Do Reputation Systems Undermine Trust? Divergent Effects of Enforcement Type on Generalized Trust and Trustworthiness[1]

Ko Kuwabara
Columbia University

Research shows that enforcing cooperation using contracts or tangible sanctions can backfire, undermining people's intrinsic motivation to cooperate: when the enforcement is removed, people are less trusting or trustworthy than when there is no enforcement to begin with. The author examines whether reputation systems have similar consequences for generalized trust and trustworthiness. Using a web-based experiment simulating online market transactions (studies 1 and 2), he shows that reputation systems can reinforce generalized trust and trustworthiness, unlike contractual enforcement or relational enforcement based on repeated interactions. In a survey experiment (study 3), he finds that recalling their eBay feedback scores made participants more trusting and trustworthy. These results are predicated on the diffuse nature of reputational enforcement to reinforce perceptions of trust and trustworthiness. These results have implications for understanding how different forms of governance affect generalized trust and trustworthiness.

> Pierre Omidyar founded eBay on a simple idea: People are basically good. This fundamental belief created a completely new kind of marketplace, forever transforming e-commerce.
>
> (eBay.com)

The problem of trust is a fundamental fact of social and economic life: from time to time, people engage in exchange of various goods and resources with clients and colleagues, buyers and sellers, teams or organizations without knowing whether, when, and how specific terms of exchange will be honored.

[1] This research was partially supported by National Science Foundation grant 0602212. The author thanks Brandy Aven, Batia Wiesenfeld, Stephan Meier, and Sigrid Seutens for generous feedback on earlier drafts. Direct correspondence to Ko Kuwabara, 703 Uris

0002-9602/2015/12005-0003$10.00

To enable successful exchange in such situations, actors often rely on formal governance, such as contracts, hierarchies, or regulations that provide monitoring and sanctioning to ensure cooperation between people who might not sufficiently trust each other (Williamson 1981). A growing body of research has found, however, that many institutional devices for enforcing cooperation can backfire, undermining people's intrinsic motivation to cooperate (Molm, Takahashi, and Peterson 2000; Bohnet, Frey, and Huck 2001; Malhotra and Murnighan 2002; Fehr and Rockenbach 2003; Mulder et al. 2006; Bowles 2008). Under strong, conspicuous forms of enforcement, such as formal contracts or monetary sanctions, acts of cooperation may be perceived to be motivated by external incentives or constraints rather than intrinsic motives—that is, in order to avoid punishment or attain rewards rather than out of one's own goodwill or moral responsibility. An ironic consequence is that, when the enforcement is removed, people are less trusting or cooperative than when there is no enforcement to begin with.

In the current research, I examine whether reputational enforcement through peer-to-peer feedback might have similar consequences as contractual enforcement, making people less trusting or trustworthy toward strangers in the absence of enforcement. There is perhaps no better test of *generalized* trust and trustworthiness—for example, the belief that "people are basically good"—than online markets ruled by anonymity and pervasive automation, where durable relations are notably absent and virtual strangers come to exchange goods and services halfway across countries, often without the prospect of future interactions. To mitigate the risks of anonymous transactions, many online markets maintain reputation systems, such as eBay's Feedback Forum, that provide platforms for collecting, aggregating, and displaying numerical ratings and reviews of various products and services. Following eBay's lead, peer-to-peer reputation systems have become one of the most ubiquitous institutions on the Internet. Users and scholars alike have marveled at how effectively even the simplest reputation system can curtail fraud and sustain trust in markets that are too large and decentralized for contractual enforcement or centralized policing (Kollock 1994; Resnick et al. 2000; Dellarocas 2003*a*; Bolton, Katok, and Ockenfels 2004). Yet, such optimism may be questioned if reputation systems are eroding our trust and trustworthiness offline.

I also compare reputational enforcement to another important form of governance: relational enforcement through direct reciprocity in repeated interactions (Axelrod 1984). In the absence of external enforcement, people naturally repeat interactions with specific others in order to curtail mal-

Hall, 2033 Broadway, Columbia Business School, New York, New York 10027. E-mail: kk2558@columbia.edu

feasance and reduce uncertainty on their own. However, such relations can lock people in and inhibit exploration that may cultivate generalized trust and trustworthiness toward unfamiliar others.

My theoretical argument builds on the idea that generalized trust and trustworthiness toward strangers develop through gradual exposure to strangers under conditions of moderate social risk that test and affirm people's perceptions about themselves and each other as intrinsically or dispositionally trusting or trustworthy, even in the absence of strong enforcement (Molm et al. 2000; Macy and Sato 2002). These conditions, I argue, are more likely under reputational enforcement than either contractual or relational enforcement or in the absence of any enforcement. Anonymous exchanges are too risky without enforcement and too safe under contractual enforcement to cultivate trust and trustworthiness. In comparison, reputational enforcement provides social feedback that can reinforce intrinsic motivations (Deci, Koestner, and Ryan 1999) instead of weakening them by eliminating risk entirely. Unlike relational enforcement, reputational enforcement also promotes exposure to new partners.

Across three sets of web-based experiments, I find compelling support for my central prediction that reputational enforcement—compared to contractual or relational enforcement—can reinforce both generalized trust and trustworthiness. These results have direct implications for understanding the consequences of market designs and institutions for user experiences (Dellarocas 2003*a*; Nissenbaum 2004; Pavlou and Gefen 2004). Although online feedback mechanisms have been studied extensively (e.g., Kollock 1999; Resnick et al. 2000; Dellarocas 2003*a*; Bolton, Greiner, and Ockenfels 2012; Diekmann et al. 2014), how they affect people's behaviors and decisions outside of particular systems or communities has been largely neglected. As online economies continue to grow in scale and form, and as IT-based solutions pervade our lives inside and outside organizations, understanding how they shape our lives, both online and offline, is an important task for social scientists in this increasingly digital and borderless age (Cook, Snijders et al. 2009).

More generally, by comparing different forms of enforcement that prevail in markets (reputational), hierarchies (contractual), and networks (relational) (Powell 2003), current research contributes to our understanding of how different forms of governance affect generalized trust and trustworthiness. Although social control through peer enforcement (e.g., Durkheim 1893; Homans 1961; Coleman 1990) and institutional governance (Williamson 1981; Powell 2003; Cook, Hardin, and Levi 2005) has long been a central topic in sociology, far less attention has been paid to how the effects of enforcement might "spill over" to exchange decisions and actions outside of specific relations or institutions.

THEORETICAL BACKGROUND

The Problem of Generalized Trust and Trustworthiness

How trust develops between strangers has been a topic of considerable interest across social sciences (e.g., Putnam 1993; Kramer and Tyler 1996; La Porta et al. 1997; Yamagishi 1998; Hardin 2002; Cook, Snijders et al. 2009). In contrast to personalized trust that develops between particular individuals through repeated interactions, generalized trust refers to trust toward strangers based on dispositional tendencies to trust unfamiliar others (Rotter 1971), a general belief in human benevolence (Kosugi and Yamagishi 1998), or moral obligation to "trust as if people are trustworthy" (Uslaner 2002). Similarly, generalized trustworthiness refers to intrinsic motivations or dispositional tendencies to refrain from cheating and betraying strangers on the basis of a moral faith that trustworthiness is part of social obligations toward people (Simpson and Eriksson 2009).

While much light has been cast on the development of trust between particular individuals over repeated interactions (e.g., Axelrod 1984; Kramer 1999; Molm et al. 2000), scholars continue to debate how such experiences of personalized trust might generalize to trust and trustworthiness toward strangers in other contexts. According to the social learning perspective, generalized trust develops as individuals extrapolate from their personal experiences in localized settings—such as schools, neighborhoods, or civic associations—to form certain expectancies about similar others or contexts (Putnam 2000; Stolle 2001; Hardin 2002; Glanville 2004; Glanville and Paxton 2007). Through repeated interactions with different friends, neighbors, coworkers, or community members, people learn what types of people can be trusted. In this view, generalized trust is grounded in the basic process of *social perception* in which people actively learn to discern who is or is not trustworthy by making sense of observable traits and signals, learned information, or other telltale signs of character (Frank 1988; Macy and Skvoretz 1998).

The social learning perspective has been challenged on the grounds, however, that localized interactions—whether through civic associations, sociopolitical affiliations, or informal networks of friends and neighbors—are generally too homogeneous and cohesive to cultivate generalized trust toward people in more diverse social settings outside of embedded relations (Uslaner 2002, 2012). Localized interactions may be safe grounds for personalized trust, but it is theoretically unclear how they might cultivate beliefs that "most people" are trustworthy. What, for instance, do people learn from one book club, about members of another book club or a bowling team or a new neighborhood? In an extensive review of studies using survey data to examine generalized trust, Uslaner (2002) in fact finds no

compelling evidence that participation in civic, fraternal, religious, political, or recreational associations increases generalized trust.[2]

It is also not clear how the logic of social learning might account for generalized trustworthiness, that is, why some people resist taking advantage of others. Learning to discern who is trusting or not is a moot issue, and it may in fact encourage dishonesty, because the sequential nature of trust exchange ensures that the truster (e.g., buyer) must commit to the exchange first. Rather, the critical issue for the trusted party is learning to be trustworthy and to refrain from cheating regardless of who the truster is; in this regard, social perception is agnostic about who refrains from exploiting others and why.

An alternative hypothesis links generalized trust and trustworthiness to reinforcement learning (Macy and Sato 2002) rather than social learning—that is, to the reinforcement of *self-perception* rather than social perception about others. The logic of self-perception is simple: people form attitudes and beliefs by observing their own actions and inferring what caused them, just as they form impressions or opinions about others by observing their actions (Bem 1967). Thus, rather than learning to discern who is or is not trustworthy, people come to view themselves as a trusting or trustworthy person by observing their own acts and learning to value trust and trustworthiness as a moral or personal virtue (Uslaner 2002). In this view, trust is motivated not by the expected utility of trusting someone who may turn out to be trustworthy but by a sense of self that identifies with acts of trust, that is, by the conviction that one is a trusting person. Self-perceptions of trust do not preclude strategic decisions or cost-benefit calculus based on social perceptions about others, but they can precede or anchor such decisions. It is not unlike the concept of dispositional trust, based on stable and innate personality attributes (Rotter 1971; Smith 2010), but self-perception emphasizes situational changes in beliefs rather than stable traits. For instance, a recent study by Grand and Dutton (2012) demonstrates that simply reflecting on being a giver (self-perception) makes people more giving than reflecting on receiving benefits from others (social perception) in social exchange. Moreover, and in contrast to perceptions about other people's trustworthiness, perceptions about one's own trustworthiness can directly promote acts of trustworthiness by invoking a self-image of oneself as someone who resists opportunism, even in situations in which dishonesty yields a greater immediate payoff.

[2] It should be noted that Uslaner's (2002) results are still far from conclusive. Reviewing the literature, Paxton (2007) notes that empirical support for the relationship between trust and civic association remains rather mixed and calls for greater attention to different forms of associations.

Self-perceptions of trust and trustworthiness are often traced to socialization through early parenting, religious upbringing, or moral education (Rotter 1971; Luhmann 1979; Fukuyama 1995). Uslaner (2002) argues that this form of trust—trusting as if people are trustworthy—is a moral value, linked closely to one's identity and general faith in the goodwill of the other. Exchange theory suggests, however, that perceptions of trust and trustworthiness, including self-perceptions, may also be reinforced by successful exchanges that yield material payoffs (Macy and Sato 2002), emotional "buzz" (Lawler 2001), or social approval (Willer 2009) under conditions that lead actors to attribute cooperative acts to themselves and each other rather than external enforcement. Such conditions, I argue below, are more pervasive under reputational systems than under contracts.

Contractual Enforcement

Between strangers, many transactions occur only under formal contracts that reduce the risk of exploitation through binding agreements that impose relatively direct and tangible sanctions in case of violations or negligence. Ironically, contractual enforcement—and the threat of tangible sanctions in general—can backfire, reducing cooperation once enforcement is removed to a greater extent than when there are no contracts or sanctions to begin with (Malhotra and Murnighan 2002; Fehr and Rockenbach 2003; Mulder et al. 2006). One reason is that formal enforcement may interfere with people's social perception through which they come to attribute each other's cooperation to their intrinsic motives or dispositions to refrain from opportunism. Although the goal of enforcement is to reduce risks of exploitation, the development of trust depends critically on risk, because only in the presence of potential exploitation can acts of cooperation be reliably or meaningfully attributed to each other's goodwill and benign intentions. Under binding agreements, cooperation may be perceived to be coerced or externally motivated because people cannot adequately demonstrate their willingness to place or honor trust. As a consequence, once enforcement is removed, people realize that they actually have not learned to trust each other or to identify who is trustworthy or not, thus withholding cooperation (Malhotra and Murnighan 2002; Simpson and Eriksson 2009). In a market with a heterogeneous population of honest and dishonest sellers, enforcement can also compromise social perceptions of trust by reducing the need to safeguard against the risk of exploitation, thus "crowding in" dishonest sellers who, in the absence of enforcement, may be driven out of the market by wary buyers (Akerlof 1970; Bohnet et al. 2001). The threat of enforcement might keep dishonest types in check, but once enforcement is removed, the market begins to collapse unless people quickly learn to identify honest and dishonest types rather than withdrawing from the market altogether.

Another reason why formal enforcement might sustain cooperation but reduce trust is that it can interfere with people's self-perceptions of their own trust and trustworthiness. By this account, trusters might attribute their own decisions to place and honor trust to the threat of enforcement and come to perceive themselves as less trusting or trustworthy than they actually are (Kramer 1999; Simpson and Eriksson 2009). By (mis)attributing their own actions to extrinsic incentives, people lose their intrinsic motivation to act prosocially (Gneezy and Rustichini 2000; Gneezy, Meier, and Rey-Biel 2011), much like students losing interest in the very subjects in which they care excelling if they receive financial incentives for good grades (Lepper, Greene, and Nisbett 1973; Deci et al. 1999).

Reputational Enforcement

In online markets that are too large and decentralized for formal enforcement by centralized agents, reputation systems have proven to be a highly effective solution to the problem of trust. Peer-to-peer transactions are significantly riskier online than in face-to-face exchanges, because people must take a leap of faith in faceless strangers halfway across the Internet, separated in both time and space. Yet, remarkably, eBay has become the largest peer-to-peer market in the history of mankind without offering so much as a warranty or central policing, only a simple reputation system that allows buyers and sellers to rate each other (+1 or −1) and leave comments (e.g., "Nice person to do business with! Highly recommended") after transactions. EBay posts users' feedback points next to their screen names to help mitigate the risk of online transactions and attract future buyers and sellers. Each user can receive only one feedback point from a given partner, and they are not retractable once posted.

A substantial body of research attests to the effectiveness of even the simplest reputation systems. Research on eBay finds robust effects of both positive and negative feedback points on likelihood of sales and final auction prices (e.g., Bajari and Hortaçsu 2003; Resnick et al. 2006), while laboratory studies show that displaying no more than binary feedback (a positive or negative point) from just one previous interaction suffices to sustain high levels of cooperation between strangers engaged in anonymous, one-shot exchanges (e.g., Keser 2003; Bolton, Katok, and Ockenfels 2005). Following eBay's lead, online reputation systems have become a ubiquitous feature of online markets and communities (Kollock 1999; Resnick et al. 2000).

Thus far, research on online reputation systems has focused almost exclusively on dyadic trust between particular buyers and sellers (e.g., Resnick et al. 2006; Dellarocas 2010), although few studies have gone beyond this scope to examine the role of reputation systems in promoting generalized

trust toward the general community of sellers in a particular market (Pavlou and Gefen 2004). The scope of my interest is even broader: how people's exposure to reputational enforcement affects their generalized trust and trustworthiness outside of particular reputation systems or markets. This means my focus is not on the effects of particular reputations associated with particular individuals or communities but on the effects of reputation systems. Do reputation systems make people more or less intrinsically trusting and trustworthy—that is, even in the absence of feedback mechanisms?

While contracts and reputation systems present alternative forms of enforcement, one key difference is that reputational enforcement is generally more diffuse than contractual enforcement. In contrast to contractual enforcement, reputational sanctioning—posting and sharing feedback—is much less direct and tangible. Providing feedback has no direct material consequences for the recipient, only indirect and often marginal consequences insofar as transactions are enforced by other community members who collectively monitor and aggregate reputational information about each other, and victims of malfeasance receive no material reparations from submitting feedback. Thus, the effectiveness of reputational enforcement depends on whether others in the market will actually observe the seller's reputation and act appropriately (e.g., avoid or penalize sellers with negative reputations) in future transactions (Greif et al. 1998; Dellarocas 2003*b*).

This diffuse nature of reputational enforcement has important consequences for generalized trust. Although trust grows under risk (Molm et al. 2000), unenforced transactions present too much risk while contractual enforcement creates too little risk to cultivate perceptions of trust and trustworthiness. Under such conditions, people are unlikely to make strong inferences about each other's trustworthiness or their own willingness to trust, because they may be better off avoiding unenforced transactions altogether, or, given contracts, they might enter exchanges with strangers too readily without proper due diligence and discretion. By comparison, reputation systems create conditions of relatively moderate risk, enough to spur measured cooperation without placing blind bets. Under such conditions, market transactions can expose people gradually to exchange opportunities with strangers in ways that test and affirm their own sense of trust and integrity. Although simulation studies have explored this logic (Macy and Sato 2002), the current research, to my knowledge, is the first to empirically test idea.

Another difference between contracts and reputations is that social signaling plays a much more prominent role in reputational enforcement. Signals are observable actions or markers that identify people with certain unobservable traits and qualities, such as a college degree that reflects one's intelligence (Spence 1974). Contracts are primarily designed to enforce co-

operation by providing *negative sanctions* for *noncompliance* regardless of unobservable traits or intentions, and they do more to mask than to reveal signals of trust or trustworthiness: under contracts, it is difficult (and unnecessary) to tell who is actually trustworthy or not. If contracts are optional, the very act of proposing contracts can signal distrust rather than trust and dampen the partner's motivation to honor or reciprocate trust (Fehr and Falk 2002; Puranam and Vanneste 2009). In comparison, the primary function of many reputation systems is to create *positive signals* that not only help identify who is trustworthy but also frame transactions in terms of finding people to trust rather than constraining their behavior (Kollock 1999).[3] Although negative feedback is used from time to time, online feedback is overwhelmingly positive (Resnick et al. 2006; Bolton et al. 2012). Thus, people in reputation systems evaluate one another and enter transactions primarily on the basis of positive reputations that elicit perceptions of quality, trustworthiness, and collective social approval rather than negative sanctions that presume strangers should not be trusted. And yet, because reputational enforcement is diffuse, reputation systems still require buyers to place a certain amount of trust in each seller, even if they can observe seller reputations ex ante. In this sense, contracts are designed to replace or substitute for trust (Cook et al. 2005), whereas reputation systems are designed to invoke and complement trust, reinforcing people's perception that they are willing to trust strangers in the absence of strong enforcement.

The informational nature of social feedback in reputational enforcement is critical for trustworthiness as well. Overjustification theory (Deci et al. 1999, 2001) suggests that tangible incentives (e.g., monetary rewards or fines) can diminish intrinsic motivations, but intangible incentives (e.g., verbal feedback such as praise) can enhance them. Tangible rewards are generally perceived to be more controlling or coercive, particularly when they are expected, inhibiting attributions to intrinsic motivation. By this token, contracts that impose sanctions for specific terms of exchange are likely to be regarded as more controlling, undermining prosocial motives (Malhotra and Murnighan 2002; Simpson and Eriksson 2009). Studies show that even weak incentives based on small monetary fines or rewards can erode intrinsic motivations (Tenbrunsel and Messick 1999; Gneezy and Rustichini 2000).

In contrast, intangible rewards provide informational feedback that affirms one's intrinsic motivation and sense of competence or worth without controlling people's behavior. For example, reputational enforcement is based on feedback points that aggregate over time into a reputation, often

[3] I thank an *AJS* reviewer for suggesting this point.

earning special status (such as stars), reinforcing the person's moral values or identity, much like a student who becomes an "A student" instead of merely acing tests.[4] Furthermore, it is noteworthy that reputational enforcement is often unpredictable, since posting feedback is largely voluntary in many reputation systems. It has been known since Skinner's (1969) work on operant conditioning that behavioral reinforcement is significantly more effective on variable rather than regular schedules, that is, when feedback is somewhat unpredictable or unexpected.

Hence, unlike contractual enforcement that might interfere with both social and self-perceptions of trust and trustworthiness, the combination of diffuse sanctions and social feedback in reputation systems can create microconditions that facilitate the development of both generalized trust and trustworthiness. From these ideas, I expect reputational enforcement to produce higher generalized trust (hypothesis 1) and trustworthiness (hypothesis 2) than contractual enforcement.

Relational Enforcement

In the absence of contractual or reputational enforcement, actors often engage in repeated interactions with familiar partners, which enables relational enforcement through reciprocity (Gouldner 1960; Axelrod 1984; Granovetter 1985). Through positive reciprocity, actors reward each other by offering continued cooperation; through negative reciprocity, actors punish each other by reducing cooperation or terminating relationships. Over time, exchange partners develop mutual knowledge, identification, and commitment that safeguard against the risks and uncertainties of one-shot, unenforced exchanges.

In game theory, relational enforcement, or relational contracting, is often modeled as repeated games with private reputation gleaned from direct interactions rather than public or third-party reputation aggregated from others' experiences (Wilson 1997). In this sense, relational enforcement is a form of reputational enforcement. However, the relational bonds that develop between people through repeated interactions, including trust, are relation specific and limited to particular partners. Indeed, studies have demonstrated negative effects of particularized trust on generalized trust: strong ties to familiar others can reduce trust toward outsiders (Hayashi and Yamagishi 1998; Yamagishi, Cook, and Watabe 1998; Stolle 2001). Conversely,

[4] Kollock (1999) has made a similar observation about the role of positive vs. negative reputations (also see Cook and Hardin 2001; Yamagishi et al. 2009). Whereas negative reputations can create incentives to change or relinquish one's identity (e.g., by relocating to a new market or rebranding oneself), positive reputations play an important role in affirming and stabilizing one's existing identity.

if generalized trust is reinforced through gradual exposure to strangers, we should expect reputational enforcement to produce greater generalized trust than relational enforcement in fixed dyads (hypothesis 3). Likewise, if generalized trustworthiness is reinforced by social feedback from different partners, then, reputation systems that are specifically designed to aggregate feedback from different partners should produce more generalized trustworthiness than relational enforcement (hypothesis 4).

OVERVIEW OF THE STUDIES

To test the four hypotheses, I conducted three sets of web-based experiments designed to simulate particular aspects of reputational enforcement. In all three studies, transactions were modeled after the investment game (Berg, Dickhaut, and McCabe 1995), in which one actor is assigned to the role of the truster and the other the trustee. In each round, the truster is given an endowment (e.g., 10 points) and asked to send or entrust any portion of it to the trustee. The trustee receives a multiple (typically 3×) of the entrustment and is asked whether to share any portion of it with the truster. This task has been used extensively to study generalized trust (e.g., Glaeser et al. 2000) and trustworthiness (e.g., Simpson and Eriksson 2009). Using this paradigm, study 1A tested hypotheses 1 and 3 for generalized trust, and study 1B tested hypotheses 2 and 4 for generalized trustworthiness, after exposing participants to different types of enforcement over a series of anonymous exchanges. Study 2 modified study 1 to rule out alternative explanations (e.g., contrast effect) and further explore how enforcement affects generalized trust and trustworthiness using self-reported measures. Finally, in order to probe the theoretical mechanism more directly and provide external validity, study 3 used actual eBay users to show that their feedback scores from eBay can reinforce their self-perceptions of trust and trustworthiness.

No one study was designed to directly test between social and self-perception since they can co-occur in many situations, and both explanations converge closely on the same predictions. Study 3 manipulates, but does not measure, self-perception. In the discussion section below, I consider several reasons why self-perception still offers a more compelling explanation than social perception for my results across the three studies and for exchange theory more generally.

Finally, it is important to distinguish the predicted effects of reputational enforcement that reinforce perceptions of trust and trustworthiness from dependence on reputational information that reduces people's willingness to engage in transactions outside of reputation systems. For many people, it is hard to imagine going to a new restaurant, watching a new movie, or shopping online without consulting consumer reviews first. It is

no surprise that such people would be more likely to avoid new services or businesses when reputation systems are unavailable. What has not been tested thus far, however, is whether they are still more trusting or trustworthy as a result of their experiences in reputation systems than they would otherwise be without prior exposure to reputational enforcement.

STUDY 1A: DO REPUTATION SYSTEMS UNDERMINE TRUST?

Study 1A was designed to test hypotheses 1 and 3 for trust using investment games to simulate transactions in an online market for "clients" hiring "agents" for "professional services" under different types of enforcement. Following the standard paradigm for studying the spillover effects of enforcement (Bohnet et al. 2001; Malhotra and Murnighan 2002; Simpson and Eriksson 2009), the experimental treatments first exposed participants to a series of transactions with anonymous partners under contractual, relational, or reputational enforcement before presenting several additional rounds of one-shot exchange without any enforcement. This pretest-posttest design helps evaluate how participants change after exposure to different types of enforcement. To isolate the effects of enforcement, partners in both studies were simulated and programmed to offer full cooperation. The transactions were framed as online transactions for high- or low-quality services.[5]

Experimental control was crucial for testing my hypotheses under four conditions that are difficult to achieve in natural settings: (1) creating comparable conditions for contractual, relational, and reputational enforcement, (2) removing enforcement at a specific point while holding other market conditions constant, (3) measuring the effects of enforcement on trust and trustworthiness separately, and (4) ensuring high levels of cooperation by partners. The third condition is critical because my substantive goal is to understand how different types of enforcement affect individual buyers and sellers rather than a community of buyers and sellers as a whole. The fourth condition specifies an important scope condition for my predictions: the enforcement system must be (perceived as) reasonably effective if it is to invoke perceptions of trust and trustworthiness (Pavlou and Gefen

[5] Although this is a departure from markets for goods like eBay, services and products are treated as interchangeable for the purpose of this study insofar as they both present moral hazard. I used services in order to facilitate attributions to people rather than products. Cook, Cheshire et al. (2009) note that the basis of evaluation for transactions for service vs. goods in online markets may be different. Buyers care more about the motivation than the competence of the seller to provide high-quality service, but they care more about the competence than the motivation of seller to deliver high-quality goods. Such issues are absent in the current study.

2004; Cheshire 2011). Anecdotal and empirical evidence shows that many online markets with reputation systems, including eBay, are indeed overwhelmingly safe despite enduring concerns about fraud and deception (Kollock 1999; Resnick et al. 2000; Diekmann et al. 2009).

Design and Procedures

One hundred and one students (58 male, 20.5 years old) completed the study. They were recruited using a standard protocol for behavioral experiments, via mailing lists and flyers posted around campus, to participate in an "online market study" that promised $14–$20 in cash, based on performance, paid upon full completion of the study "for just a few minutes of your time each day for a total of less than 60 minutes over the course of five to seven days, played from any computer with Internet access." After signing up for the study, participants were directed by e-mail to a custom web page that guided them through the consent form, experimental instructions, and practice trials before they were "randomly" assigned to play the client role in a series of online transactions with anonymous exchange partners playing the agent role. In each transaction, the client was given an endowment of 10 points for hiring an agent and decided how many hours of service, $0 \le s \le 10$, to hire the agent for at 1 point per hour. In turn, the agent decided whether to provide high- or low-quality service for the total hours hired. High-quality service earned the client $3s$ points and the agent s points, while low-quality service earned the client $.5s$ points and the agent $2s$ points in each transaction. Any remaining amount $10 - s$ was added to the client's total earnings. Participants were told that the full payoffs were public knowledge to both clients and agents. Thus, the dilemma for the client was deciding how many hours to hire the agent for without knowing the quality of service that would be delivered. The client made a profit only if the agent provided high-quality service.

Each participant was assigned randomly to one of four experimental conditions, or "markets." In the *unenforced* condition ($N = 30$), the market consisted of a series of one-shot transactions without contractual, reputational, or relational enforcement. In the *contract* condition ($N = 23$), each transaction was automatically placed under a contract that imposed, at no cost to the client, a penalty of 10 points on the agent with a 90% probability for providing low-quality service. In the *fixed-partner* condition ($N = 24$), participants repeated exchanges with the same agent. In the *reputation* condition ($N = 24$), each agent was given a numerical reputation score, and clients were given an opportunity to rate each transaction with binary feedback (+ or –). Participants were given information only about the specific market to which they were assigned.

Enforcement can be said to undermine trust if removing it results in lower levels of trust compared to the unenforced (baseline) condition. Thus, each condition consisted of two phases: participants were told that the experiment would last up to 15 rounds, but after completing eight transactions in their respective markets (phase 1), they were asked to complete additional transactions in the unenforced condition (phase 2). The dependent measure of trust was hours of service purchased in phase 2.

In previous studies (e.g., Bohnet et al. 2001), enforcement was simply removed without pretext. However, it is unlikely that reputation systems are suddenly removed from real online markets. A more plausible scenario is that people (are forced to) pursue alternative transactions outside of reputation systems for better prices or customers; for instance, one might resort to Craigslist if a particular item is not available on eBay. Thus, participants were told that, because "no new agents are currently available in your assigned market," they need be transferred to a different market, namely, the unenforced condition with a different set of participants. Apart from ensuring greater realism, this also minimizes the "history" effect—the possibility that participants might simply continue to behave as if they were still in the same market with the same population of partners as in phase 1 before the removal of the enforcement. Participants were not told that they would be transferred to the unenforced market until round 9 or that they would complete exactly three more transactions in phase 2. Table 1 summarizes the experimental design.

In order to isolate the effects of enforcement on trust from partner effects, all participants except those in the fixed-partner condition faced the same sequence of agents (table 2), all preprogrammed to deliver high-quality service in every round. In the reputation condition, all agents were associated with largely positive reputation scores in order to ensure that the reputation system was perceived to be reasonably effective. However,

TABLE 1
SUMMARY OF THE EXPERIMENTAL DESIGN IN STUDY 1A

Condition	Phase 1	Phase 2
Unenforced	Unenforced transactions with random partners	Unenforced transactions with random partners
Contracts	Random partners under contractual enforcement	Unenforced transactions with random partners
Fixed partner . . .	Fixed partner under relational enforcement	Unenforced transactions with random partners
Reputation	Random partner under reputational enforcement	Unenforced transactions with random partners

NOTE.—Phase 1 is rounds 1–8; phase 2 is rounds 9–11.

TABLE 2
AGENTS IN STUDY 1A

Round	Agent ID	Transactions Completed*	Positive Feedback Points (%)*
1 . . .	3E	17	100
2 . . .	8A	24	100
3 . . .	12C	20	100
4 . . .	7B	21	95.2
5 . . .	5C	18	100
6 . . .	14A	23	91.3
7 . . .	9B	18	100
8 . . .	4D	26	100
9 . . .	2C	NA	NA
10 . . .	5B	NA	NA
11 . . .	10A	NA	NA

NOTE.—Agent ID did not change in the fixed-partner condition in rounds 1–8.

* Information was presented in the reputation condition only.

agents in rounds 4 and 6 were given a negative point for realism and to see whether participants are paying attention to the reputation scores.

Contracts and reputation systems exist in various forms. For instance, as Durkheim (1893) noted, some contracts are repressive or retributive (penalizing offenders) while others are restitutive (compensating victims). For the purpose of the current research, I focused on retributive enforcement in order to ensure that the enforcement mechanisms were as comparable as possible across conditions except in their defining differences. For instance, although both voluntary and involuntary contracts have been shown to undermine cooperation (Malhotra and Murnighan 2002), the contracts in the current studies were involuntary and free, provided in every transaction at no cost to the client. This ensures comparability to many reputation systems that are also involuntary and free (i.e., every transaction could be rated at no real cost). The critical difference is that enforcement resulted in the agent incurring direct financial penalties under contractual terms, whereas reputational enforcement imposed no such penalties that reduced participant earnings directly.[6] Because agents were programmed to always fully cooperate, however, no contractual penalties were actually imposed on agents, thus holding constant direct material consequences of enforcement and varying only their perceived effects.

[6] This can create differences in final earnings that may affect participants' trust. However, because partners were programmed to only cooperate, no such differences actually emerged.

Setting

Participants completed the entire experiment online by logging into a custom website. Sometime after each transaction, the server sent e-mail to the participant notifying that the result of the most recent transaction was now available. The results screen showed how many hours of service and at what level of quality the agent provided (always high), the client's total earning from the transaction and the client's net cumulative earnings. At this point, clients in the contract condition were told that the agent satisfied the terms of the contract (thus, the agent could not be punished), while clients in the reputation condition were asked to submit a rating. Those in the other conditions received no such message. Next, the screen matched the client with a new agent for the next transaction (or the same agent in the fixed-partner condition). Participants were given 10 hours to complete each transaction and five days to complete all 11 transactions.[7]

The experiment was conducted online to create greater psychological realism, including perceptions of risk that stem from asynchronous, online interactions. The trade-off is loss of experimental control. Outside of the laboratory, participants are more likely to talk to each other, get distracted, or drop out. So long as noise is random across the conditions, however, such issues are unlikely to bias the results. For precaution, participants were instructed not to discuss the study with others until the end of the semester. During debriefing, they were asked whether they discussed the study with anyone. No such case was reported.

Results and Discussion

Figure 1 shows the mean level of trust (the number of hours commissioned) in each round by condition. In phase 1 (rounds 1–8) trust was substantially higher in all of the treatment conditions compared to the baseline level in the unenforced condition, which hovered between 30% and 40%. Not surprisingly, the contract condition produced nearly full cooperation in all rounds. In contrast, trust rose more slowly in both the fixed-partner and the reputation conditions. Trust in the reputation condition dips markedly in response to the agent with a negative reputation in round 6 but recovers immediately in the next round. These observations were submitted to Tobit

[7] If the client did not initiate the next transaction within 10 hours, the transaction expired, and the server matched the client with the next agent. Up to two e-mail reminders were sent out to make sure transactions did not expire. Among participants who completed the entire study, only 0.8% of all transactions expired, 14 participants let at least one transaction expire, four participants let two transactions expire, and no one let more than two transactions expire. In statistical analyses, expired rounds were denoted as 0 points in trust and dummied as expired.

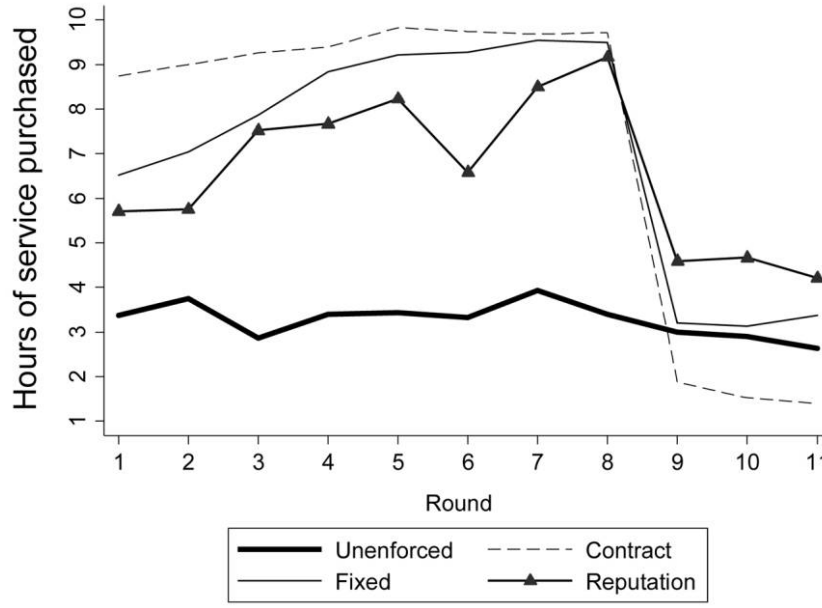


FIG. 1.—Levels of trust (hours of service purchased) by condition in study 1A

regression to account for the dependent variable censored at 10 maximum hours, with standard errors clustered for repeated measures.[8] The results indicate that, in phase 1, the contract condition produced more trust than the fixed-partner condition ($\beta = 4.41$, $P = .002$), which produced more trust than the reputation condition ($\beta = .75$, $P = .001$), which produced more trust than the unenforced condition ($\beta = .97$, $P < .001$).

To test hypotheses 1 (reputation > contracts) and 3 (reputation > fixed relations), I examined trust in phase 2. As shown in figure 1, the trust levels dropped precipitously after round 8 in all treatment conditions. The contract condition resulted in both the biggest drop and subsequently the lowest level of trust of all, significantly below the unenforced condition ($\beta = -2.16$, $P = .006$), consistent with previous research on the detrimental effects of contracts. The fixed-partner condition showed no difference from the unenforced condition ($\beta = -.22$, $P = .59$), but it was significantly higher than the contracts condition ($\beta = 2.30$, $P = .01$). In contrast, trust in the reputation condition remained roughly 1.6 service hours above the unenforced condition ($\beta = .43$, $P < .001$) and the fixed-partner con-

[8] All tests reported are two-tailed. Race, gender, age, and major had no significant effects and are not reported.

dition ($\beta = .54$, $P = .01$). Altogether, these results provide strong support for both hypothesis 1 and hypothesis 3.

It may be argued that rounds 9–11 in the treatment conditions are better compared to the first three rounds of the unenforced condition than the last three, when participants in the treatment conditions are first exposed to the unenforced condition. Redoing the comparisons accordingly shows robust support for hypotheses 1 and 3: rounds 9–11 in the reputation condition > rounds 1–3 in the unenforced condition ($\beta = .28$, $P = .002$); rounds 1–3 in the unenforced condition = rounds 9–11 in the fixed-partner condition ($\beta = -.15$, $P = .67$); and rounds 1–3 in the unenforced condition > rounds 9–11 in the contracts condition ($\beta = 2.55$, $P < .001$). The analyses were also repeated using all 11 rounds of the unenforced condition against the last three rounds of the treatment conditions, producing stronger differences.

STUDY 1B: DO REPUTATION SYSTEMS UNDERMINE TRUSTWORTHINESS?

Design and Procedures

To test hypotheses 2 and 4 for trustworthiness, study 1B modified the experimental design above by assigning participants to the role of agents against (simulated) clients who decided whether to hire them for 10 hours of service. If a client chose to hire the agent, the agent decided how many hours of high- ($1 \le h \le 10$) or low-quality service ($10 - h$) to provide. High-quality service earned the agent h points and the client $2h$ points, while low-quality service earned the agent $3(10 - h)$ points and the client h points per transaction.

As in study 1A, three enforcement conditions were created in addition to the unenforced condition ($N = 21$), with minor differences. In the contracts condition ($N = 18$), high-quality service was enforced with 90% probability, which cost the agent 4 points for each hour of low-quality service delivered. In the fixed-partner condition ($N = 23$), the client was programmed to withhold cooperation (commission) in the next round with the inverse probability of the number of high-quality service hours provided, $1 - (h/10)$. For instance, after receiving seven hours of high-quality service (thus three hours of low-quality service), the client chose zero hours in the next round with 30% probability. In the reputation condition ($N = 21$), the agent's reputation score was aggregated from all rounds as a percentage of positive points. Each agent began with a null reputation score and received a positive feedback point with (unbeknownst to participants) 10% probability for each hour of high-quality service rendered and a negative feedback point otherwise. For example, providing six hours of

high-quality service earned a positive feedback point with 60% probability. In addition, clients were programmed to reject the agent, on the basis of the most recent feedback point received. Specifically, the agent was rejected and rematched with another client with 65% probability after receiving a negative rating and 5% after a positive rating. Agents were notified by e-mail every time they were rejected, and they forfeited the exchange round entirely after three rejections. Although rejections did not cost the agents, and less than 2% of rounds were forfeited, they served to reinforce the idea that clients trusted agents on the basis of reputation scores.

The same pretest-posttest design as study 1A was used to compare trustworthiness across conditions. The instructions specified that the experiment would last up to 30 transactions, but each agent actually completed 20 transactions before phase 2, in which they played an additional three transactions without enforcement. In phase 2, agents were hired in every round; no one was rejected. More transaction rounds were provided in study 1B to ensure enough rounds to accrue comparable numbers of feedback points as in study 1A. Except those in the fixed-partner condition, participants were told that they would not be matched with the same client more than once.

Eighty-three participants (39 men, 20.04 years old) completed the study. They were recruited for cash payment as in study 1A. After signing up, participants were led to a web page that guided them through consent, instructions, and practice trials before they were assigned to the agent role. All transactions took place online over a period of up to 10 days. At the beginning of each transaction, the server contacted the participant via e-mail that a particular client has chosen him or her for 10 hours of service. The next screen asked how many hours of high-quality service to deliver, calculated the agent's net earnings from the transaction, and completed the transaction by ostensibly releasing an e-mail message to the client informing the agent's decision. To ensure timely decisions, e-mail reminders were sent out to participants to respond to their clients, as in study 1A. In study 1B, participants were told that letting a transaction expire results in providing zero hours of high-quality service and were penalized with a negative rating or contract enforcement. Among those who completed the study, less than 3% of all transactions expired.

Results and Discussion

Figure 2 plots the number of high-quality service hours provided in each round. As expected, the treatment conditions sustained significantly higher levels of effort than the unenforced condition during phase 1 (rounds 1–20), and the contract condition produced almost full cooperation from the beginning. Tobit regression correcting for repeated measures shows that the

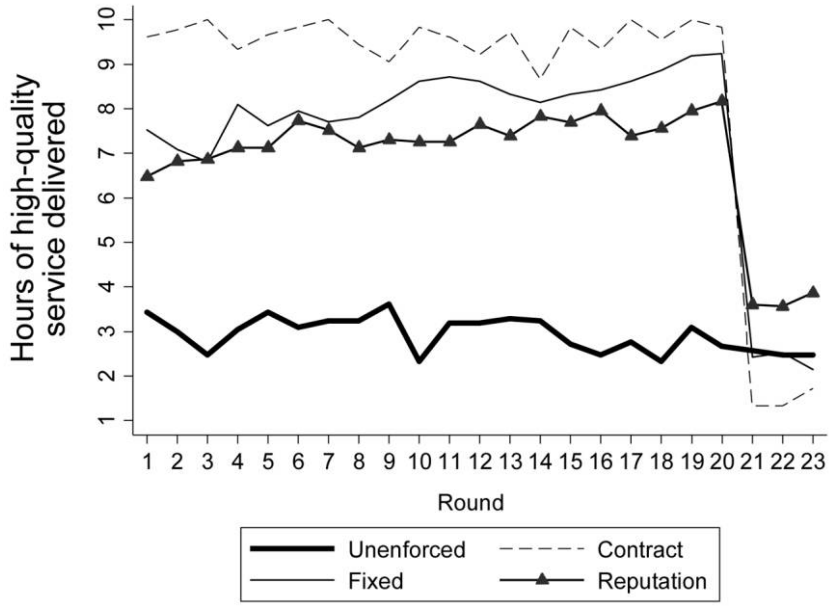


FIG. 2.—Levels of trustworthiness in study 1B (hours of high-quality service delivered).

fixed-partner condition was slightly but consistently less effective than the contracts ($\beta = -2.65$, $P < .001$) in sustaining trustworthiness, and the reputation system even less so, relative to the fixed-partner condition ($\beta = -1.743$, $P = .025$).

In phase 2 (rounds 21–23), participants were removed from their treatment conditions and matched with new partners in the unenforced condition, resulting in trustworthiness dropping sharply in all treatment conditions. As expected, the biggest drop occurred in the contract condition, falling marginally below the baseline level in the unenforced condition ($\beta = -1.76$, $P = .078$). The fixed-partner condition dropped to the baseline level ($\beta = .055$, $P = .82$) but still significantly above the contract condition ($\beta = .95$, $P = .02$). The reputation system fell the least, remaining significantly above the baseline level ($\beta = .57$, $P = .03$), the fixed partner condition ($\beta = 1.19$, $P = .01$), and the contracts condition ($\beta = 2.19$, $P < .01$). These results converge with study 1A to support the prediction that reputation systems can produce greater generalized trustworthiness than contractual enforcement (hypothesis 2) as well as relational enforcement (hypothesis 4).

Once again, a reasonable concern is whether the unenforced condition in rounds 21–23 provides valid comparisons to the treatment conditions

in phase 2. Using the first three rounds of the unenforced condition instead reveals no statistically significant difference between the reputation and unenforced conditions ($\beta = -.52$, $P = .12$). However, a comparison using all 23 rounds of the unenforced condition reveals a significant positive effect of the reputation condition ($\beta = .62$, $P = .048$). Thus, a conservative conclusion is that reputational enforcement may not necessarily produce higher levels of trustworthiness than unenforced exchanges, but it can sustain trustworthiness from falling as much as it does when contractual or relational enforcement is removed.

STUDIES 2A AND 2B: RULING OUT CONTRAST EFFECTS

In study 1A, participants in phase 2 withdrew trust in inverse proportion to the level of cooperation they displayed in phase 1. Although this pattern is consistent with the idea that risk-taking is necessary to cultivate generalized trust, it could be due to the effect of contrast in perceived risk from phase 1 to 2—the less risk clients felt in phase 1, the more risk they felt in phase 2 in contrast—rather than the direct effect of risk-taking in phase 1. Similarly, the patterns in study 1B could be due to the contrast between the levels of enforcement agents felt in phase 1 versus 2 rather than the differences in the nature of enforcement in phase 1. Study 2 was designed to address these concerns as well as to sample from a different, more urban university, five years after study 1, to demonstrate the robustness of my original findings.

Design and Procedures

The experimental design and procedures were identical to studies 1A and 1B, with two modifications. First, contractual enforcement was reduced from a 90% to a 50% enforcement rate, which reduced trust and trustworthiness to the level of reputational enforcement in a pilot study. If the patterns in study 1 are due entirely to contrast effects in perceived risk, we should observe no difference in phase 2 between contracts at 50% enforcement and the reputation system.[9] Second, I administered a postexperimental survey to further evaluate how enforcement affected participants' generalized trust and trustworthiness. One to two days after completing the online transactions, participants playing the agent's role in study 2A ($N = 62$, 34 men, 20.8 years old) responded to a six-item generalized social trust scale (Yamagishi and Yamagishi 1994). The questions asked how much participants agree that "most people are basically honest [or trustworthy,

[9] To ensure comparable levels of risk in phase 1 across conditions, the negative reputation score in round 6 (see table 2) was removed.

good and kind, trustful of others]," "I am trustful," and "most people will respond in kind when they are trusted by others." Responses were measured on seven-point scales and aggregated into a single index (Cronbach's $\alpha = .93$).

In study 2B ($N = 58$, 32 men, 20.2 years old), participants playing the client's role responded to questions adopted from the Machiavellianism scale (Mach IV; Christie and Geis 1970) to measure people's attitude toward opportunistic behavior: "It is hard to get ahead without cutting corners here and there," "Honesty is the best policy in all cases," and "Most people who get ahead in the world lead clean, moral lives."[10] Mach IV has been shown to predict untrustworthiness in investment games (Gunnthorsdottir, McCabe, and Smith 2002). Responses were reverse coded to measure "anti-Machiavellianism" (Cronbach's $\alpha = .83$). Three additional questions were added to measure reputational concerns: "One should not be overly concerned with his or her own image if one wants to be successful," "A person's reputation is not very useful in judging his or her true character," and "I care about what others think of me, even if I won't see them again." The first two questions were adopted from Yamagishi and Yamagishi (1994) and reverse coded. Responses were measured on a seven-point scale (Cronbach's $\alpha = .90$).

Results

Figure 3 plots the number of high-quality service hours commissioned (study 2A) or provided (study 2B) in each round. As intended, there was no difference between the contract and the reputation conditions in either trust ($\beta = .40$, $P = .68$) or trustworthiness ($\beta = .56$, $P = .37$) in phase 1.[11] However, the reputation condition produced greater trust than the contracts condition after removing enforcement in phase 2 (rounds 12–15; $\beta = 1.56$, $P = .028$). Similarly, the reputation condition produced greater trustworthiness than the contracts condition after removing enforcement (rounds 20–23; $\beta = 1.14$, $P = .046$). These patterns converge with the results from study 1 and rule out effects of contrast in risk between phases 1 and 2, since no differences in phase 1 were found across conditions in either study.

Results from the postexperimental questionnaire provide additional support for hypotheses 1 (reputation > contracts) and 3 (reputation > fixed relations). As figure 4*A* shows, participants in the contracts condition reported

[10] Mach IV is based on Machiavelli's *The Prince* and contains questions measuring manipulativeness and cynicism also, but those questions were precluded.

[11] To examine the effects of enforcement over a longer period of time in phase 2, phase 2 ended in round 15 instead of 14 in study 2A, and phase 1 ended in round 19 instead 20 in study 2B.

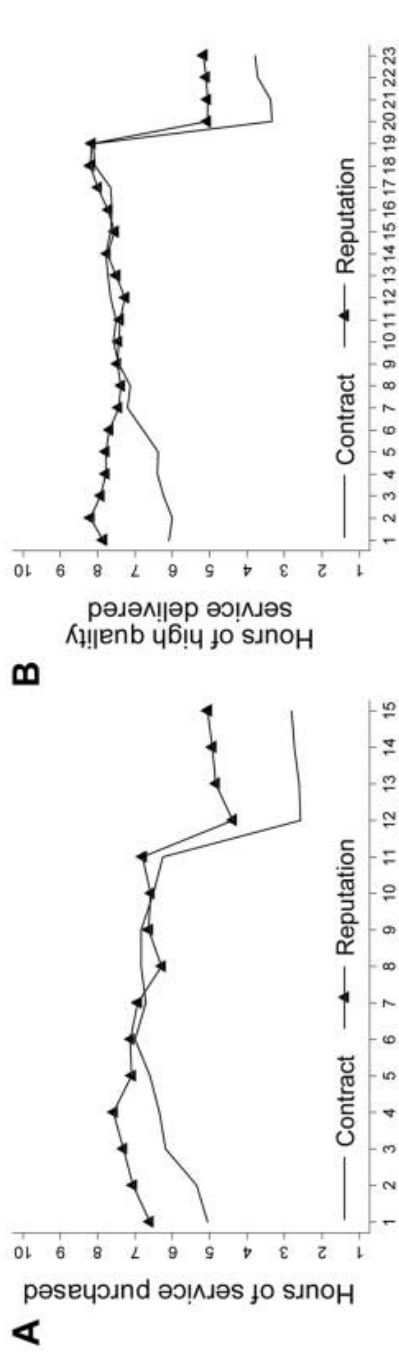
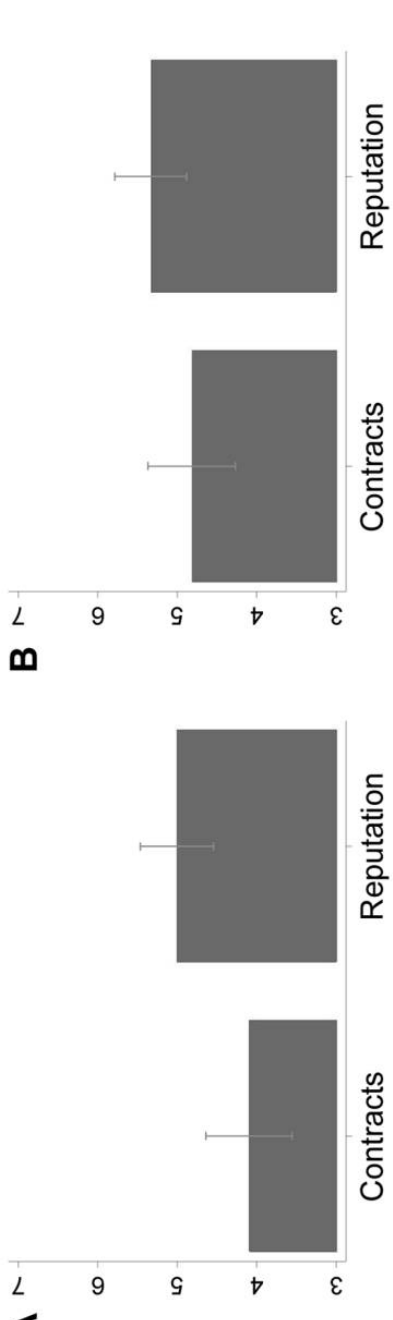


FIG. 3.—*A*, Levels of trust from study 2A; *B*, trustworthiness from study 2B. Enforcement was removed after round 11 in study 2A and round 19 in study 2B.



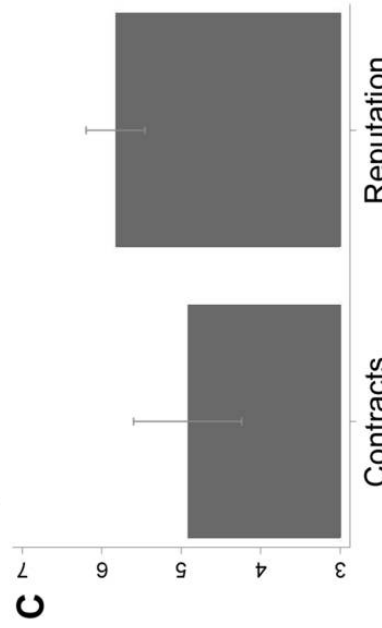


FIG. 4.—*A*, Self-reported generalized trust; *B*, attitude toward opportunism (Machiavellianism, reverse coded); and *C*, concerns about reputations from the postexperimental questionnaires in studies 2A and 2B. Likert-type seven-point scales were used for all items. Error bars indicate SDs.

lower levels of generalized trust (4.10 ± 1.48) than those in the reputation condition (5.01 ± 1.26; $t(60) = 2.61$, $P = .012$). In study 2B, participants from the reputation condition reported greater disapproval of opportunism (anti-Machiavellianism; fig. 4*B*; 5.33 ± 1.19) than those from the contracts condition (4.82 ± 1.45), although this difference did not reach significance ($t(56) = 1.49$, $P = .14$). However, those from the reputation condition reported greater reputational concerns (5.83 ± .97) than those from the contracts condition (4.92 ± 1.79; $t(56) = 2.40$, $P = .02$; fig. 4*C*). Controlling for age, race, and sex did not change the statistical significance of these results.

STUDY 3: DO EBAY USERS BECOME MORE TRUSTING OR TRUSTWORTHY?

Studies 1 and 2 found that participants show greater generalized trust and trustworthiness after even relatively brief exposure to reputation systems than to contracts in simulated transactions. These results are predicated on the idea that reputational enforcement promotes perception of trust and trustworthiness between strangers, including self-perception. However, I have not directly tested for self-perception as a possible theoretical mechanism. Study 3 was designed to address this limitation by manipulating self-perception of generalized trust and trustworthiness and to provide external validity by recruiting actual eBay users to see whether exposure to a real life feedback system promotes trust and trustworthiness.

Design and Procedures

Study 3 used a survey experiment in a 2 (role: truster vs. trustee) × 2 (self-perception: salient vs. nonsalient) design to manipulate self-perceptions of trust and trustworthiness. Participants were 475 people (31.27 years old, 41% male) recruited from Amazon Mechanical Turk (MTurk) to complete a short survey. The study was advertised as restricted to people with an active eBay account.[12]

The instruction stated that the survey consisted of two unrelated studies: (1) background questions about the respondent's sex, age, race, education level, and eBay profile, including number of positive and negative feedback

[12] MTurk is an online labor market for crowdsourcing various microtasks to human "workers" for small payments (typically a few cents). Today, it claims hundreds of thousands of active workers and is finding increasing popularity among researchers for recruiting participants for simple studies quickly and cheaply from a considerably more diverse population than is typically available on college campuses. Studies show significant convergence between results from MTurk and lab experiments (for overviews, see Buhrmester, Kwang, and Gosling 2011; Mason and Suri 2012).

points received in the last 12 months, what percentage of their feedback points was earned as a buyer rather than a seller, and the user name, and (2) an investment game with another anonymous person from MTurk to measure respondents' generalized trust or trustworthiness. For the investment game, respondents were randomly assigned to the truster or the trustee role. Those playing the role of the truster were given $5, which could be sent to an anonymous responder in increments of 50 cents. Their entrustments were tripled, which responders divided between themselves and their senders. Those playing the trustee were told that their truster had sent them $3.50 out of $5 (hence receiving $11.50 to divide). Participants received 10 cents for completing the survey. In addition, they received an additional payment with 25% probability based on randomly matching their response in the investment game with another participant's.

Feedback points were elicited to predict respondents' decisions in the investment game as a function of their exposure to eBay's reputation system. One possibility is that the more positive feedback a respondent has earned, the more often that the reputation system has presumably reinforced his self-perceived trust and trustworthiness. This idea is problematic for two reasons, however. One is selection bias: trusting or trustworthy people may be more active on eBay. Another reason is that a variety of services against fraud have been available to eBay users in recent years, including escrow services or third-party protection by credit card companies. Because these services are tantamount to contractual enforcement with binding terms, reputation scores from eBay may reflect effects of contractual rather than reputational enforcement.

The real reason for eliciting reputation scores was to experimentally manipulate the salience of self-perceived trust and trustworthiness. To this end, the survey was randomized to ask the background questions (including feedback scores from eBay) first or present the investment game first. Soliciting background information first was intended to invoke self-perceptions of trust or trustworthiness on the basis of one's own feedback score before playing the investment game. To reinforce this manipulation, respondents were asked to report their feedback score and "spend a minute to think about how you earned this score. What percentage of your feedback points did you earn as a buyer [seller]?" Such manipulations have been used in experiments on honesty (Mazar, Amir, and Ariely 2008); for instance, people are less likely to lie or misreport on insurance forms if they are asked to sign their names before, rather than after, completing the form (Shu et al. 2012). Similarly, I predicted that respondents would show greater trust and trustworthiness after reporting their positive feedback scores, controlling for negative scores.

Respondents' usernames from eBay were used to verify the reported feedback scores against their public profiles on eBay. Sixty-two people were

removed due to missing or incorrect information. Of the remaining sample, 245 people were active eBay users in the past 12 months; the others were coded as inactive. In the final data set, 36% were female, 62% were white, 32% Asian, 2.6% Hispanic, and 1.8% black; the average age was 31 ± 10.76; and 85.5% had a college degree or above. The average number of positive reputation points was 135 (SD = 698.6) and that of negative points was .49 (SD = 1.51).[13] Of the reported transactions in the last 12 months, 63.5% were as buyers.

Results

Following studies 1 and 2, I used Tobit regression to predict decisions in the investment game. Models 1 and 2 in table 3 show the results for respondents in the role of the truster. As shown, being an eBay user by itself had no effect, whether respondents reported their reputation scores first or not. However, having positive feedback points shows a positive effect ($\beta = .71$, SE = .25, $P < .01$), but only when it was reported first, before the investment game. I obtained substantively identical results for trustworthiness in models 4 and 5.

Because total positive feedback scores include points earned as a buyer or seller or both, they may conflate self-perceptions of trust and trustworthiness. To address this issue, I broke down total positive scores into positive scores by role, using self-reports of percentage of transactions as a buyer versus seller. Using these predictors instead of total positive feedback score, I found patterns suggesting self-perception. Specifically, for those in the truster role in the investment game, having positive points as a buyer was positive and significant ($\beta = .76$, SE = .31, $P = .014$), but having positive points as a seller was not ($\beta = -.10$, SE = .22, $P = .66$) when the background questions were presented first (model 3); neither variable was significant when the investment game was presented first (not shown). In mirror contrast, for those in the trustee role in the investment game, having positive points as a seller was positive and significant ($\beta = 1.46$, SE = .52, $P = .006$), but having positive points as a buyer was not ($\beta = .76$, SE = .47, $P = .11$) when the background questions were presented first (model 6); neither variable was significant when the investment game was presented first. Thus, reputations for trust and trustworthiness both affected decisions to place and honor trust in the investment game, but only when people reflected on their reputations in the relevant role.

[13] The large standard deviations are due to one person with 13,109 positive points and 212 negative points. Excluding this person did not change my main results.

TABLE 3

TOBIT REGRESSION PREDICTING TRUST AND TRUSTWORTHINESS IN AN INVESTMENT GAME, BEFORE AND AFTER REPORTING EBAY FEEDBACK SCORES

	TRUST IN IG			TRUSTWORTHINESS IN IG		
	Model 1 IG First	Model 2 Qs First	Model 3 Qs First	Model 4 IG First	Model 5 Qs First	Model 6 Qs First
Is an active eBay user	.86	−1.78	−.26	1.91	−1.96	−1.58
	(1.10)	(1.31)	(1.15)	(2.75)	(2.06)	(1.72)
Positive feedback points (logged)	.10	.71**		−.67	1.57**	
	(.26)	(.25)		(.69)	(.57)	
Negative feedback points	−.04	.41	.31	−.04	−.05	−.06
	(.21)	(.30)	(.29)	(.03)	(.03)	(.03)
Positive feedback as buyer (logged)			.86*			.76
			(.30)			(.47)
Positive feedback as seller (logged)			−.10			1.46**
			(.22)			(.52)
Constant	4.31*	1.90	4.04**	8.25**	6.21	5.57**
	(1.81)	(1.84)	(.69)	(3.40)	(3.25)	(.80)
N	116	115	115	96	86	86
Pseudo R^2	.01	.06	.04	.02	.03	.03
Log likelihood	−261.27	−251.47	−255.93	−264.59	−233.41	−232.94

NOTE.—IG = investment game; Qs = background questions. SEs are in parentheses. All models controlled for gender, age, race, and education level. All two-tailed tests.

* $P < .05$.

** $P < .01$.

DISCUSSION

Three sets of studies supported my claim that reputational enforcement can create exchange conditions that promote generalized trust and trustworthiness. Whereas binding contracts are designed to substitute for trust and trustworthiness by providing direct sanctions, reputations can complement them by reinforcing people's self-perception that they are willing to trust strangers or honor their trust, even in the absence of strong enforcement. This is in part because reputational enforcement is indirect and diffuse enough to mitigate the risks of exchange without eliminating them completely, in effect creating an intermediate condition between unenforced exchanges that are too risky to enter and contractual enforcement that makes transactions too safe yet too restrictive to nurture generalized trust or trustworthiness.

Reputational enforcement also produced greater generalized trust and trustworthiness than fixed-partner relations did. After repeated exchanges with the same partner, people were no more trusting or trustworthy than those in the unenforced condition. In contrast, people exposed to reputational enforcement were more likely to place or honor trust in unenforced transactions thereafter. Through repeated interactions, people develop trust toward particular others, but not to generalized others.

These results are predicated on the idea that different forms of governance have different effects on perceptions of generalized trust and trustworthiness, including self-perceptions. Even in one-shot, anonymous exchanges with limited opportunities to learn about each other, actors can learn about themselves and change their self-perceptions through experience and reflection, viewing themselves as more and more trusting under certain conditions and less and less trusting under other conditions. While my results are consistent with this argument, it is important to note that none of the studies presented here provide a direct measure of self-perception, only behavioral consequences of self-perception. Measuring self-perception directly is not as straightforward as measuring social perception, because self-reports of self-perception are likely to be biased in rather self-serving ways; many people are reluctant to admit that they see themselves as distrustful or untrustworthy people. Furthermore, in many situations, social and self-perception can occur concurrently as people attempt to make sense of their situations from multiple perspectives to understand themselves and each other. Thus, the current studies should not be interpreted as conclusive or exclusive tests of self-perception to rule out social perception.

Still, there are several reasons for why self-perception offers a more compelling explanation for the three studies than social perception. First, my studies focused on one-shot exchanges (except the fixed-partner conditions) with anonymous virtual partners. As research on attribution theory

has shown (Kelley 1967), such situations are more conducive to learning about oneself than about others in general or in ways that might generalize to yet unknown others, because people can observe themselves, but not specific others, across situations. Second, with respect to trustworthiness, social perception is a moot issue in investment games, because the second mover knows whether or exactly how much the first mover has decided to trust or not, and social perception is agnostic about whether people will actually refrain from cheating others. In comparison, self-perceptions of trustworthiness can directly promote acts of trustworthiness by invoking a sense of self that identifies with being trustworthy, even in situations in which cheating yields a greater immediate payoff. Thus, self-perception can better explain why participants in the reputation conditions were more trustworthy in all three studies. Third, with respect to study 3, social perception seems less likely, given the mechanics of feedback points in reputation systems like eBay. Feedback points are given to a buyer because the buyer was trusting, not because the seller was trustworthy. One's feedback points should therefore reflect more directly on oneself (self-perception) rather than on exchange partners (social perception). In this sense, it is difficult to explain why receiving positive feedback points should make people perceive others to be more trustworthy.

Empirically, the idea that reputational enforcement can promote generalized trust and trustworthiness extends two previous studies in particular. In Malhotra and Murnighan (2002), participants were exposed to several rounds of prisoner's dilemmas with an anonymous but fixed partner under binding versus nonbinding contracts. They found that, unlike binding contracts that coerce cooperation but undermine trust, nonbinding contracts can sustain relatively high levels of cooperation without diminishing intrinsic motivations to cooperate once enforcement is removed, because offering nonbinding contracts signals positive intentions to enter transactions and offer cooperation without imposing strong external constraints that eliminate risk completely. Reputational enforcement seems to work in a similar fashion, motivating cooperation without coercing it by signaling trust and trustworthiness.

Bohnet and Huck (2004) measured trust and trustworthiness in one-shot investment games after exposure to three experimental conditions: unenforced one-shot exchanges (baseline) versus unenforced exchanges with fixed partners versus one-shot exchanges under a reputation system. They found that, after removing reputations, trust and trustworthiness both dropped to the baseline level, but not higher, thus showing no reinforcement of trust or trustworthiness under reputational enforcement. One possible reason is that, in their free-play design with real human (rather than simulated) partners, cooperation under enforcement rarely exceeded

50%, possibly because their reputation system was much shorter in scope, tracking only one previous round rather than aggregating over the entire history of transactions. This suggests that reputation systems may need to be effective enough to ensure high levels of cooperation in order to also reinforce trust and trustworthiness.[14]

Theoretically, the ideas about different forms of enforcement in the current research echo recent efforts by exchange theorists to specify how different forms of exchange can invoke different attributions and perceptions that promote or inhibit the development of trust (e.g., Molm et al. 2000; Lawler 2001). In negotiated exchange, actors engage in direct bargaining over fixed resources to reach specific and binding agreements on the division of benefits. Each episode of exchange is a discrete event consisting of joint action by both parties under contractual enforcement. In comparison, reciprocal exchange generally involves sequences of unilateral giving and receiving in which actors provide favors or benefits to each other separately and without explicit bargaining. The norm of reciprocity creates diffuse obligations to return favors, but when and how to reciprocate is unspecified.

The current research builds on these ideas in several ways. First, I examined bilateral transactions, a hybrid form of exchange with features of both negotiated and reciprocal exchange. To wit, transactions were modeled using the investment game (Berg et al. 1995) in which one actor is assigned to the role of the truster and the other the trustee. Like reciprocal exchange, the investment game occurs without explicit bargaining, communication, or guarantee of repayment, creating risk of exploitation. Unlike reciprocal exchange, however, the investment game is a type of negotiated exchange characterized by a bilateral flow of benefits (Molm et al. 2000). Transactions can therefore take on characteristics of negotiated or reciprocal exchange, depending, for instance, on how they are enforced. Formal enforcement using contracts or material sanctions can make transactions more similar to negotiated exchange by ensuring reciprocity and eliminating the risk of exploitation, with direct consequences for trust and trustworthiness. Despite their relevance to markets, transactions have been overlooked in exchange theory as a special case of negotiated exchange (Kuwabara 2011).

Second, using the investment game allowed me to examine trust and trustworthiness separately. While trust has been studied extensively by exchange theorists, trustworthiness has received less attention (but see Hardin 2002; Simpson and Eriksson 2009). One reason is that trustworthiness is

[14] My study differs from Bohnet and Huck (2004) in other ways as well. First, they did not examine contractual enforcement. Second, they did not experimentally isolate trust and trustworthiness from each other. Thus, buyer behaviors may be driven by seller behaviors (or vice versa) rather than enforcement effects.

irrelevant in negotiated exchange under binding agreements that ensures full compliance, while in reciprocal exchange, trustworthiness is conflated with reciprocal acts of trust since "the same act can complete one exchange and initiate another" (Molm et al. 2000, p. 1400). However, understanding bilateral transactions requires understanding trustworthiness as a critical but discrete component of the exchange.

Third, I examined generalized rather than personalized trust and trustworthiness by considering one-shot exchanges. While exchange theory has focused almost exclusively on repeated exchange, I show that the logic of social exchange can be extended to one-shot, episodic exchange outside of fixed relations, by understanding self-perception as a process that occurs in parallel to social perception.

Fourth, the pretest-posttest design evokes the emerging work on the sequencing effects of different exchange forms (Cheshire, Gerbasi, and Cook 2010; Molm, Whitham, and Melamed 2012). Molm et al. (2012) argue that beginning an exchange relation with negotiated exchange sensitizes actors to conflictual aspects of exchange and inhibits the development of solidarity, even after they switch to reciprocal exchange, compared to relations consisting exclusively of reciprocal exchange.[15] My findings are consistent with this argument: the contracts conditions in which participants exchanged under binding terms weakened both trust and trustworthiness in unenforced transactions in phase 2, compared to both the fixed-partner or unenforced conditions. One important difference, however, is that my central mechanism is not sensitivity or contrast to conflictual aspects of negotiated exchange, as Molm et al. (2012) argue, but the lack of risk under binding agreements. Both mechanisms can reduce solidarity (Molm, Collett, and Schaefer 2007).

Finally, my results speak to emerging ideas about fragile versus resilient trust. Molm et al. (2009) show that, in fixed relations, reciprocal exchange is more likely to produce trust that is resilient to occasional mistrust, whereas negotiated exchange produces trust that is relatively fragile, because trust on the basis of binding agreement can last only so long as the terms of exchange are clearly met. Similarly, my studies show that gradual exposure to (moderate) risk under reputational enforcement is more likely than contractual enforcement to promote resilient trust that persists even in the absence of enforcement by inoculating people against risk.

It is notable that, as the current research as well as other similar studies show, relatively brief exposure to different types of enforcement in an arti-

[15] Their experiments provided support for this prediction, but only for actors in power-disadvantaged relations who may be more sensitive to the risks and conflicts in exchange relations. Although my studies did not manipulate structural power in exchange networks, one-shot exchanges may have similar effects.

ficial setting is sufficient to induce systematic behavioral changes that persisted over, in my case, periods of several hours or a couple of days. Recent studies show that changes in prosocial behavior induced by self-perception can be surprisingly persistent (Gneezy and Rustichini 2000; Grant and Dutton 2012). In real life, many people experience more sustained exposure to reputational enforcement, accruing hundreds and thousands of feedback points as repeat customers of eBay, Yelp, Amazon, and other online markets with feedback mechanisms. Still, it is unlikely that the observed changes in trust and trustworthiness are permanent because reinforced behaviors can be unlearned over time. As study 3 suggests, self-perceptions are also situational, activated in certain contexts but not others. Without deliberately reflecting on their own reputation scores, respondents were no more trusting or trustworthy regardless of their prior history on eBay.

The current experiments were designed to test the internal validity of my theoretical arguments under the most elemental conditions by using experiments to isolate theoretically and psychologically relevant features of reputation systems that distinguish them from other forms of enforcement. Although the results suffice to show that reputation systems *can* reinforce generalized trust and trustworthiness, reputation systems exist in many forms across different markets and communities. More research is needed to enrich our understanding of how and when reputation systems promote trust and trustworthiness under more varied conditions. First, how much cooperation is needed to reinforce trust and trustworthiness? In my study, transaction partners were simulated, given fixed reputations, and programmed to offer full cooperation in every round. Although this is not an entirely unrealistic portrayal of many real world markets like eBay where transactions have been overwhelmingly positive and successful (Kollock 1999; Resnick et al. 2000), fraud and mistrust do occur from time to time. How much fraud a reputation system can tolerate before it starts to collapse and erode people's trust is an increasingly critical question for better understanding the efficacy of reputation systems in light of growing concerns about cybersecurity and dubious reputation management practices (Bilton 2011).

Second, my experiments simulated very simple reputation systems with only singular numerical ratings. How my results might extend to more varied types of reputation systems with multidimensional ratings or predominantly qualitative evaluations is an important question for future research. On the one hand, it is plausible that the addition of verbal feedback or more fine-grained ratings might enhance people's experience of trust (Pavlou and Dimoka 2006). On the other hand, it is also plausible that reputation systems can become coercive enough to erode generalized trust and trustworthiness (Nissenbaum 2004). Designing an effective reputation system—or any benign institution, for that matter—is a balancing act.

CONCLUSION

Despite the substantial body of research demonstrating the efficacy of reputation systems for sustaining cooperation, hardly any research has considered their consequences for people's actions outside of such systems. Indeed, although sociologists have long been interested in various forms of governance of exchange relations, very little attention has been paid to how a particular system of enforcement might affect behaviors and decisions in other domains. This gap in our understanding of the spillover effects of enforcement systems comes to bear particularly in light of studies showing that formal enforcement using contracts or tangible sanctions can erode intrinsic motivation to cooperate. Whether reputation systems have similar consequences is a question with significant implications for research on social exchange and governance as well as for market design and social policies as online markets continue to grow in scope and scale (Nissenbaum 2004; Cook et al. 2009).

By examining different forms of governance that are prevalent in markets (reputation), hierarchies (contracts), and networks (repeated exchange), the current research sheds new light on conditions under which trust and trustworthiness that people develop in local interactions might become generalized beyond specific relations, organizations, or markets. While studies have linked generalized trust and trustworthiness to a wide range of societal benefits, from civic participation and volunteerism (Putnam 2000; Uslaner 2002) to social integration (Smith 2010; Uslaner 2012) to economic growth (Putnam 1993; Knack and Keefer 1997), identifying the mechanisms through which generalized trust and trustworthiness develop has remained rather elusive. Scholars have posited social learning processes in which trust that emerges from local interactions with familiar partners generalizes to other contexts (Putnam 2000; Glanville and Paxton 2007), although which social and institutional conditions enable such processes remain contested (Uslaner 2002; Glanville 2004). The current study points to reputational enforcement as a potentially important mechanism for the development of generalized trust and trustworthiness. Reputations create conditions of moderate risk by identifying and matching people with reliable exchange partners outside of repeated interactions. Through this process of entering and consummating exchanges with new partners on the basis of reputational enforcement, people increasingly come to perceive themselves as trusting and trustworthy.

Although such processes may be more formalized in online reputation systems, they are also ubiquitous in offline markets and communities (Greif 1989; Raub and Weesie 1990; Kollock 1994; Hillmann and Aven 2011). My research therefore points to an important role that market transactions may play in promoting generalized trust and trustworthiness. Economic

exchange is often assumed to undermine prosocial behaviors and sentiments by introducing instrumental rationality into the moral and affective fabric of social relations (Molm et al. 2000; Zelizer 2005; Ingram and Zou 2008). My results show, however, that people can develop trust and trustworthiness toward anonymous strangers on the basis of arm's length, transactional exchanges that test and affirm their sense of trust and integrity under conditions of moderate risk.

REFERENCES

Akerlof, George A. 1970. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics* 84:488–500.

Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic.

Bajari, Patrick, and Ali Hortaçsu. 2003. "The Winner's Curse, Reserve Prices and Endogenous Entry: Empirical Insights from Ebay Auctions." *Rand Journal of Economics* 3:329–55.

Bem, Daryl J. 1967. "Self-Perception: An Alternative Interpretation of Cognitive Dissonance Phenomena." *Psychological Review* 74:183–200.

Berg, Joyce, John Dickhaut, and Kevin McCabe. 1995. "Trust, Reciprocity, and Social History." *Games and Economic Behavior* 10:122–42.

Bilton, Nick. 2011. "The Growing Business of Online Reputation Management." *New York Times*, April 4.

Bohnet, Iris, Bruno S. Frey, and Steffen Huck. 2001. "More Order with Less Law: On Contract Enforcement, Trust, and Crowding." *American Political Science Review* 95:131–44.

Bolton, Gary, Ben Greiner, and Axel Ockenfels. 2012. "Engineering Trust: Reciprocity in the Production of Reputation Information." *Management Science* 59:265–85.

Bolton, Gary E., Elena Katok, and Axel Ockenfels. 2004. "How Effective Are Online Reputation Mechanisms? An Experimental Study." *Management Science* 50:1587–1602.

———. 2005. "Cooperation among Strangers with Limited Information about Reputation." *Journal of Public Economics* 89:1457–68.

Bowles, Samuel. 2008. "Policies Designed for Self-Interested Citizens May Undermine 'the Moral Sentiments': Evidence from Economic Experiments." *Science* 320:1605–9.

Buhrmester, Michael, Tracy Kwang, and Samuel D. Gosling. 2011. "Amazon's Mechanical Turk." *Perspectives on Psychological Science* 6:3–5.

Cheshire, Coye. 2011. "Online Trust, Trustworthiness, or Assurance?" *Daedalus* 140: 49–58.

Cheshire, Coye, Alexandra Gerbasi, and Karen S. Cook. 2010. "Trust and Transitions in Modes of Exchange." *Social Psychology Quarterly* 73:1–20.

Christie, Richard, and Florence L. Geis. 1970. *Studies in Machiavellianism*. New York: Academic Press.

Coleman, James S. 1990. *Foundations of Social Theory*. Cambridge, Mass.: Belknap Press of Harvard University Press.

Cook, Karen, Russell Hardin, and Margaret Levi. 2005. *Cooperation without Trust?* New York: Russell Sage.

Cook, Karen S., Coye Cheshire, Alexandra Gerbasi, and Brandy Aven. 2009. "Assessing Trustworthiness in Providers of Online Goods and Services." Pp. 189–214 in *eTrust: Forming Relationships in the Online World*, edited by Karen S. Cook, Chris Snijders, Vincent Buskens, and Coye Cheshire. New York: Russell Sage.

Cook, Karen S., and Russell Hardin. 2001. "Norms of Cooperativeness and Networks of Trust." Pp. 327–47 in *Social Norms*, edited by Michael Hechter and Karl-Dieter Opp. New York: Russell Sage.

Cook, Karen S., Chris Snijders, Vincent Buskens, and Coye Cheshire. 2009. *eTrust: Forming Relationships in the Online World*. New York: Russell Sage.

Deci, Edward L., Richard Koestner, and Richard M. Ryan. 1999. "A Meta-analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation." *Psychological Bulletin* 125:627–68.

———. 2001. "Extrinsic Rewards and Intrinsic Motivation in Education: Reconsidered Once Again." *Review of Educational Research* 71:1–27.

Dellarocas, Chrysanthos. 2003*a*. "Building Trust Online: The Design of Robust Reputation Reporting Mechanisms in Online Trading Communities." In *Information Society or Information Economy? A Combined Perspective on the Digital Era*, edited by Georgios I. Doukidis, Nikolaos Mylonopoulos, and Nancy Pouloudi. Hershey, Pa.: Idea Book.

———. 2003*b*. "The Digitization of Word-of-Mouth: Promise and Challenges of Online Feedback Mechanisms." *Management Science* 49:1407–24.

———. 2010. "Online Reputation Systems: How to Design One That Does What You Want." *MIT Sloan Management Review* 51 (3).

Diekmann, Andreas, Ben Jann, Wojtek Przepiorka, and Stefan Wehrli. 2014. "Reputation Formation and the Evolution of Cooperation in Anonymous Online Markets." *American Sociological Review* 79:65–85.

Diekmann, Andreas, Ben Jann, and David Wyder. 2009. "Trust and Reputation in Internet Auctions." Pp. 139–65 in *eTrust: Forming Relationships in the Online World*, edited by Karen S. Cook, Chris Snijders, Vincent Buskens, and Coye Cheshire. New York: Russell Sage.

Durkheim, Émile. 1893. *The Division of Labor in Society*. New York: Free Press.

Fehr, Ernst, and Armin Falk. 2002. "Psychological Foundation of Incentives." *European Economic Review* 46:687–724.

Fehr, Ernst, and Bettina Rockenbach. 2003. "Detrimental Effects of Sanctions on Human Altruism." *Nature* 422:137–40.

Frank, Robert H. 1988. *Passions within Reason: The Strategic Role of the Emotions*. New York: Norton.

Fukuyama, Francis. 1995. *Trust: Social Virtues and the Creation of Prosperity*. New York: Free Press.

Glaeser, Edward L., David I. Laibson, Jos A. Scheinkman, and Christine L. Soutter. 2000. "Measuring Trust." *Quarterly Journal of Economics* 115:811–46.

Glanville, Jennifer L. 2004. "Voluntary Associations and Social Network Structure: Why Organizational Location and Type Are Important." *Sociological Forum* 19: 465–91.

Glanville, Jennifer L., and Pamela Paxton. 2007. "How Do We Learn to Trust? A Confirmatory Tetrad Analysis of the Sources of Generalized Trust." *Social Psychology Quarterly* 70:230–42.

Gneezy, Uri, Stephan Meier, and Pedro Rey-Biel. 2011. "When and Why Incentives (Don't) Work to Modify Behavior." *Journal of Economic Perspectives* 25:191–209.

Gneezy, Uri, and Aldo Rustichini. 2000. "Pay Enough or Don't Pay at All." *Quarterly Journal of Economics* 115:791–810.

Gouldner, Alvin W. 1960. "The Norm of Reciprocity: A Preliminary Statement." *American Sociological Review* 25:161–78.

Granovetter, Mark. 1985. "Economic Action and Social Structure: The Problem of Embeddedness." *American Journal of Sociology* 91:481–510.

Grant, Adam, and Jane Dutton. 2012. "Beneficiary or Benefactor." *Psychological Science* 23:1033–39.

Greif, Avner. 1989. "Reputation and Coalitions in Medieval Trade: Evidence on the Maghribi Traders." *Journal of Economic History* 49:857–82.
Greif, Avner, Margaret Levi, Jean-Laurent Rosenthal, and Barry R. Weingast. 1998. "Self-Enforcing Political Systems and Economic Growth: Late Medieval Genoa." Pp. 23–63 in *Analytic Narratives*, edited by R. H. Bates. Princeton, N.J.: Princeton University Press.
Gunnthorsdottir, Anna, Kevin McCabe, and Vernon Smith. 2002. "Using the Machiavellianism Instrument to Predict Trustworthiness in a Bargaining Game." *Journal of Economic Psychology* 23:49–66.
Hardin, Russell. 2002. *Trust and Trustworthiness*. New York: Russell Sage.
Hayashi, Nahoko, and Toshio Yamagishi. 1998. "Selective Play: Choosing Partners in an Uncertain World." *Personality and Social Psychology Review* 2:276–89.
Hillmann, Henning, and Brandy L. Aven. 2011. "Fragmented Networks and Entrepreneurship in Late Imperial Russia." *American Journal of Sociology* 117:484–38.
Homans, George C. 1961. *Social Behavior: Its Elementary Form*. New York: Harcourt Brace.
Ingram, Paul, and Xi Zou. 2008. "Business Friendships." *Research in Organizational Behavior* 28:167–84.
Kelley, Harold H. 1967. "Attribution Theory in Social Psychology." Pp. 192–238 in *Nebraska Symposium on Motivation*, edited by D. Levine. Lincoln: University of Nebraska.
Keser, Claudia. 2003. "Experimental Games for the Design of Reputation Management Systems." *IBM Systems Journal* 42:498–506.
Knack, Stephen, and Philip Keefer. 1997. "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation." *Quarterly Journal of Economics* 112:1251–88.
Kollock, Peter. 1994. "The Emergence of Exchange Structures: An Experimental Study of Uncertainty, Commitment, and Trust." *American Journal of Sociology* 100:313–45.
———. 1999. "The Production of Trust in Online Markets." *Advances in Group Processes* 16:99–123.
Kosugi, Motoko, and Toshio Yamagishi. 1998. "Generalized Trust and Judgments of Trustworthiness." *Japanese Journal of Psychology* 69:349–57.
Kramer, Roderick. 1999. "Trust and Distrust in Organizations: Emerging Perspectives, Enduring Questions." *Annual Review in Psychology* 50:569–98.
Kramer, Roderick M., and Tom R. Tyler. 1996. *Trust in Organizations: Frontiers of Theory and Research*. Thousand Oaks, Calif.: Sage.
Kuwabara, Ko. "Cohesion, Cooperation, and the Value of Doing Things Together." American Sociological Review 76:560–80.
La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert W. Vishny. 1997. "Trust in Large Organizations." *American Economic Review* 87:333–38.
Lawler, Edward J. 2001. "An Affect Theory of Social Exchange." *American Journal of Sociology* 107:321–52.
Lepper, Mark R., David Greene, and Richard E. Nisbett. 1973. "Undermining Children's Intrinsic Interest with Extrinsic Reward: A Test of the 'Overjustification' Hypothesis." *Journal of Personality and Social Psychology* 28:129–37.
Luhmann, Niklas. 1979. *Trust and Power*. London: Wiley.
Macy, Michael W., and Yoshimichi Sato. 2002. "Trust, Cooperation, and Market Formation in the U.S. and Japan." *Proceedings of the National Academy of Sciences* 99:7214–20.
Macy, Michael W., and John Skvoretz. 1998. "The Evolution of Trust and Cooperation between Strangers: A Computational Model." *American Sociological Review* 63:638–60.
Malhotra, Deepak, and J. Keith Murnighan. 2002. "The Effects of Contracts on Interpersonal Trust." *Administrative Science Quarterly* 47:534–59.

Mason, Winter, and Siddharth Suri. 2012. "Conducting Behavioral Research on Amazon's Mechanical Turk." *Behavior Research Methods* 44 (1): 1–23.
Mazar, Nina, On Amir, and Dan Ariely. 2008. "The Dishonesty of Honest People: A Theory of Self-Concept Maintenance." *Journal of Marketing Research* 45:633–44.
Molm, Linda D., Jessica L. Collett, and David R. Schaefer. 2007. "Building Solidarity through Generalized Exchange: A Theory of Reciprocity." *American Journal of Sociology* 113:205–42.
Molm, Linda D., Nobuyuki Takahashi, and Gretchen Peterson. 2000. "Risk and Trust in Social Exchange: An Experimental Test of a Classical Proposition." *American Journal of Sociology* 105:1296–1427.
Molm, Linda D., Monica M. Whitham, and David Melamed. 2012. "Forms of Exchange and Integrative Bonds: Effects of History and Embeddedness." *American Sociological Review* 77:141–65.
Mulder, Laetitia B., Eric van Dijk, David De Cremer, and Henk A. M. Wilke. 2006. "Undermining Trust and Cooperation: The Paradox of Sanctioning Systems in Social Dilemmas." *Journal of Experimental Social Psychology* 42:147–62.
Nissenbaum, Helen. 2004. "Will Security Enhance Trust Online, or Supplant It." Pp. 155–88 in *Trust and Distrust in Organizations: Dilemmas and Approaches*, edited by Roderick M. Kramer and Karen S. Cook. New York: Russell Sage.
Pavlou, Paul A., and Angelika Dimoka. 2006. "The Nature and Role of Feedback Text Comments in Online Marketplaces: Implications for Trust Building, Price Premiums, and Seller Differentiation." *Information Systems Research* 17:392–414.
Pavlou, Paul A., and David Gefen. 2004. "Building Effective Online Marketplaces with Institution-Based Trust." *Information Systems Research* 15:37–59.
Paxton, Pamela. 2007. "Association Memberships and Generalized Trust: A Multilevel Model across 31 Countries." *Social Forces* 86:47–76.
Powell, Walter W. 2003. "Neither Market nor Hierarchy." Pp. 104–17 in *The Sociology of Organizations: Classic, Contemporary, and Critical Readings*, edited by Michael J. Handel. Thousand Oaks, Calif.: Sage.
Puranam, Phanish, and Bart S. Vanneste. 2009. "Trust and Governance: Untangling a Tangled Web." *Academy of Management Review* 34:11–31.
Putnam, Robert D. 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, N.J.: Princeton University Press.
———. 2000. *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon & Schuster.
Raub, Werner, and Jeroen Weesie. 1990. "Reputation and Efficiency in Social Interactions: An Example of Network Effects." *American Journal of Sociology* 96:626–54.
Resnick, Paul, Richard Zeckhauser, Eric Friedman, and Ko Kuwabara. 2000. "Reputation Systems." *Communications of the ACM* 43:45–48.
Resnick, Paul, Richard Zeckhauser, John Swanson, and Kate Lockwood. 2006. "The Value of Reputation on Ebay: A Controlled Experiment." *Experimental Economics* 9:79–101.
Rotter, Julian B. 1971. "Generalized Expectancies for Interpersonal Trust." *American Psychologist* 26:443.
Shu, Lisa L., Nina Mazar, Francesca Gino, Dan Ariely, and Max H. Bazerman. 2012. "Signing at the Beginning Makes Ethics Salient and Decreases Dishonest Self-Reports in Comparison to Signing at the End." *Proceedings of the National Academy of Sciences* 109:15197–200.
Simpson, Brent, and Kimmo Eriksson. 2009. "The Dynamics of Contracts and Generalized Trustworthiness." *Rationality and Society* 21:59–80.
Skinner, B. F. 1969. *Contingencies of Reinforcement: A Theoretical Analysis*. New York: Appleton-Century-Crofts.
Smith, Sandra Susan. 2010. "Race and Trust." *Annual Review in Sociology* 36:453–75.

Spence, A. Michael. 1974. *Market Signaling: Informational Transfer in Hiring and Related Processes*. Cambridge, Mass.: Harvard University Press.
Stolle, Dietlind. 2001. "Clubs and Congregations: The Benefits of Joining an Association." Pp. 202–44 in *Trust in Society*, edited by K. S. Cook. New York: Russell Sage.
Tenbrunsel, Ann E., and David M. Messick. 1999. "Sanctioning Systems, Decision Frames, and Cooperation." *Administrative Science Quarterly* 44:684–707.
Uslaner, Eric M. 2002. *The Moral Foundations of Trust*. Cambridge: Cambridge University Press.
———. 2012. *Segregation and Mistrust: Diversity, Isolation, and Social Cohesion*. Cambridge: Cambridge University Press.
Willer, Robb. 2009. "Groups Reward Individual Sacrifice: The Status Solution to the Collective Action Problem." *American Sociological Review* 74:23–43.
Williamson, Oliver E. 1981. "The Economics of Organization: The Transaction Cost Approach." *American Journal of Sociology* 87:548–77.
Wilson, Rick K. 1997. "'Liar, Liar': Cheap Talk and Reputation in Repeated Public Goods Settings." *Journal of Conflict Resolution* 41:695–717.
Yamagishi, Toshio. 1998. *The Structure of Trust: The Evolutionary Games of Mind and Society*. Tokyo: University of Tokyo Press.
Yamagishi, Toshio, Karen S. Cook, and Motoki Watabe. 1998. "Uncertainty, Trust and Commitment Formation in the United States and Japan." *American Journal of Sociology* 108:165–94.
Yamagishi, Toshio, Masafumi Matsuda, Noriaki Yoshikai, Hiroyuki Takahashi, and Yukihiro Usui. 2009. "Solving the Lemons Problem with Reputation." Pp. 73–109 in *eTrust: Forming Relationships in the Online World*, edited by Karen S. Cook, Chris Snijders, Vincent Buskens, and Coye Cheshire. New York: Russell Sage.
Yamagishi, Toshio, and Midori Yamagishi. 1994. "Trust and Commitment in the United States and Japan." *Motivation and Emotion* 18:9–66.
Zelizer, Viviana A. Rotman. 2005. *The Purchase of Intimacy*. Princeton, N.J.: Princeton University Press.