The Crisis of Trust and Trustworthiness in Organizations

Marc A. Cohen, Ph.D.
Faculty Fellow, Center for Business Ethics
Professor of Management and Philosophy
Seattle University
June 2, 2021
Executive Summary

We’re surrounded by news of a crisis of trust within and between organizations. The benefits of trust are widely reported in academic work and also in the more popular business press: trust is said to increase efficiency and productivity, improve communication and creativity, increase employee satisfaction and commitment, and encourage teamwork. But there is still a crisis—why? There are two reasons. First, the pressure to perform creates fear among managers and organizations, and, in response, both attempt to control others to reduce risk, instead of trusting one another. Second, managers and organizations hold negative assumptions about their employees being unwilling to work without close oversight and supervision.

In response to these problems, this paper outlines empirical research intended to give managers and organizations confidence in building a culture of trust.

- The experimental research shows that managers and organizations can create a culture of trust by trusting others. As one sociologist puts it, summarizing a body of research, there is a “positive effect (trustworthy behavior) from supervisors treating employees as if they are reliable even when there is relatively little information about their actual character or competence.”

- Surveillance systems and potential sanctions (to provide material disincentive) are widely used in the organizational context to control employees. But these systems can backfire, employees often react negatively to being distrusted; employees respond to such systems by reducing work effort.

- There are positive effects of trust, both economically and psychologically. Within a culture of trust managers and employees are not guarding against opportunistic behavior.
The Crisis of Trust in Organizations

We’re surrounded by news of a crisis of trust within and between organizations. What should we think, and what should we do?

One starting point about this crisis comes from the title of a recent news article: “If Workers Slack Off, the Wristband Will Know. (And Amazon has a Patent For It)” (Yeginsu, 2018). And, if anything, the use of workplace monitoring has grown dramatically during the pandemic (Golden & Chemi, 2020), because companies want to make sure that remote workers are actually working. One company provides software that will randomly photograph employees’ computer screens one, two, or three times every ten minutes (see also Blackman, 2020). To be sure, there are exceptions, some managers and some organizations have built cultures of trust. But there’s a vast literature telling managers and organizations to foster trust; the books wouldn’t sell and the message wouldn’t count as news if there wasn’t a crisis.

The benefits of trust are widely reported: one paper in *Harvard Business Review* identifies four benefits in organizations: increased efficiency and productivity, improved communication and creativity, increased satisfaction and commitment, and teamwork is encouraged (Cardona & Wilkerson, 2009).

One widely discussed example from management describes these positive impacts in a particular work setting: Between 1962 to 1982 General Motors produced cars in a factory in Freemont, California, that was widely considered to be the worst of GM’s plants. “The work force in those days had a horrible reputation, frequently going out on strike (sometimes wildcat strikes), filing grievance after grievance and even sabotaging quality. Absenteeism routinely ran over 20%” (Shook, 2010). Workers drank on the job and had sex in the plant; one employee described bringing a thermos of screwdrivers (the alcoholic drink, not the tools) to work (NPR, 2010). “There were cars with engines put in backwards, cars without steering wheels or brakes” (NPR, 2010). After the plant closed in 1982, Toyota agreed to reopen and operate the plant as part of a joint venture with GM, the New United Motor Manufacturing, Inc., known as NUMMI — to produce the Toyota Corolla and also some GM cars. Honda and Nissan were already manufacturing cars in the United States and, by entering the joint venture, Toyota hoped to quickly learn how to adapt to the U.S. operating environment. General Motors hoped to learn about the Japanese manufacturing techniques and apply lessons in its other factories (NPR, 2010).

This example is used in business school curricula as an example of changing business culture, but the change in the business culture in the Freemont factory was the result. Changing the production process, the way people worked, was the mechanism: Toyota gave assembly-line workers the means to produce high-quality cars by giving them the autonomy to fix mistakes when they occur. The workers responded positively to this trust placed in them. Within one year the Freemont facility was GM’s best factory in terms of quality with the same workers.

But managers and organizations nevertheless struggle to trust. Why? The pressure to perform creates fear among managers and organizations, and in response both attempt to control others to reduce risk. A recent (different) article from *Harvard Business Review* puts it clearly: “Pressure to reach performance targets and control costs sometimes leads managers to do things that
unintentionally signal a lack of trust. When these pressures are great, many managers become focused on their own job security and respond by constricting control. This can lead to the type of thinking that focuses on only securing bottom-line outcomes, which often come at the expense of other priorities, like developing relationships and empowering employees to make independent decisions (Brower et al., 2017). This passage captures the motivations, and it’s also interesting because of the dishonesty it (perhaps accidentally) suggests. Managers (and organizations) deliberately constrict control because of fear, and managers might prefer that employees don’t notice, they might prefer that employees miss the signal—the signal is unintentional” in that sense—but that signal is accurate.

Notice that the explanation here—managers and organizations constrict control because of fear—has nothing to do with the particular employees; the manager and/or the organization are motivated by fear, and he or she (or it, the organization) thinks that controlling the work and controlling employees will deliver better results.

Managers and organizations also struggle to trust (or just refuse to trust) because they make negative assumptions about their employees. Consider these passages from one of the earliest papers on workplace monitoring: “Economic theory, in particular principal-agency theory, assumes that in work relations individuals pursue their own interest and expend work effort to the point where net utility is maximized…. Agents [employees] relentlessly exploit every opportunity to east their work burden, as long as the principles [managers] do not react and punish them so severely that their net utility from shirking is decreased” (Frey, 1993). Given this expectation, we put wristbands on our works or monitor their computer screens.

But, pushing back against this trend, managers and organizations need to understand three things about trust.

**Trust Works**

Trust is widely studied using an experiment involving players sending money back and forth: player one is given money; he or she can send some to player two (or keep all of it); the experimenter triples any amount sent; and then player two can return some money to player one (or not). Both players know the rules and the starting point, how much player one is given, but the players don’t interact or meet.

The experiment makes it possible to measure trust quantitatively: if player one sends all of his/her funds, then player one completely trusts player two. But if player one sends a small percentage, then player one’s level of trust is low. And the experiment is designed to capture three core intuitions. First, cooperation can make both players better off. If player one sends all of his/her money, the total amount to be shared is tripled. But—this is the second intuition—sending money makes player one vulnerable because player two can keep the money. So, third, cooperation depends on player one trusting player two and sending money, and it depends on player two acting in a trustworthy way, returning funds.

My recent paper—written with Seattle University colleague Mathew Isaac—studied outcomes in this experiment using different conditions (Cohen & Isaac, 2021). In the first condition, player one
started with $20 and trusted player two, so player one sent all $20. Player two received $60. Our subjects participated as player two, we asked them how much they would return. In a separate condition, player one started with $2000 but didn’t trust player two, so player one only sent 1%, $20. Again, player two (different subjects in the role of player two) received $60, and we asked the subjects how much they would return.

Subjects received the same amount in both conditions so we can compare their decisions. Subjects returned more in the trusted condition, on average about $27, compared to an average of about $15 in the distrusted condition. Notice that player one benefited by trusting: after the transfers player one has $27 instead of the original $20; this is the first intuition mentioned earlier. But player one was vulnerable: a few subjects kept the entire $60, leaving player one with $0. We can conclude that being trusted/distrusted affected the subjects’ decisions. (The statistical analysis in the paper is more complicated because it involves a control condition. We show that player twos’ decisions in the trusted condition differed significantly from the control that did not involve trust or distrust, and player twos’ decisions in the distrusted condition also differed from the control—so there is a trust effect and a distinct distrust one. We’d be happy to send the paper to any readers interested in parsing the statistical details!)

There are two important points for managers and for business decision-making in this experiment.

First, when we think about trusting another person or another company, the decision often turns on that other person’s or that other company’s likely behavior, or on that other person’s/other company’s character. Is the job applicant trustworthy? Is the potential supplier one that has performed well in the past? But the finding outlined here, that player two responds differently in the two conditions, shows that trustworthiness is not a stable trait or fixed disposition. Instead, player two’s decision, how much to return to player one, depended on the way player one behaved. If player one trusted player two and sent all of his/her funds, then player two acted in a trustworthy way and returned a significant amount. Player two acted in a cooperative way. Academics sometimes refer to the finding here—if player one trusted player two, then player two acted in a trustworthy way—using the slogan, trust begets trustworthiness. But if player one didn’t trust player two and sent only a nominal amount, then player two acted in a way that punished player one—sending back little or nothing. Indeed, 40% of subjects in the distrusted condition sent nothing back! This sort of response on player two’s part is called “negative reciprocity” in the academic literature.

How, then, does a manager create a culture of trust? By trusting others. As mentioned above, one reason we might hesitate to trust, despite the benefits for the persons involved and for organizations in business contexts, is that we are afraid to make ourselves vulnerable. But the findings just outlined should give us confidence: trusting others affects their thinking, it affects their actions in a positive way, so the vulnerability is much less than we think.

**Distrust Doesn’t Work**

This is the second—and related—lesson from the experiment. As noted, subjects in our experiment returned less when they were distrusted, 40% of subjects returned nothing. And the distrust effect was stronger: subjects responded more strongly (in the negative direction) to being distrusted...
compared to the positive effect of being trusted. To understand this, we asked our subjects to explain their decisions, why did they return little or nothing? They said they were responding directly to being distrusted, they felt insulted, and also to the unfairness involved. And it’s important to note here that we didn’t label the conditions, the subjects saw that they were distrusted—because player one sent only $20, 1% of his/her funds—and reacted.

This finding is consistent with literature on surveillance and control in the workplace. For example, the paper quoted above—explaining the negative attitude about employees in economic theory—is by economist Bruno Frey. He found that monitoring workers indicates distrust and backfires. Employees who feel that their own self-determination is “unduly” restricted, and/or when managers’ attempts at control signal a negative judgment about the employee, employees responded by reducing work effort. The pragmatic question here is what constitutes “undue” restriction, and the most direct answer is that control creates undue restriction when it is seen as expressing disrespect (Frey, 1993). One subsequent study of call center employees found further evidence for seeing respect at work: employees did not respond negatively to monitoring when they perceived the monitoring to be fair or, put the other way, employees would look for opportunities to exploit management when the felt that management was “unfair” and “uncaring” (Nagin et al., 2002).

Moreover, surveillance and controlling systems have other negative effects important to note. Another an early paper in this space, by organizational psychologist Lloyd Strickland (1958), measured supervisors’ attitudes toward subordinates; one subordinate was monitored, the other not, and at the end of the experiment the supervisor trusted the unmonitored subordinate more even though their outcomes had been the same. Strickland shows that the presence of a control mechanism affected subjects’ judgments of others: the supervisor-subjects saw the employee-subjects who were monitored as being compliant and supervisors saw the others, who were not monitored, as being trustworthy. More recently, two other organizational psychologists reported a parallel finding: after agreeing to a contract, observers explained behavior in terms of compliance with reference to the contract (Malhotra & Murnighan, 2002). The presence of control mechanisms again affected subjects’ interpretations about the behavior of others.

The effect here depends on the signals involved. Surveillance systems and potential sanctions (to provide material disincentive) are used—in the organizational context—because management assumes that persons are not trustworthy. These systems can therefore provide a signal that other persons cannot be trusted so reduce trust.

And, two management professors used experiments to show that the presence of control systems can affect the way we interpret behavior and close off space for considering ethics and trust: “The presence of a sanction system will lead to judgments that the decision is more business in nature, whereas the absence of such a decision will lead to the perception that the decision is relatively more ethical” (Tenbrunsel & Messick, 1999). A decision that is “more business in nature” involves calculation of costs and benefits for different actions, and the outcome will depend on the nature/scope of the penalties involved for non-cooperative behavior; in contrast to the relying on norms and reciprocity, and trust, in “more ethical” decisions.

So, in short, the manager who defaults to surveillance and control—the manager who “constricts
control”—will likely find that employees respond negatively. And surveillance and control systems can have other negative effects on organizations.

**Control Still Plays a Role**

What do we do in practice, how is a manager to combine trust and control? Some academics think that trust and control substitute for one another. Within organizations, managers who trust their employees don’t need to actively monitor their performance, whereas managers who do not trust their employees tend to rely on surveillance and control systems. Again, “If Workers Slack Off, the Wristband Will Know.” And joint venture partners that trust one another do not need contracts with legally binding clauses; trust replaces the contract and enables work to proceed more quickly, and also with less expense. But partners who do not trust need contracts.

Others think that trust and control are compatible. The point is often put in abstract terms. When A and B work together, A might worry that B won’t act in a trustworthy way. So A will demand a contract that imposes penalties on B for not fulfilling B’s obligation; the idea is that B will act in a trustworthy way to avoid the penalty (in economic terms, the cost changes the “payoff” to B for non-cooperation). And then, knowing this, A will be more likely to trust B. So control mechanisms are thought to facilitate trust. Less abstractly, on this line of thought, a contract between two companies might be required to make both parties feel safe enough to enter a joint venture at the beginning of a relationship, one in which trust could develop gradually, over time.

But even if trust and control are compatible, control has negative effects. This puts managers and organizations in a difficult position, surely trust is a matter of degree, we need both trust and control.

One important example comes from work by Rosalind Klein Woolthius and colleagues (2005), long-time trust researchers, who studied a joint venture between two large pharmaceutical companies. The individual companies “already had a long and trusting relationship”; both companies “explicitly mentioned their trust to be a factor of utmost importance in their decision to collaborate”; but they nevertheless drew up an “extensive contract” that specified obligations in “great detail,” and they also formed a joint venture to formalize control over the relationship.

Woolthius and her colleagues note, the existing trust relationship should have reduced the need for contractual safeguards, trust should have substituted for the contract. But this was not the case. Why? Woolthius and her colleagues suggest that the contract played a different kind of role, not connected with preventing opportunistic among the partners; instead, the contract “was a sign of the commitment… and a tool for coordination, simply to avoid misunderstandings in the project’s management (comparable to minutes of a meeting).” So here the contract is best understood as making the terms of the partnership explicit. It serves as a sign of commitment rather than as a signal of distrust. And using a contract in this way is compatible with building a culture of trust because the contract is not being used to control—but to clarify a joint understanding.

Some might see a communication issue here: control mechanisms must be presented in the right way to avoid the negative signal. But the negative signal is not something that can be hidden. Employees easily see through the rhetoric of management. Instead, managers should be able to explain to employees why they need certain monitoring systems; joint venture partners should be
able to say what positive role a contract plays—explaining why those don’t reflect distrust—and then rely on their partners to see the business logic. Doing so is to treat partners and employees with respect.

This requirement connects to the example above from Bruno Frey: employees react negatively when the find surveillance and control as “unduly” restrictive, where undue means disrespectful.

**Explaining the Positive Effects**

On the standard economic account, when transaction partners trust one another, they can avoid time-consuming and expensive processes involved in negotiating contracts. So trust reduces transaction costs. One academic paper puts the point this way: “From a transaction cost perspective, the most compelling argument for the superior efficiency of interorganizational relationships that involve trust is simply that trust reduces the inclination to guard against opportunistic behavior (i.e., deliberate misrepresentation) on the part of the exchange partners” (Zaheer et al., 1998). The same line of thought applies within organizations: when an organization or just particular managers trust employees, they—the organization, the managers—can dispense with control systems for monitoring workers, saving time and money.

But there is also a psychological dimension. From the same paper: “Partners in relational exchange that have forged a high level of interorganizational trust are more likely to give each other the benefit of the doubt and greater leeway in mutual dealings. Such leeway will tend to reduce the scope, intensity, and frequency of dysfunctional conflict.”

The key word in that passage is “dysfunctional.” Interpersonal trust makes persons feel safer sharing ideas, participating in team assignments, and disagreeing, and these behaviors can lead to more conflict. But in the context of a trust relationship, persons feel safe (they are not guarding against opportunistic behavior), that allows conflicts to surface—and to get resolved. So we might see more conflict, but it the effect is not negative. Moreover, in the context of trust relationships persons can accept disagreement and at the same time maintain positive work relationships, again because they are not expecting others to act in ways that harm, persons in trust relationships are not expecting others to take from them, undermine them, blame them, etc.

That is the perspective from the management discipline, focused on instrumental benefits. My academic work emphasizes another kind of explanation, another perspective.

As background, there’s dispute among academics about how to best understand trust. Some see trust relationships as bets, literally, so when A trusts B to do X, A bets on B to do X because the potential benefits outweigh the possible risks. My own work has pushed back against this way of thinking and emphasized the moral dimension of trust. On my view, trust relationships are given structure by commitments. So, when A trusts B to do X, A *invites* B to acknowledge and accept an obligation to do X, and when—or if—B accepts the invitation, B takes on that obligation. A then relies on B’s commitment. An intuitive example: I ask my neighbor to pick up my mail while I’m on vacation (the invitation), he says “yes” (the commitment), and I trust him to do so (relying on that commitment). The whole process could be explicit, B makes a promise to A and A relies on that promise; my neighbor says “yes.” Or the obligations could be implicit; maybe in some
settings co-workers have a set of unstated obligations to one another. And note that the invitation could work in the opposite direction: my neighbor might offer to pick up my mail, inviting me to trust him to do so.

Further details aside, when A makes a trust invitation to B, A treats B as a partner. A’s invitation includes B in the moral community, defined as the set of persons who rely on one another’s commitments. Treating someone in that way is inherently, intrinsically motivating. One way to see this is to look at the opposite: when A will not trust B, A tells B that B lacks the ability and/or the integrity required to be a partner in some way. A insults B. That is why distrust invites negative reactions.

**The Limits of Trust**

But we shouldn’t be pollyannaish. Several years after the 2008-2009 financial crisis, I was asked to write a book review for a volume of essays—all by academics—on restoring trust in the financial services industry (Cohen, 2015).

The book’s fourteen essays provided an overview of the crisis, outlined a set of theoretical questions, and described a path forward. The central theme was this: relying on regulation to punish bankers and financial services firms can only foster (what the authors called) weak trustworthiness. And regulation can always be gamed, so this strategy isn’t “stable” enough. Instead, we need to foster strongly trustworthy behavior on the part of bankers and firms, which requires that bankers and firms acknowledge obligations and responsibilities toward others, and then act in ways motivated by those obligations. As one essay put it: the problem is that customers expect banking professionals to act in ways consistent with their, the customer’s, interests, but the bankers and the firms see transactions as governed by a different norm, buyer beware. To restore trust, we have to reframe the transaction so that the banker sees that strong trustworthiness—acting in ways consistent with the customer’s interests—is demanded.

My review was critical. I asked, “[I]f if we can’t weakly trust finance professionals to abide by regulation, and if we can’t trust finance professionals to honestly disclose risks, how [can the academics] convince those same finance professionals to re-frame their work and acknowledge an obligation to their customers?” They can’t. The bankers and the firms don’t see a problem. That’s why we get platitudes when CEOs testify in front of congress, but we don’t see financial services firms changing their behavior. And we don’t see social sanctions, disgraced bankers being shunned by others.

There was another banking scandal in 2012, just as the global financial crisis was receding. London investment banks had been systematically lying about the index used to price loans, the London Inter-Bank Offered Rate, LIBOR. The lying affected the interest rates applied to loans to corporations and consumers around the world. The banks involved were able to buy securities, manipulate LIBOR to increase the value of those securities, and then sell them at profit. Financial reporter James Surowiecki commented at the time, “the track record of the banking industry over the past two decades doesn’t inspire confidence in its devotion to the truth or to the public interest. The Barclays traders, for instance, sent e-mails casually thanking their colleagues for lying [about the rates they paid, distorting LIBOR], and sometimes talked with their supervisors about their
plans [to profit from such lies], revealing a culture in which deception was simply part of how things got done.” Bank regulators had expected the banks and the bankers to be honest, if only motivated to protect their own individual and the banks’ reputations. Bank regulators had expected to bankers and the banks to cultivate a reputation for honesty and trustworthy behavior. They were wrong. So Suroweicki draws the only conclusion possible: we need “intrusive and overbearing” regulation with aggressive enforcement.

The same conclusion applies with respect to the financial crisis.

Maybe this is not the conclusion you’d expect from a trust researcher, from me, but it fits the data. One long-time trust researcher, a sociologist, surveyed the literature on the positive effects of trust on the part of managers, within organizations, and she found that there is a “positive effect (trustworthy behavior) from supervisors treating employees as if they are reliable even when there is relatively little information about their actual character or competence” (Cook et al., 2007). This conclusion is consistent with the material presented in the first part of this paper, that trust begets trustworthiness. But in the case of the financial services industry, we have information about character, we have a pattern of behavior to study. And the pattern of behavior is consistently negative.

So trust is not an unalloyed good. And when persons or firms demonstrate untrustworthy behavior then trust in them is misplaced. It’s irrational.

Recent news only reinforces this conclusion with respect to banks and bankers.

The investment bank Goldman Sachs offered the following assurance to investors between 2006 and 2010: “Our reputation is one of our most important assets. As we have expanded the scope of our business and our client base, we increasingly have to address potential conflicts of interest, including situations where our services to a particular client or our own proprietary investments or other interests conflict, or are perceived to conflict, with the interest of another client.” And, “We have extensive procedures and controls that are designed to identify and address conflicts of interest.” And, “our client’s interests always come first.” During that period Goldman sold securities knowing that they had been constructed to fail, to benefit a hedge fund. A group of investors, led by the Arkansas Teacher Retirement system, sued Goldman for $13 billion, alleging that the bank violated these commitments to customers. providing false assurances. (Goldman acknowledged that it provided “incomplete information” to investors and paid $550 million penalty in 2011, which did not provide compensation to customers (Wall Street Journal, 2021).

The case is now at the U.S. Supreme Court. The plaintiffs are trying to recover losses during the financial crisis, but the case itself turns on a technical point in the law about representations to stockholders, and news accounts reported that some of the Court justices were confused by the issues at stake (Bravin, 2021; Liptak, 2021). What is striking is that Goldman Sach’s lawyer argued that those were “exceptionally generic and aspirational statements,” and so everyone knew that they created no substantive obligations on the firm’s part.

This is not a crisis of trust. It’s a crisis of trustworthiness: the banks and the firms have proven to be fundamentally untrustworthy, they fail to acknowledge obligations to customers, and so do not
deserve our trust.

That said, this is mostly a story about—and a warning about—about the large investment banks, it isn’t true of all banks (and I don’t say this just because I worked as a commercial loan office and as a strategist for a large bank). But it’s important here to see the problem of trust in organizational contexts, the positive role that we can all take in creating cultures of trust in organizations, the way trusting others can affect their behavior, and at the same time see that the problem of untrustworthy behavior is not solved by trusting more.

References


