence and integrity attributions for trust repair has itself only just begun), (2) the literature's focus thus far on newly formed relationships, which may not offer an optimal context for benevolence attributions to play a role, and (3) the lack of a substantive theoretical basis for considering how benevolence attributions compare to attributions of competence and integrity for trust repair (although we attempt to remedy this issue by leveraging the research of Trafimow and Trafimow [1999]).

have focused on newly formed relationships. Thus, research is needed to assess whether the present reasoning will generalize to relationships of longer duration. This does not imply that the present analysis should be dismissed, since this research program's various theoretical underpinnings (e.g., identity negotiation, interpersonal perception, and attribution theory) have, themselves, been validated with longer-term relationships. Moreover, it is essential to note that our organizational relationships can often be quite impersonal in nature. Network research, for example, has repeatedly emphasized that organizational members possess far more weak ties—that is, with relative strangers—than strong ties—that is, with those with whom they have close personal relationships (e.g., Granovetter, 1995). Additionally, research on fault lines provides the basis for expecting that it is far more likely that trust will be violated in relationships of shorter rather than longer duration, because parties in newly formed relationships are less likely to have had the opportunity to develop mutual understandings (e.g., Swann, Milton, & Polzer, 2000). Therefore, although the extent to which one can generalize research on newly formed (or impersonal) relationships to relationships of longer duration remains a concern, the present conceptualization represents a useful starting point for research on trust repair that, like many theories in the organizational sciences, simply requires further testing and elaboration.

Third, our focus on broader mechanisms has entailed our overlooking potentially meaningful individual differences that might provide insight into the likely outcomes of various trust repair efforts (e.g., whether some of us are simply more willing to accept apologies than others). Moreover, other than noting the need to explore how different emotions may affect the strength of trustors' and trustees' beliefs, our cognitive focus has led us to pay little attention to the potentially critical role that emotions can play in trust decisions (Dunn & Schweitzer, 2005) and, hence, the implications of affect for trust repair. Finally, this analysis has not examined the influence of a broader array of situational factors that may affect the importance of these three levels, such as the nature of parties' relationships or the effects of national or organizational culture.

Broader Implications

Despite these limitations, we hope this conceptual framework will provide a useful starting point for pursuing a wide array of new research questions, practitioner-oriented decisions, and subsequent conceptual elaborations. One key implication that has been underscored by this work is the potential for trustors to play a far more active role in the trust repair process than prior research has described. Indeed, the opportunities for trustors not only to respond to their trustees' trust repair efforts but also to implement various efforts on their own (to support their beliefs regarding their trustees' untrustworthiness, or even dispel them when sufficiently motivated to preserve the relationship) illustrate the vital need to account for the initiatives of both trustees and trustors in any analysis of trust repair.

A second implication is that by revealing the potential for the interpersonal context to affect the viability of its three levels, this framework highlights the need to broaden our attention beyond the efforts of trustees and trustors to consider how the larger context may affect the likelihood of trust repair. Until now, these contextual influences have been almost entirely ignored by the literature, since it has focused on gauging the effects of specific trust repair efforts for largely undifferentiated actors. However, to the extent that various aspects of the context can significantly bolster, diminish, or even reverse the effects of such responses, they may ultimately deserve as much attention as the specific tactics parties employ.

Third, we must acknowledge the potential for these insights to be used for both good and ill. Those who wish to repair trust may certainly offer the best response at their disposal based on an honest assessment of the facts. However, some individuals may choose instead to ignore the truth in favor of supplying a response based simply on its efficacy. Thus, we encourage researchers and practitioners alike to explore this domain with caution, moral fortitude, and sensitivity. Despite this concern, researchers must recognize that temptations to repair trust inappropriately exist regardless of whether we know much about the process or not. With such knowledge, each side (both trustor and "potentially" unethical trustee) should be better able to make informed decisions in this escalating "arms

race." Therefore, given the choice, this option of maximizing knowledge for both sides strikes us as far more acceptable than that of sticking our heads in the sand.

A final implication of this work is that it has allowed us to integrate a seemingly disparate set of findings from a wide array of literature and to convey how these findings might form a cohesive whole. It is in this sense that this research program responds to an often-levied criticism that research in organizational behavior represents more of a garbage can of findings than a body of knowledge that in any way builds. As the present effort reveals, the potential for such accumulation in our field certainly exists, the pursuit of which can help reconcile a range of apparent contradictions. Thus, we hope that future research efforts will take note of such opportunities and, when possible, strive for such convergence.

REFERENCES

Abramson, L. Y., Seligman, M. E. P., & Teasdale, J. 1978. Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology,* 87: 49–74.

Aquino, K., Tripp, T. M., & Bies, R. J. 2001. How employees respond to personal offense: The effects of blame attribution, victim status, and offender status on revenge and reconciliation in the workplace. *Journal of Applied Psychology,* 86: 52–59.

Aquino, K., Tripp, T. M., & Bies, R. J. 2006. Getting even or moving on? Power, procedural justice, and types of offense as predictors of revenge, forgiveness, reconciliation, and avoidance in organizations. *Journal of Applied Psychology,* 91: 653–668.

Barber, B. 1983. *The logic and limits of trust.* New Brunswick, NJ: Rutgers University Press.

Baumeister, R. F., & Jones, E. E. 1978. When self-presentation is constrained by the target's knowledge: Consistency and compensation. *Journal of Personality and Social Psychology,* 36: 608–618.

Bedau, H., & Radelet, M. 1987. Miscarriages of justice in potentially capital cases. *Stanford Law Review,* 40: 21–179.

Bell, B. E., & Loftus, E. F. 1989. Trivial persuasion in the courtroom: The power of (a few) minor details. *Journal of Personality and Social Psychology,* 56: 669–679.

Bies, R. J., & Tripp, T. M. 1996. Beyond distrust: Getting even and the need for revenge. In R. M. Kramer & T. R. Tyler (Eds.), *Trust in organizations: Frontiers of theory and research:* 246–260. Thousand Oaks, CA: Sage.

Bottom, W. P., Gibson, K., Daniels, S., & Murnighan, J. K. 2002. When talk is not cheap: Substantive penance and expressions of intent in rebuilding cooperation. *Organization Science,* 13: 497–513.

Boyer, P. 2001. *Religion explained: The evolutionary origins of religious thought.* New York: Basic Books.

Brass, D. J. 1992. Power in organizations: A social network perspective. In G. Moore & J. A. Whitt (Eds.), *Research in politics and society,* vol. 4: 295–323. Greenwich, CT: JAI Press.

Butler, J. K., Jr., & Cantrell, R. S. 1984. A behavioral decision theory approach to modeling dyadic trust in superiors and subordinates. *Psychological Reports,* 55: 19–28.

Chen, S., Lee-Chai, A. Y., & Bargh, J. A. 2001. Relationship orientation as a moderator of the effects of social power. *Journal of Personality and Social Psychology,* 80: 173–187.

Cook, T., & Elmer, N. 1999. Bottom-up versus top-down evaluations of candidates' managerial potential: An experimental study. *Journal of Occupational and Organizational Psychology,* 72: 423–439.

Copeland, J. T. 1994. Prophecies of power: Motivational implications of social power for behavioral confirmation. *Journal of Personality and Social Psychology,* 67: 264–277.

Crant, J. M., & Bateman, T. S. 1993. Assignment of credit and blame for performance outcomes. *Academy of Management Journal,* 36: 7–27.

Darley, J. M., & Fazio, R. H. 1980. Expectancy confirmation processes arising in the social interaction sequence. *American Psychologist,* 35: 867–881.

Das, T. K., & Teng, B.-S. 1998. Between trust and control: Developing confidence in partner cooperation in alliances. *Academy of Management Review,* 23: 491–512.

de Quervain, D. J. F., Fischbacher, U., Trever, V., Schellhammer, M., Schnyder, U., Buck, A., & Fehr, E. 2004. The neural basis of altruistic punishment. *Science,* 305: 1254–1258.

Dirks, K. T., & Ferrin, D. L. 2002. Trust in leadership: Meta-analytic findings and implications for organizational research. *Journal of Applied Psychology,* 87: 611–628.

Dirks, K. T., Kim, P. H., Cooper, C. D., & Ferrin, D. L. 2005. *Understanding the effects of substantive responses on trust following a transgression.* Paper presented at the annual meeting of the Academy of Management, Honolulu.

Dirks, K. T., & Skarlicki, D. 2009. The relationship between being perceived as trustworthy by coworkers and individual performance. *Journal of Management,* 35: 136–157.

Dunn, J. R., & Schweitzer, M. E. 2005. Feeling and believing: The influence of emotion on trust. *Journal of Personality and Social Psychology,* 88: 736–748.

Eisenstadt, D., & Leippe, M. R. 1994. The self-comparison process and self-discrepant feedback: Consequences of learning you are what you thought you were not. *Journal of Personality and Social Psychology,* 67: 611–626.

Elangovan, A. R., & Shapiro, D. L. 1998. Betrayal of trust in organizations. *Academy of Management Review,* 23: 547–566.

Elsbach, K. D. 1994. Managing organizational legitimacy in

the California cattle industry: The construction and effectiveness of verbal accounts. *Administrative Science Quarterly,* 39: 57–88.

Emerson, R. M. 1962. Power-dependence relations. *American Sociological Review,* 27: 31–40.

Ferrin, D. L., Kim, P. H., Cooper, C. D., & Dirks, K. T. 2007. Silence speaks volumes: The effectiveness of reticence in comparison to apology and denial for responding to integrity- and competence-based trust violations. *Journal of Applied Psychology,* 92: 893–908.

Fragale, A. R., Rosen, B., Xu, C., & Merideth, I. 2009. The higher they are, the harder they fall: The effects of wrongdoer status on observer punishment recommendations and intentionality attributions. *Organizational Behavior and Human Decision Processes,* 108: 53–65.

French, J. R. P., Jr., & Raven, B. 1959. The bases of social power. In D. Cartwright (Ed.), *Studies in social power:* 150–167. Ann Arbor: University of Michigan Press.

Geiss, F. L. 1993. Self-fulfilling prophecies: A social psychological view of gender. In A. E. Beall & R. J. Sternberg (Eds.), *The psychology of gender*: 9–54. New York: Guilford Press.

Georgesen, J. C., & Harris, M. J. 1998. Why's my boss always holding me down? A meta-analysis of power effects on performance evaluations. *Personality and Social Psychology Review,* 2: 184–195.

Granovetter, M. 1995. *Getting a job: A study in contacts and careers* (2nd ed.). Chicago: University of Chicago Press.

Hendry, S. H., Shaffer, D. R., & Peacock, D. 1989. On testifying on one's own behalf: Interactive effects of evidential strength and defendant's testimonial demeanor on mock jurors' decisions. *Journal of Applied Psychology,* 74: 539–545.

Hodgins, H. S., & Liebeskind, E. 2003. Apology versus defense: Antecedents and consequences. *Journal of Experimental Social Psychology,* 39: 297–316.

Jones, E. E. 1986. Interpreting interpersonal behavior: The effects of expectancies. *Science,* 234: 41–46.

Jones, E. E., & Harris, V. A. 1967. The attribution of attitudes. *Journal of Experimental Social Psychology,* 3: 1–2.

Jones, E. E., & Nisbett, R. E. 1972. The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior*: 79–94. Morristown, NJ: General Learning Press.

Jussim, L., Soffin, S., Brown, R., Ley, J., & Kohlhepp, K. 1992. Understanding reactions to feedback by integrating ideas from symbolic interactionism and cognitive evaluation theory. *Journal of Personality and Social Psychology,* 62: 402–421.

Kassin, S. M., & McNall, K. 1991. Police interrogations and confessions: Communicating promises and threats by pragmatic implication. *Law and Human Behavior,* 15: 233–251.

Kassin, S. M., & Wrightsman, L. S. 1980. Prior confessions and mock juror verdicts. *Journal of Applied Social Psychology,* 10: 133–146.

Kassin, S. M., & Wrightsman, L. S. 1981. Coerced confessions, judicial instruction, and mock juror verdicts. *Journal of Applied Social Psychology,* 11: 489–506.

Kee, H. W., & Knox, R. E. 1970. Conceptual and methodological considerations in the study of trust and suspicion. *Conflict Resolution,* 14: 357–366.

Kelley, H. H. 1973. The process of causal attribution. *American Psychologist,* 28: 107–128.

Kim, P. H. 1997. Strategic timing in group negotiations: The implications of forced entry and forced exit for negotiators with unequal power. *Organizational Behavior and Human Decision Processes,* 71: 263–286.

Kim, P. H., Diekmann, K. A., & Tenbrunsel, A. E. 2003. Flattery may get you somewhere: The strategic implications of providing positive vs. negative feedback about ability vs. ethicality in negotiation. *Organizational Behavior and Human Decision Processes,* 90: 225–243.

Kim, P. H., Dirks, K. T., Cooper, C. D., & Ferrin, D. L. 2006. When more blame is better than less: The implications of internal vs. external attributions for the repair of trust after a competence- vs. integrity-based trust violation. *Organizational Behavior and Human Decision Processes,* 99: 49–65.

Kim, P. H., Ferrin, D. L., Cooper, C. D., & Dirks, K. T. 2004. Removing the shadow of suspicion: The effects of apology vs. denial for repairing ability- vs. integrity-based trust violations. *Journal of Applied Psychology,* 89: 104–118.

Kim, P. H., & Fragale, A. R. 2005. Choosing the path to bargaining power: An empirical comparison of BATNAs and contributions in negotiation. *Journal of Applied Psychology,* 90: 373–381.

Kim, P. H., Pinkley, R. L., & Fragale, A. R. 2005. Power dynamics in negotiation. *Academy of Management Review,* 30: 799–822.

Kipnis, D. 1972. Does power corrupt? *Journal of Personality and Social Psychology,* 24: 33–41.

Lakatos, I. 1978. *The methodology of scientific research programmes: Philosophical papers.* Cambridge: Cambridge University Press.

Lax, D. A., & Sebinius, J. K. 1986. *The manager as negotiator: Bargaining for cooperation and competitive gain.* New York: Free Press.

Lewicki, R. J., & Bunker, B. B. 1996. Developing and maintaining trust in work relationships. In R. M. Kramer & T. R. Tyler (Eds.), *Trust in organizations: Frontiers of theory and research*: 114–139. Thousand Oaks, CA: Sage.

Lewicki, R. J., McAllister, D. J., & Bies, R. J. 1998. Trust and distrust: New relationships and realities. *Academy of Management Review,* 23: 438–458.

Lewicki, R. J., & Wiethoff, C. 2000. Trust, trust development, and trust repair. In M. Deutsch & P. Coleman (Eds.), *The handbook of conflict resolution: Theory and practice:* 86–107. San Francisco: Jossey-Bass.

Lewis, J. D., & Weigert, A. 1985. Trust as a social reality. *Social Forces,* 63: 967–985.

Malhotra, D., & Murnighan, J. K. 2002. The effects of contracts on interpersonal trust. *Administrative Science Quarterly,* 47: 534–559.

McAllister, D. J. 1995. Affect- and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of Management Journal,* 38: 24–59.

McClure, J. 1998. Discounting causes of behavior: Are two reasons better than one? *Journal of Personality and Social Psychology,* 74: 7–20.

McCullough, M. E., Worthington, E. L., Jr., & Rachal, K. C. 1997. Interpersonal forgiving in close relationships. *Journal of Personality and Social Psychology,* 73: 321–336.

McKnight, D. H., Cummings, L. L., & Chervany, N. L. 1998. Initial trust formation in new organizational relationships. *Academy of Management Review,* 23: 473–490.

Meindl, J. R., Ehrlich, S. B., & Dukerich, J. M. 1985. The romance of leadership. *Administrative Science Quarterly,* 30: 78–102.

Meyerson, D., Weick, K. E., & Kramer, R. M. 1996. Swift trust and temporary groups. In R. M. Kramer & T. R. Tyler (Eds.), *Trust in organizations: Frontiers of theory and research:* 166–195. Thousand Oaks, CA: Sage.

Mikulincer, M., & Nizan, B. 1988. Causal attribution, cognitive interference, and the generalization of learned helplessness. *Journal of Personality and Social Psychology,* 55: 470–478.

Morrison, E. W., & Robinson, S. L. 1997. When employees feel betrayed: A model of how psychological contract violation develops. *Academy of Management Review,* 22: 226–256.

Nakayachi, K., & Watabe, M. 2005. Restoring trustworthiness after adverse events: The signaling effects of voluntary "hostage posting" on trust. *Organizational Behavior and Human Decision Processes,* 97: 1–17.

Neale, M. A., & Northcraft, G. B. 1991. Behavioral negotiation theory: A framework for conceptualizing dyadic bargaining. *Research in Organizational Behavior,* 13: 147–190.

Nease, A. A., Mudgett, B. O., & Quinones, M. A. 1999. Relationships among feedback sign, self-efficacy, and acceptance of performance feedback. *Journal of Applied Psychology,* 84: 806–814.

Nesdale, A. R., Rule, B. G., & McAra, M. 1975. Moral judgments of aggression: Personal and situational determinants. *European Journal of Social Psychology,* 5: 339–349.

Nickerson, R. S. 1998. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology,* 2: 175–220.

Nisbett, R. E., Caputo, C., Legant, P., & Maracek, J. 1973. Behavior as seen by the actor and as seen by the observer. *Journal of Personality and Social Psychology,* 27: 154–164.

Ohbuchi, K., Kameda, M., & Agarie, N. 1989. Apology as aggression control: Its role in mediating appraisal of and response to harm. *Journal of Personality and Social Psychology,* 56: 219–227.

Overbeck, J. R., Tiedens, L. Z., & Brion, S. 2006. The powerful want to, the powerless have to: Perceived constraint moderates causal attributions. *European Journal of Social Psychology,* 36: 479–496.

Pancer, S., Brown, S., & Barr, C. 1999. Forming impressions of political leaders: A cross-national comparison. *Political Psychology,* 20: 345–368.

Penrod, S., & Cutler, B. 1995. Witness confidence and witness accuracy: Assessing their forensic relation. *Psychology, Public Policy, and Law,* 1: 817–845.

Podsakoff, P. M., & Farh, J.-L. 1989. Effects of feedback sign and credibility on goal setting and task performance. *Organizational Behavior and Human Decision Processes,* 44: 45–67.

Rattner, A. 1988. Convicted but innocent: Wrongful conviction and the criminal justice system. *Law and Human Behavior,* 12: 283–293.

Reeder, G. D., & Brewer, M. B. 1979. A schematic model of dispositional attribution in interpersonal perception. *Psychological Review,* 86: 61–79.

Riess, M., Rosenfeld, P., Melburg, V., & Tedeschi, J. T. 1981. Self-serving attributions: Biased private perceptions and distorted public descriptions. *Journal of Personality and Social Psychology,* 41: 224–231.

Riordan, C. A., Marlin, N. A., & Kellogg, R. T. 1983. The effectiveness of accounts following transgression. *Social Psychology Quarterly,* 46: 213–219.

Robinson, S. L. 1996. Trust and breach of the psychological contract. *Administrative Science Quarterly,* 41: 574–599.

Robinson, S. L., & Rousseau, D. M. 1994. Violating the psychological contract: Not the exception but the norm. *Journal of Organizational Behavior,* 15: 245–259.

Ross, D. F., Ceci, S. J., Dunning, D., & Toglia, M. P. 1994. Unconscious transference and mistaken identity: When a witness misidentifies a familiar but innocent person. *Journal of Applied Psychology,* 79: 918–930.

Rothbart, M. 1981. Memory processes and social beliefs. In D. L. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior:* 145–181. Hillsdale, NJ: Lawrence Erlbaum Associates.

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. 1998. Not so different after all: A cross-discipline view of trust. *Academy of Management Review,* 23: 393–404.

Schindler, P. L., & Thomas, C. C. 1993. The structure of interpersonal trust in the workplace. *Psychological Reports,* 73: 563–573.

Schlenker, B. R. 1980. *Impression management: The self-concept, social identity, and interpersonal relations.* Monterey, CA: Brooks/Cole.

Schlenker, B. R., & Darby, B. W. 1981. The use of apologies in social predicaments. *Social Psychology Quarterly,* 44: 271–278.

Schlenker, B. R., Pontari, B. A., & Christopher, A. N. 2001. Excuses and character: Personal and social implica-

tions of excuses. *Personality and Social Psychology Review,* 5: 15–32.

Schwartz, G., Kane, T., Joseph, J., & Tedeschi, J. T. 1978. The effects of remorse on the reactions of a harm-doer. *British Journal of Social Psychology,* 17: 293–297.

Schweitzer, M. E., Hershey, J. C., & Bradlow, E. T. 2006. Promises and lies: Restoring violated trust. *Organizational Behavior and Human Decision Processes,* 101: 1–19.

Scott, M. B., & Lyman, S. M. 1968. Accounts. *American Sociological Review,* 33: 46–62.

Shapiro, D. L. 1991. The effects of explanations on negative reactions to deceit. *Administrative Science Quarterly,* 36: 614–630.

Sigal, J., Hsu, L., Foodim, S., & Betman, J. 1988. Factors affecting perceptions of political candidates accused of sexual and financial misconduct. *Political Psychology,* 9: 273–280.

Sitkin, S. B. 1995. On the positive effect of legalization on trust. In R. J. Bies, R. J. Lewicki, & B. H. Sheppard (Eds.), *Research on negotiations in organizations:* 185–217. Greenwich, CT: JAI Press.

Sitkin, S. B., & Roth, N. L. 1993. Explaining the limited effectiveness of legalistic "remedies" for trust/distrust. *Organization Science,* 4: 367–392.

Snyder, C. R., & Higgins, R. L. 1988. Excuses: Their effective role in the negotiation of reality. *Psychological Bulletin,* 104: 23–35.

Snyder, C. R., Higgins, R. L., & Stucky, R. J. 1983. *Excuses: Masquerades in search of grace.* New York: Wiley.

Snyder, M., & Stukas, A. A., Jr. 1999. Interpersonal processes: The interplay of cognitive, motivational, and behavioral activities in social interaction. *Annual Review of Psychology,* 50: 273–303.

Strenta, A. C., & Kleck, R. E. 1984. Physical disability and the perception of social interaction: It's not what you look at but how you look at it. *Personality and Social Psychology Bulletin,* 10: 279–288.

Swann, J. W. B. 1987. Identity negotiation: Where two roads meet. *Journal of Personality and Social Psychology,* 53: 1038–1051.

Swann, J. W. B., Milton, L. P., & Polzer, J. T. 2000. Should we create a niche or a fall line? Identity negotiation and small group effectiveness. *Journal of Personality and Social Psychology,* 79: 238–250.

Taylor, S. E., & Brown, J. D. 1988. Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin,* 103: 193–210.

Tomlinson, E. C., Dineen, B. R., & Lewicki, R. J. 2004. The road to reconciliation: Antecedents of victim willingness to reconcile following a broken promise. *Journal of Management,* 30: 165–187.

Trafimow, D., & Trafimow, S. 1999. Mapping perfect and imperfect duties onto hierarchically and partially restrictive trait dimensions. *Personality and Social Psychology Bulletin,* 25: 687–697.

Weber, J. M., Malhotra, D., & Murnighan, J. K. 2005. Normal acts of irrational trust: Motivated attributions and the trust development process. *Research in Organizational Behavior,* 26: 75–101.

Weber, M. 1947. *The theory of social and economic organization.* New York: Free Press.

Weiner, B., Amirkhan, J., Folkes, V. S., & Verette, J. A. 1987. An attributional analysis of excuse giving: Studies of a naive theory of emotion. *Journal of Personality and Social Psychology,* 52: 316–324.

Wood, R. E., & Mitchell, T. R. 1981. Manager behavior in a social context: The impact of impression management on attributions and disciplinary actions. *Organizational Behavior and Human Decision Processes,* 28: 356–378.

Zaheer, A., Lofstrom, S., & George, V. 2002. Interpersonal and organizational trust in alliances. In F. Contractor & P. Lorange (Eds.), *Cooperative strategies and alliances: What we know 15 years later:* 347–377. London: Elsevier Science.

Zanna, M. P., & Pack, S. J. 1975. On the self-fulfilling nature of apparent sex differences in behavior. *Journal of Experimental Social Psychology,* 11: 583–591.

Peter H. Kim (kimpeter@usc.edu) is an associate professor of management and organization at the University of Southern California's Marshall School of Business. He received his Ph.D. in organizational behavior from Northwestern University's Kellogg Graduate School of Management. His research concerns the dynamics of interpersonal perceptions and their implications for workgroups, negotiations, and dispute resolution.

Kurt T. Dirks (dirks@wustl.edu) is a professor of organizational behavior at the Olin Business School at Washington University in St. Louis. He received his Ph.D. from the University of Minnesota. His research is focused on trust in the workplace. He also conducts research on feelings of ownership.

Cecily D. Cooper (cecily@miami.edu) is an assistant professor of management at the University of Miami School of Business. She received her Ph.D. in organizational behavior from the Marshall School of Business at the University of Southern California. Her research focuses on trust—particularly, how individuals and organizations build and maintain the trust of perceivers—as well as the role of fairness and social influence in the development of manager-subordinate relationships.