

# Chapter 1

## Trust and Distrust in Organizations: Dilemmas and Approaches

RODERICK M. KRAMER AND KAREN S. COOK

**F**OR MORE than a decade now, the topic of trust has been at the center of scholarly research on organizations. The ascension of trust as a major focus of research in the organizational sciences reflects in no small measure a large body of evidence documenting the substantial and varied benefits that accrue when trust is in place within organizational boundaries. Coupled with this strong evidence of the benefits of trust is an acute appreciation of the problematic nature of organizational trust. Although trust may be a desirable resource, it is often a fragile and elusive one.

Interest in trust as an important dimension of organizational functioning was initially stimulated by a number of influential works on social trust, including most notably Robert Putnam's (1993) provocative findings implicating trust as a critical factor in explaining the origins of civic engagement and the development of democratic regimes in Italian communities. Putnam's argument attracted wide attention and was followed up by Francis Fukuyama's (1995) comparative survey of evidence that trust plays a critical role in societal functioning. In the years following these initial landmark volumes, a number of additional works (for example, Cook 2001; Hardin 2002; Putnam 2000; Sztompka 1999) have added to the stock of studies of trust in the social sciences. Viewed in the aggregate, these contributions converge on the conclusion that trust often constitutes an important resource within social systems.

Although much of the initial evidence regarding the benefits of trust as a social resource pertained, strictly speaking, to social systems writ large, the organizational implications of these arguments and empirical findings did not go unnoticed (for example, Kramer and Tyler 1996; Lane and Bachman 1998; Sitkin et al. 1998). Within organizational settings, the virtues of trust as a social resource have been discussed primarily on three levels. The first major theme has been its constructive effect with respect to reducing transaction costs within organizations. Second, the role trust plays in spontaneous sociability among organizational members has been explored. Third, there has been appreciation of how trust facilitates appropriate (that is, adaptive) forms of deference to organizational authorities. Much of the appreciation of the merits of trust as a social resource, it should be noted, followed from a corresponding disenchantment with traditional organizational theories of managerial "command and control." Organizational theories emphasizing authority and hierarchy seemed less readily applicable to the flatter and less centralized organizational forms that began to proliferate in recent decades. Moreover, they were seen as less relevant for organizations embedded in high-velocity environments in which rapid change necessitated swift assessment and action. Finally, the socially differentiated character of contemporary organizations presented special challenges to conventional notions of organizing and managing social relations with harmony.

In concert, these trends helped push theory and research on trust into the mainstream of the academic literature on organizations. In addition, the topic took on new relevance for American citizens after the events of September 11, 2001. Almost immediately, questions were raised about the extent to which they should or could trust the systems on which they rely for well-being and security. Almost immediately, the American public was confronted with a series of sudden assaults on its trust in organizations. The stunning collapse of Enron, Worldcom, and Arthur Andersen, and the Catholic church scandals forced a nationwide search for answers to fundamental questions about trust and trustworthiness. These events, coming so close together and from so many distinct arenas of social life, assaulted our confidence in the trustworthiness of the organizational systems on which we rely. In some cases, these events have generated deep ambivalence and even pervasive distrust, which challenges the fundamental legitimacy of professional and managerial authority.

A primary aim of the present volume is to help us take stock of where we stand on these questions. Construed broadly, the chapters in this book all examine problems and prospects for trust within and across organizational boundaries. In putting this diverse collection together, we had a number of goals. The first was to contribute to the development of a better understanding of the antecedents and consequences of trust in

contemporary organizational contexts. Organizational forms and practices are undergoing rapid change and reinvention, raising important new questions about how trust operates across these innovative forms. Thus, one goal of this book is to assess the current state of theory and research on problems of trust and distrust in organizational settings.

We also wanted to bring together scholars who study trust from a variety of different perspectives with a penchant for different kinds of data. In particular, we wanted to bring together in one volume researchers from a wide range of disciplines, including social psychologists, sociologists, political scientists, and economists. Given the complexity of trust-related phenomena in organizations, we felt it was essential to encourage conversations that cross disciplinary boundaries.

We also included in our volume researchers who, although they may share common theoretical interests, nonetheless approach their questions using different methodologies. Thus, we made an effort to achieve a mix of both quantitative and qualitative methods of research. Finally, and perhaps most important, we were committed to exploring problems and prospects for trust within the context of new, interesting, and important organizational settings and forms. Thus, we included scholars who study problems of trust and distrust in a variety of diverse and interesting contexts, such as doctor-patient relationships, social work relationships, geographically dispersed teams, virtual teams, and the Internet. Our emphasis on context was motivated by our strong presumption that the contours of trust and distrust are likely to vary in subtle but important and interesting ways across organizational settings. In their theorizing, academics understandably value and seek parsimonious explanations and generalized accounts of phenomena. As a countermeasure to this push for parsimony and generalization, however, we felt it was important at this stage of theorizing to seek out the distinctive features of trust-related phenomena as they appear in different organizational arenas. We took seriously Hardin's (2002) argument that trust is a three-part relation defined not only by the characteristics of the truster and trustee, but also the specific transactional domain or context in which their relation is embedded. According to Hardin's (2002) "encapsulated interest" account, an actor A trusts actor B with respect to some specific behavior. Context helps to determine the extent to which A considers B to be trustworthy in terms of the matter at hand.

Organizing such a diverse set of chapters presented something of a challenge, especially given the deliberately diverse mix of theorists and their distinctive interests and concerns. Obviously there are a number of different cross-cutting dimensions that we might have used to cluster some chapters together and to separate others. We ultimately settled on an organizational scheme that focuses on the context in which the trust relationship is embedded and the challenges surrounding the construction of trust in various settings. In part I are the chapters that deal with trust

in relationships that can be classified as primarily hierarchical such as employer-employee, supervisor-worker, professional-client, and doctor-patient relationships. Issues of motivation, monitoring, commitment, and trust loom large in these chapters. At the heart of these investigations are the following questions: Do trust and power really mix—and if so how? What are the determinants of trustworthiness in hierarchical relations? How is trust maintained in the face of power differences?

In part II we move to the more collective level to focus on the role of trust between individuals or organizational agents in teams and networks. How does trust develop in distributed work settings, especially when team members are geographically dispersed? How do we determine the trustworthiness of those we do not see or interact with frequently on a face-to-face basis? Does trust create a safety net for team failure? These are some of the questions addressed in this section of the volume.

In part III we turn to the general challenges of trust building in various environments. What problems do people face in their efforts to create or foster trustworthiness and thus seed trust in social relations? What of betrayal? How do we understand and cope with the many paradoxes of trust? We begin with the chapters that discuss the role of power and hierarchy in trust relations.

## **Part I: Trust and Hierarchy**

Despite the proliferation of new organizational forms, the hierarchical relation remains one of the most important and prevalent forms of organizing found in contemporary organizations. Such relationships are interesting arenas in which to examine trust-related judgments and behaviors because by definition the parties in the relationship are unequal in power or status. Accordingly, the first set of papers examines aspects of trust within hierarchical relationships, an important class of relations in many social and political contexts.

Kurt Dirks and Daniel Skarlicki investigate the importance of trust in leader-follower relationships. In particular, they examine how trust in leaders contributes to effective functioning within groups and organizations. The authors note that although the importance of trust has long been recognized by leadership theorists, there is a dearth of systematic theory and research showing precisely how trust in leaders actually contributes to effective group functioning. In reviewing the extant literature pertaining to this topic, Dirks and Skarlicki note that much of the theory and research to date on trust in leaders is typically *relationship-based* or *character-based*.

Relationship-based trust perspectives emphasize the importance of how followers construe their relationship to a leader. The relationship or

the social bond between leaders and followers includes such things as salience of a shared identity, recognition of a common background, and a history of cooperative interaction. Character-based perspectives focus on leaders' characteristics, specifically, the impact of leaders' perceived characteristics on the level of trust followers have in them. According to this perspective, followers' trust in leaders is influenced by their perceptions of the leaders' fairness, trustworthiness, or competence.

Dirks and Skarlicki comment that these perspectives have different implications for thinking about the consequences of trust in leaders. From a relationship-based perspective, leaders and followers exist in an exchange relation, in which positive sentiments and the behaviors those sentiments foster can be reciprocated, resulting in the building of mutual trust over time. In the character-based perspective, followers act in accord with the leader's preferences to the extent that the characteristics they attribute to leaders justify such action. Thus, followers will be willing to engage in behaviors that put them at risk when they attribute competence and benevolence to their leaders.

Despite the fact that some research has tried to untangle these issues, many questions remain. In particular, we know relatively little about what types of trust an organization should focus its efforts on building. For example, are there more advantages to be gained from fostering trust in the top management team, thus creating an organizational culture of trust? Or should more effort be focused on proximal relations (supervisor-subordinate dyads) and let trust build from the ground up? Finally, the authors take on the thorny question of why it is that trust in leaders is so hard to build and sustain? Drawing on existing theory and evidence, they propose that it is important to consider the interdependent effects of leader variables that might influence followers' willingness to engage in some form of desired behavior (such as organizational citizenship behavior, or OCB). For example, it may be essential that a leader be perceived as high in integrity and benevolence in order for followers to engage in OCB.

To evaluate this possibility empirically, Dirks and Skarlicki sampled employees from the financial services department of a large bank. Consistent with their theoretical expectation, the authors found that the relationship between leader integrity and citizenship behavior was strong and positive only when leader and organizational benevolence was high. At moderate or low levels of benevolence, the relationship between leader integrity and such behavior was insignificant, indicating a close link between benevolence and perceived integrity.

John Brehm and Scott Gates examine the complex and problematic trust relations that exist among social work supervisors, the workers they supervise, and the latter's clients. They begin by noting that there is at root a fundamental problem of trust between social work supervisors

and the social workers under them. Although supervisors are legally responsible for the actions of those under their supervision, civil service regulations and other legal constraints limit the ability of these supervisors to exercise vigilant and thorough supervision. In a word, they are accountable, yet their hands are in some respects tied. This relationship is further complicated by the fact that the relationship between social workers and those they supervise is protected by confidentiality agreements. Thus, supervisors must work partly in the dark when it comes to knowing how well those they supervise are actually carrying out their duties in a credible and responsible way.

To analyze trust dynamics in this complex relationship, the authors take as an analytic point of departure David Kreps's (1990) well-known trust-honor game. Brehm and Gates introduce a variation of this game, which they call the *trust-honor and reform* game. The game assumes that supervisors confront the choice as to whether to monitor or not monitor their social workers. Social workers in turn can decide to pull or not pull a given case. Pulling a case refers to the discretionary power (based on threat) social workers have to make their clients ineligible for state support when a client does something that merits such action. Clients can respond by either reforming or not reforming their behavior. Issues of trust and distrust reverberate throughout this triadic relationship, as all parties strives to pursue their interests and protect themselves from the prospect of misplaced trust.

Using this basic model as a conceptual platform, Brehm and Gates then show how trust in this complex triadic relationship can be understood in terms of the updating of beliefs among the interdependent actors. In particular, using survey data from social workers and their supervisors, Brehm and Gates attempt to evaluate this model empirically (because of privacy policies, client data unfortunately were not available).

The results of their study suggest that trust between supervisors and their subordinates is contingent in no small measure upon the availability of opportunities for supervisors to interact with and train their subordinates. Second, trust entails support and support entails protection. In particular, they found that subordinates' sense of protection from outside interference was a strong predictor of trust in the supervisor. Thus, the ability of the supervisor to insulate subordinates from debilitating organizational politics and disruptive case interference were found to be very important to the trust-building process. Finally, and perhaps most important, the authors found that trust is consequential in the sense that social workers who trusted their supervisors behaved differently from those who did not. In particular, trust was shown to enhance perceived discretion on the part of the social workers, resulting in a greater expenditure of effort.

Another important form of hierarchical relationship in which trust has long been assumed to be important is the relationship between physi-

cians and their patients. This relationship has been undergoing rapid change in recent years, as medical care has been increasingly delivered in managed-care settings rather than fee-for-service practices. Yet little is currently known about how about such settings actually influence trust and distrust between doctors and those they treat. Accordingly, Karen Cook, Roderick Kramer, David Thom, Irena Štepanikova, Stefanie Bailey Mollborn, and Robin Cooper investigate some of the determinants of perceived trustworthiness in physician-patient relationships in managed-care settings. They conceptualize the relationship between physicians and patients as one of mutual or reciprocal trust, noting that although there has been a fair amount of research on the determinants of patient trust in physicians, much less work has examined the conditions under which physicians trust or distrust their patients, and the consequences of such trust or distrust.

Using qualitative data from interviews and focus groups involving patients and physicians, Cook et al. (2003) find that patients trust physicians who demonstrate care, empathy, and respect for them as individuals, not just as “cases.” Important indicators of these sentiments are sustained eye contact, open body language, active listening, as well as the opportunity for the patient to participate in decision making relevant to their treatment. The findings also indicate that certain demographic characteristics such as gender, age, and culture or country of origin matter in the trust relationship because they affect the nature of the interpersonal communication of information. Finally, these authors explore how the features of the managed care settings in which patients receive their health care affect the levels of trust in their physicians. Such factors include a perceived increase in time constraints by both parties, lack of continuity of care, increased diffusion of responsibility and the perceived conflict of interest that emerges when physicians are widely viewed as constrained to offer lower-cost health care options such as fewer diagnostic tests or limitations on access to specific drugs or therapies.

Chapters by Gary Miller and John Darley, respectively, examine the role trust plays in another hierarchical relationship: employer-employee relations. In particular they investigate the effect of trust on the voluntary effort of the workers. There has been a strong presumption in the trust literature that high-trust relationships are more productive than low-trust relationships. The authors explore this general thesis, although they approach the problem of voluntary effort from distinctly different perspectives, Miller draws on insights from rational-choice theory. Darley eschews rational choice to focus on social identity and more emotional determinants of trust.

In his chapter, Gary Miller addresses the important and enduring question of how best to align worker incentives and effort. Specifically,

a primary aim of his paper is to examine the nature of the “tit-for-tat” exchange between managers and workers. His focus is on relationships capable of yielding Pareto improvements (all parties are happier) in organizations in which work contracts are based on the monitoring of individual actions rather than outcomes.

In building his argument, Miller first demonstrates the Pareto-suboptimality of effort-based monitoring. Drawing on the work of Akerlof, he suggests that Pareto improvements can be achieved if both supervisors and those they supervise enter into an informal “gift exchange” relation, whereby the supervisors tacitly agree to exchange leniency (relaxed surveillance) for increased voluntary effort by the workers. Supervisors benefit in such an exchange, of course, because of the reduced costs of monitoring. In turn, workers benefit by avoiding having someone breathing down their necks. Employing the Folk Theorem, Miller argues that this tacit arrangement works because each actor can make a credible threat that effectively guarantees or enforces the other’s cooperation or compliance. The supervisor has the threat potential to increase monitoring should the worker’s effort fall below some desired threshold. The worker in turn has the credible threat of merely “working to rule” (providing only a minimal or suboptimal level of effort). Using this logic, Miller thus argues that what looks on the surface like a process of simple and straightforward gift exchange could actually be the rational response of supervisors and workers to the other’s implicit threats. Miller goes on to argue that exchange is one of the sustainable cooperative equilibria in a cooperative game in a repeated game situation.

Miller then examines the “control paradox,” whereby more control leads to less effort. He illustrates the implication of his argument in several important domains, including public agencies, congressional oversight, public versus private schools, and regulatory environments. In some ways this is the key issue faced by Brehm and Gates in their study of social workers and their supervisors. Miller concludes by explicating the role trust plays in getting to sustainable cooperation. He notes that there are multiple equilibria in such games, some of which are Pareto-superior. What might help to move to Pareto-improving equilibria? Here trust enters the equation in “switching” equilibria.

An impressive feature of Miller’s analysis is how far it carries us without our having to invoke notions of intrinsic motivation. Instead, we can assume simply that the interdependent parties to this exchange relation possess common knowledge about the equilibria being played. Because players are rational, they adjust their behavior to achieve what they desire and to avert the outcome they least desire. In this case, the worker supplies high levels of effort so as to not set off the supervisor’s “trigger” for increased monitoring. Voluntary compliance by the worker in turn

averts an increase in close monitoring. Both are happier and better off (the meaning of a Pareto improvement).

The contribution by John Darley examines the important question of the role trust plays in organizational performance and success. In order to get the work of the organization done, Darley notes, employees must be willing to comply with organizational rules and directives. They must perform their role. However, success depends on more than this. It also depends on voluntary efforts by workers, including willingness to undertake actions that fall outside their job description or formal organizational role. Motivating employees to perform these “extra-role behaviors” is, of course, at the heart of much organizational theory and research. One class of such behaviors are the organizational citizenship behaviors (OCB) investigated by Dirks and Skarlicki.

Darley reviews existing theory and research relevant to these broad issues and then moves on to address the specific issue of trust. What kind of trust makes a difference in such situations? He notes that some conceptions of trust highlight primarily its calculative dimensions. For example, I trust you simply because I recognize it is in your rational self-interest to fulfill that trust. I can trust you, moreover, because I keep my powder dry (and available) should retaliation for any detected breach of trust be required. In contrast, noncalculative modes of trust, he argues, have their roots in the social relations and ties that obtain between parties. These “noncalculative” conceptions have their theoretical roots primarily in social identity theory. As Darley points out, noncalculative trust helps us understand the conditions under which people feel that a breach or betrayal of trust has occurred. It brings to the surface the identity-based elements of trust and distrust.

## **Part II: Trust and Distrust in Teams and Networks**

In contrast to hierarchical dyadic relationships, in which the parties interact face-to-face and frequently in organizational settings, the chapters in part II examine trust and distrust between individuals in group or collective contexts. Here the parties may be more dispersed geographically or may interact much less frequently face-to-face. In some cases, the actors are anonymous, at least initially.

Chapter 7, by Helen Nissenbaum, focuses on interesting issues of trust on the Internet. Research on the psychological and social complexity of life on the Internet is still in its infancy (see Wallace 1999 and DiMaggio et al. 2001 for a review of the literature). In an important contribution to this growing literature, Nissenbaum articulates the potentially substantial benefits that might accrue should trust online be achieved. At the

same time, she notes some of the formidable difficulties associated with securing such trust.

Nissenbaum begins by suggesting that trust online could improve the quality of social life in many spheres, ranging from personal experiences and communal relationships to civic participation. She observes further that if trust can be secured, more people and institutions will accept and use online resources, leading to more investment in online capabilities, resulting in more use, and so on, in a virtuous cycle. Broadly construed, therefore, the Internet might be a significant source of social-capital building.

Although acknowledging these potential benefits of trust online, Nissenbaum then presents an intriguing analysis of some of the inherent features of online life that make such trust problematic. Among the potential obstacles to online trust she notes is the simple fact that for many users and in many parts of the world the Internet remains a relatively new and unfamiliar form of social interaction. Not surprisingly, therefore, there may be an inherent wariness in online transactions.

A second and in some respects more serious problem concerns what Nissenbaum terms *missing identities*. Much of the interaction on the Internet is inherently anonymous. This anonymity is a double-edged sword. On the one hand, it enables people to protect their identities, thereby achieving confidentiality. On the other hand, anonymity reduces greatly the opportunity for the interacting parties to access and verify identity-related information that specifically contributes to the trust-building process.

Another problematic feature of life online is what the author terms *missing personal characteristics*. Many of the differentiating cues that influence our judgments about trustworthiness, she notes, are simply missing from online transactions. For example, in the online environment we typically lack cues that would tell us something about the similarity or familiarity to us or the values we share with those with whom we are interacting online. The prospect of easy misrepresentation of these diagnostic cues—the ability to and ease with which one can lie about one's age, gender, or professional standing or training—is a serious problem, as users of some online advice services have learned.

Finally, the author concludes, the settings or contexts of information exchange on the Internet are themselves somewhat inscrutable. Online users don't know for certain the source of the messages they receive—it is relatively easy to disguise the origins of a message or inflate the credibility of an institution. As with other attributes of life online, however, inscrutability is a double-edged sword. On the one hand, it is equalizing and can liberate people from status-driven patterns of social interaction (dominance, deferral, etc.). In this respect, there is a refreshing sense of egalitarianism about chat rooms. On the other hand, inscrutability means

less ready access to crucial information about social and professional roles. An example is provided by help sites where it is hard to evaluate the expertise of the experts volunteering to offer advice or support.

Given the absence of the usual grounds for trust, Nissenbaum next considers, what foundations or bases are available for securing trust online? What steps can we take to sustain trust? The answer lies in enhancing the security of transactions on the Internet. One contributing solution is greater sophistication about access control. If we are assured that the collective is populated by responsible parties, interaction is easier and more spontaneous. A second approach is to increase the transparency of identities. If we can make users' identities transparent enough to interacting parties, they can protect themselves against betrayal and harm. Surveillance is a third weapon.

Nissenbaum concludes by posing the question, Can trust really be secured online? Is security a foundation for online trust? Nissenbaum points out that security experts are engaged in an almost Sisyphean battle in trying to close off every avenue of attack. Much like any fixed defense, the conventional sorts of firewalls and structures currently used to increase security online are inherently vulnerable to circumvention or penetration. They are breachable. As she reminds us, much of our attention has focused on the danger without, whereas far less systematic attention has been paid to the threat posed by insiders. She reminds us that the centurions at the gate may themselves be corrupt or corruptible, prone to wicked mischievousness. Pervasive distrust may be hard to avoid in such contexts.

In "Architects of Trust" Bill McEvily and Akbar Zaheer investigate network facilitators of trust and cooperation among organizational competitors in geographical clusters. Here we move to organizations or their agents as the key actors. A provocative feature of geographical clusters, they note, is that local firms often engage in acts of collaboration with each other, while at the same time competing against each other in downstream product markets. Organizational theorists have generally attributed this phenomenon to the existence of dense and overlapping networks of social, professional, and exchange ties among firms within such clusters. Repeated interactions among cluster members have produced benefits in terms of trust building, including the development of shared understandings, collective identities, and mutual confidence.

Although such an argument may seem compelling on prima facie grounds, McEvily and Zaheer point out that from an empirical perspective, many questions remain unanswered as to how such shared understandings, collective identities, and mutual confidence actually develop. Accordingly, using a qualitative approach, they investigate the development of inter-firm trust in the office furniture manufacturing industry in western Michigan. Among other things, they found that

regional institutions played a particularly important role as facilitators of trust among network members. Although unexpected, the authors were able to unpack this relationship, discovering that regional institutions were able to facilitate interaction by organizing conferences, meetings, and setting up subcommittees that brought manufacturers and suppliers together. This study thus provides important evidence for the argument that third parties can act as important conduits of trust among parties.

Roxanne Zolin and Pamela Hinds in chapter 9 explore some of the difficulties in creating and sustaining interpersonal trust in geographically distributed cross-functional work teams. Trust theorists have long argued the benefits of face-to-face interaction and direct experience with others in the trust-building process. As more and more work in organizations is done by individuals who are dispersed geographically, however, it is clear that opportunities for such critical trust-building opportunities will be reduced, and so it becomes important to understand the extent to which trust can and does develop in the absence of such experience. Zolin and Hinds investigate these issues by comparing trust development in geographically collocated versus distributed dyads of student construction-design teams. Using a longitudinal methodology, they sampled trust levels at several points in time, which enabled them to assess changes in trust levels. The authors found that trust levels in distributed dyads changed significantly less than trust levels in collocated teams. Although trust increased more in collocated teams, it also decreased more, consistent with prior research suggesting that trust levels are updated on the basis of direct experience.

Zolin and Hinds also examine how risk perceptions affect trust in geographically dispersed teams. They initially hypothesized that risk would be a significant factor in predicting trust levels for dispersed dyads, so that more risk meant less trust. Their results supported this hypothesis. Viewed in the aggregate, their findings clearly suggest that the dynamics of trust development and maintenance are different in dispersed and collocated teams. However, clearly more research is needed to document how to overcome some of the liabilities of distance for the trust-building process. Zolin and Hinds's results suggest the possibility that if distributed teams could meet initially for a brief period that led to positively structured interaction, then the team members might enjoy a relatively stable sort of trust predicated on the positive group stereotype.

In chapter 10, "Psychological Safety, Trust, and Learning in Organizations," Amy Edmondson defines psychological safety in terms of individuals' perceptions of the degree of interpersonal threat in their group or organizational environments. Thus, psychological safety encompasses people's sense of security when they take risks, such as when they make personal self-disclosures, propose new ideas, or report mistakes. With

the presence of psychological safety, there is a kind of tacit mental calculation or assessment as to whether others in a group or organization will give one the benefit of the doubt. Psychological safety thus constitutes a sort of trust in others' willingness to act benevolently toward the self as the self engages in risky behavior of some sort.

Using this definition as a starting point, Edmondson then develops the notion of *team psychological safety*, which can be conceptualized as a form of collective-level psychological safety. The author reviews recent empirical evidence on the impact of psychological safety on group functioning, using data from a variety of real groups ranging from cardiac surgery teams to teams operating in a manufacturing company. On the basis of her findings she identifies a number of antecedents of psychological safety, including important leader behaviors that facilitate the emergence and maintenance of psychological safety within teams. In particular, she suggests there are three leader behaviors that specifically promote psychological safety: accessibility, explicitly inviting input and feedback from others, and modeling openness and fallibility. Many of these same kinds of behaviors, it should be noted, have been implicated in the literature on trust in leaders, suggesting a link between trust and psychological safety. Edmondson also identifies a number of informal, emergent group dynamics that support psychological safety.

Having identified some of the antecedents of team psychological safety, Edmondson then considers some of its consequences, especially with regard to how group members are likely to interact with each other. These include the willingness to seek help, seek feedback, be vocal about errors or concerns, and engage in innovative behavior. To the extent such behaviors facilitate group learning, a strong case can be made for the role psychological safety plays in learning organizations. Future research should explore the nature of the relationship between trust and perceived psychological safety in such organizational contexts. If trust provides psychological safety, trust building may be an essential goal of team training, for example of surgical teams or fire-fighting teams.

### **Part III: Challenges to Securing and Sustaining Trust**

Given recent examples of breakdowns in trust and trustworthiness (for example, Enron, Tyco, and Arthur Andersen), it seemed appropriate that we conclude the volume with three chapters that address some of the problems of securing and sustaining trust in organizations. In "Managing Images of Trustworthiness in Organizations," Kimberly Elsbach examines this important but largely neglected topic. In particular, she explores

the impression-management strategies and self-presentational tactics managers can use to enhance impressions of their trustworthiness. In some respects, she notes, it is hardly surprising that managers accord so much importance to being perceived as trustworthy. In her review of the empirical literature, she documents a number of positive effects of perceived trustworthiness, including enhanced support for and cooperation with trusted managers. Similarly, individuals are more likely to defer to authorities and accept their views as valid when they feel a high level of trust in those authorities (Tyler 1990).

Having established the importance of perceived trustworthiness, Elsbach then elaborates on some of the tactics that can be used by managers to enhance images of their trustworthiness. These include revealing or displaying similarities between oneself and those with whom one works; and revealing dissimilarities between oneself and known untrustworthy others. Impression managers trying to persuade others of their trustworthiness can also display cues or present evidence that they are members of a group that has a reputation for competence, benevolence, or integrity. This allows one to benefit from association with the group. Elsbach notes that impression managers use a variety of physical cues, nonverbal behavior, and language to manipulate impressions of their trustworthiness when interacting with their various constituencies.

Elsbach begins by noting that to possess an image of interpersonal trustworthiness, managers must be perceived by those around them as having a number of attributes, including competence, benevolence, and integrity. But how do managers achieve such attributions? Several antecedent conditions support perceptions of trustworthiness. These include behavioral factors such as behavioral consistency, accurate and open communication, and demonstrating concern. There are also several cognitive factors associated with attributions of trustworthiness, including social categorization and perceived similarity.

Elsbach's analysis reminds us that skillful impression managers may be highly effective at cultivating the illusion of trustworthiness, even when they have little concern with actually being trustworthy. Skillful deceptors may manipulate impressions of trustworthiness in order to commit and conceal wrongdoing and may actually abuse that trust to take advantage of overly trusting or naïve others. The Enron scandal is a good example of how this abuse works in the real world, leading the *Wall Street Journal* to publish a news story with the headline "En-Ruse?" According to the article, Enron executives asked more than seventy low-level employees to go to an empty trading floor and pose as busy sales representatives to impress a group of Wall Street analysts who were visiting the company's headquarters. According to one employee who participated in the lie, "We actually brought in computers and phones, and they told us to act like we were typing or talking on the phone when

the analysts were walking through. They told us it was very important for us to make a good impression, and, if the analysts saw that the operation was disorganized, they wouldn't give the company a good rating." To enhance this illusion, the employees even brought in personal pictures to adorn the tops of their desks. The ruse worked. Although the whole charade lasted only ten minutes, it was enough time to create the impression of a dynamic, flourishing trading floor.

Elsbach notes more positively, however, that it is also vitally important for trustworthy and well-intentioned managers to be able to effectively communicate and demonstrate their trustworthiness.

J. Keith Murnighan, Deepak Malhotra, and Mark Weber in "Paradoxes of Trust" examine trust issues in the context of exchange relationships and contracts. In particular, they note that, from an attributional perspective, contracts act as double-edged swords. On the one hand, contracts provide an assurance mechanism, allowing cooperative behavior to be initiated easily. Legal scholars and organizational theorists alike have long noted the power of the "shadow of the law" as a basis for risk taking and presumptive trust in exchange relationships. In a sense the contract allows confidence-building behaviors to occur, giving trust a toehold. At the same time, contracts in certain respects make trust harder to develop because they create ambiguity in the attribution of motive. Because the existence of the contract is common knowledge to the contractees, the others' cooperative or trustworthy behavior can be attributed to the existence of the contract, a situational factor, and not to innate trustworthiness. Therefore, the behavior is less clearly diagnostic of the actor's underlying motivation, intentions, or character. Contracts thus generate situational attributions for others' cooperative behavior, which potentially impede or slow the development of interpersonal trust. Further, trust operates differently in the contract context from what the perspective of rational models of judgment and choice would lead us to expect.

Murnighan, Malhotra, and Weber offer experimental data to support their attributional impetus model. Central to this model, they posit, is the assumption that people sometimes are willing to assume large risks in an effort to develop trust (rather than the sort of gradual incremental risks assumed in many GRIT-type models of trust development) because they recognize that the targets of their trust-building efforts often are likely to judge partial or tentative trusters harshly. One of the important points they make is that people often take risks in their social exchanges in order to build trust (Cook et al. 2003).

In chapter 13, "Untangling the Knot of Trust and Betrayal," Sandra Robinson, Kurt Dirks, and Hakan Ozcelik examine how individuals' initial trust levels influence individuals' responses to their experience with a breach of trust or betrayal. Extant research makes it unclear

whether prior trust intensifies the experience of betrayal—a “the higher they are, the harder they fall” sort of argument—or acts to buffer against feelings of betrayal or mitigate the effects. The literature on betrayal suggests to the authors that individual reactions to betrayal are at the very least complex and varied.

They propose that trust might play a role in mitigating responses to such betrayal. Trust, they suggest, might function as an appraisal heuristic. They animate their analysis in terms of two dominant but conflicting logical possibilities. The first is a “love is blind” hypothesis, according to which prior trust would be expected to soften or attenuate the impact of betrayal. This hypothesis follows a kind of “idiosyncrasy credit” logic, whereby we give people some benefit of the doubt on the basis of an accumulated history of otherwise benign and positive experiences. There are several reasons why this might be the case, including cognitive consistency and the possibility that prior trust might influence the construal of actions or inactions linked to betrayal.

Robinson and his colleagues reconcile these opposing predictions and the psychological arguments in favor of them by proposing that prior trust moderates the relationship between the breach experience and the appraisal or interpretation of the breach, resulting in varied emotional and behavioral reactions.

The final chapter, by Roderick Kramer and Dana Gavrieli, looks at the presidency of Lyndon Baines Johnson and suspicion inside the Oval Office to return to the theme of part I and examine the complex relationship between power and trust. Kramer and Gavrieli argue that power affects social-information processing and social interaction in a number of ways that exert deleterious effects on the trust-building and maintenance process. They contend that it becomes harder for those in power to assess others’ true loyalty and trustworthiness. Power creates a fog or shroud of attributional ambiguity in interpersonal relationships between the powerful and those over whom they exert power.

To develop this general argument, the authors use published accounts and recently declassified data from secretly recorded White House conversations involving President Lyndon Baines Johnson and his advisers. These conversations, along with supportive archival evidence, provide a rich window into a leader’s struggles with problems of trust and distrust in his administration. Using this evidence, Kramer and Gavrieli identify a number of social-information processing biases that can lead to the breakdown of trust within power relationships. One of these is the *sinister attribution bias*, the tendency of social perceivers to overattribute malevolent motives to others, especially those in powerful positions. Another social-information processing bias is the exaggerated perception of conspiracy, the tendency to read coherent connections into individual social actions. The authors elaborate on some of the difficulties

that individuals in various positions of power encounter when trying to assess the veridicality of their own perceptions—in other words, power relations complicate the trust-building process. The pervasiveness of power relations in organizational settings makes this obstacle to trust building important to investigate.

## Conclusion

Taken together, the chapters in this volume highlight some of the complexities and subtleties of trust-related phenomena in contemporary organizational contexts. They amply demonstrate that trust takes different forms across many different settings, and that trust-building and trust-maintenance processes vary across them as well. These studies highlight the necessity and value of context-specific, “middle range” theories of trust that can be empirically tested.

A major aim of our volume was to highlight the complexities and problematics of trust and distrust in organizational contexts. Our survey does not leave one entirely sanguine about the prospects for trust.

This book had its origins in a rather pleasant academic conference. The setting for that conference was relaxing and convivial. Our conversations about trust had all of the pleasing ambience and casualness of a scientific conference in which abstract ideas were being tossed around with enthusiasm and a certain carefreeness. Recent world events have given the topic of trust more significance and even urgency. There is, we hope, a useful timeliness to our collective musings about trust and distrust.

## References

- Cook, Karen S., ed. 2001. *Trust in Society*. New York: Russell Sage Foundation.
- Cook, Karen S., Toshio Yamagishi, Coye Cheshire, Robin Cooper, Masafumi Matsuda, and Rie Mashima. 2003. “Trust Building via Risk Taking: A Cross-Societal Experiment.” Working paper.
- DiMaggio, Paul, Eszter Hargittai, W. Russell Neuman, John P. Robinson. 2001. “Social Implications of the Internet.” *Annual Review of Sociology* 27(3): 307–36.
- Fukuyama, Francis. 1995. *Trust: The Social Virtues and the Creation of Prosperity*. New York: Simon & Schuster/Free Press Paperbacks.
- Hardin, Russell, ed. 2002. *Trust and Trustworthiness*. New York: Russell Sage Foundation.
- Kramer, Roderick M., and Tom R. Tyler. 1996. *Trust in Organizations: Frontiers of Theory and Research*. Thousand Oaks, Calif.: Sage.
- Kreps, David M. 1990. “Corporate Culture and Economic Theory.” In *Perspectives on Positive Political Economy*, edited by James E. Alt and Kenneth A. Schepsle. Cambridge: Cambridge University Press.
- Lane, Christel, and Reinhard Bachman. 1998. *Trust Within and Between Organizations: Conceptual Issues and Empirical Applications*. New York: Oxford University Press.

- Putnam, Robert D. 1993. *Making Democracy Work*. Princeton, N.J.: Princeton University Press.
- . 2000. *Bowling Alone: The Collapse and Revival of American Community*. New York: Touchstone.
- Sitkin, Sim B., Denise M. Rousseau, Ronald Stuart Burt, and Colin Camerer. 1998. "Special Topic Forum on Trust in and Between Organizations." *Academy of Management Review* 23 (special issue).
- Sztompka, Piotr. 1999. *Trust: A Sociological Theory*. New York: Cambridge University Press.
- Tyler, Tom R. 1990. *Why People Obey the Law*. New Haven, Conn.: Yale University Press.
- Wallace, Patricia. 1999. *The Psychology of the Internet*. New York: Cambridge University Press.