

How Monitoring Influences Trust: A Tale of Two Faces

Maurice E. Schweitzer
Wharton School
University of Pennsylvania
Philadelphia, PA 19104
E-mail: Schweitzer@wharton.upenn.edu

Teck-Hua Ho
NUS Business School
National University of Singapore
Singapore, 119245
&
Haas School of Business
University of California at Berkeley
Berkeley, CA 94720
E-mail: dprhoth@nus.edu.sg

Xing Zhang
Singapore – ETH Center
Swiss Federal Institute of Technology in Zurich
Singapore, 138602
Email: zhangxingis@gmail.com

ABSTRACT

Monitoring changes the behavior of those who are monitored and those who monitor others. We studied behavior under different monitoring regimes in repeated trust games. We found that Trustees behaved opportunistically when they anticipated monitoring—they were compliant when they knew in advance that they would be monitored, but exploited Trustors when they knew in advance that they would not be monitored. Interestingly, Trustors failed to anticipate how strategically their counterparts would behave. Trustors misattributed the strategic, compliant behavior they observed as signals of Trustees' trustworthiness. As a result, Trustors misplaced their trust when they were unable to monitor their counterparts. Our findings stand in contrast to prior work that has found that observing compliant behavior can harm trust. We discuss the managerial implications of our results for designing and implementing monitoring systems.

INTRODUCTION

Trust plays a crucial role in individual relationships and within organizations (Lewicki, Tomlinson and Gillespie 2007). Trust promotes cooperation (Pillutla, Malhotra and Murnighan 2003) and broadly impacts workplace behavior (Dirks and Ferrin 2001; Kim, Dirks, Cooper and Ferrin 2006). Managers who develop trusting relationships with their employees are more effective.

Promoting trust and trustworthy behavior, however, represents a constant managerial challenge (Ho and Weigelt 2005; Kim, Ferrin, Cooper, and Dirks 2004; Kim, Dirks, Cooper, and Ferrin 2006; Kramer and Lewicki 2010; Yip & Schweitzer, 2015). Managers need to trust their employees to use organizational resources to accomplish workplace goals, but their trust is often misplaced. A typical organization loses about 5% of its revenue from occupational fraud, which includes employee theft and inflated expense reporting (Association of Certified Fraud Examiners 2012).

One tool that managers use to promote compliant behavior is employee monitoring. In an American Management Association survey (2014), 66% of managers reported that they monitor the internet use of their employees. Almost half of the companies surveyed reported that they use video monitoring, and nearly half (43%) monitor e-mail use or use other detailed tracking methods, such as tracking their employees' keystrokes (45%). Monitoring systems, however, are expensive and their effectiveness is limited by a scarce resource: managerial attention. As a result, managers often observe only a limited sample of employee behavior. For example, supervisors in most call centers listen to only a small fraction of the calls handled by the operators they oversee.

Monitoring systems have the potential to change behavior and perceptions, but we know surprisingly little about the consequences of monitoring systems on trust. In this article, we investigate the consequences of monitoring, and we explore how monitoring influences compliance, trust, and trustworthy behavior. Across four studies, we show that people behave strategically when they are monitored, but that observers fail to anticipate this. Instead, observers mistake strategic behavior for trustworthy behavior.

LITERATURE REVIEW

Surveillance and monitoring systems pervade the workplace (American Management Association survey, 2014), and an emerging literature has begun to explore how monitoring systems influence both perceptions and behavior (Kramer 1999). For example, prior research has found that monitoring systems can harm intrinsic motivation and communicate negative expectations (Cialdini 1996; Enzle and Anderson 1993). That is, by monitoring employees, managers may signal that they expect unmonitored employees to shirk their responsibilities (Frey 1993). As a consequence, the use of monitoring systems can signal distrust, subverting the very behaviors managers are hoping to encourage.

In some cases, however, monitoring can yield substantial benefits. Effective monitoring can not only promote desirable behaviors, but also can project perceptions of fairness (Niehoff and Moorman 1993). In addition, monitoring systems can increase trustworthy behavior (Schweitzer and Ho 2005).

We investigate compliant and trustworthy behavior within the context of dyadic relationships. We define compliance as actions one party takes that fulfill the desires of another party when those actions are observable and subject to repercussions (e.g., subject to punishment for noncompliant behavior). We define trustworthiness as actions one party takes that fulfill the desires of another party when those actions are *not* observable and hence not subject to repercussions. In this work, we devote particular attention to the relationships between trust and compliance as well as trust and trustworthiness.

We build on prior work to define trust as the willingness of one party to accept vulnerability based upon the positive expectations of another party (Mayer, Davis and Schoorman 1995). Within a dyadic relationship, a trustor decides to either trust or not trust a trustee, who subsequently acts in either a trustworthy or opportunistic way. Although several studies have explored how trust changes over time (Kim, Ferrin, Cooper, and Dirks 2004; Schweitzer, Hershey, and Bradlow 2006; Haselhuhn, Schweitzer & Wood, 2010), no prior work has explored how monitoring influences compliant, trustworthy, and trusting behavior over time.

Trust

Trust is remarkably labile (Kramer 1999; Yip and Schweitzer, 2015; Brooks, Dai and Schweitzer, 2014; Malhotra 2004; Pillutla, Malhotra and Murnighan 2003; Haselhuhn, Schweitzer and Wood 2010). Incentives (Lewicki, Tomlinson, and Gillespie 2006), reciprocity (Pillutla, Malhotra and Murnighan 2003), and even emotions (Dunn and Schweitzer 2005; Lount 2010) can profoundly influence how much individuals trust each other.

Trust is particularly influenced by repeated interactions. Repeated, positive interactions promote trust (Shapiro, Sheppard, and Cheraskin, 1992), but deception and violated trust harm trust (Schweitzer et al., 2006). The relationship between observed behavior and trust, however, is strongly moderated by context (Gambetta 1988; Levine and Schweitzer, 2015). In some cases, positive interactions fail to build trust. For example, Malhotra and Murnighan (2002) found that contracts impede trust development. When individuals acted in a compliant way after agreeing to a contract, observers attributed their positive behavior to the contract. That is, rather than making a personal attribution for the behavior they observed, observers made situational attributions and assumed that individuals had complied *because* of the contract.

One factor that may be particularly likely to influence compliance and trust is monitoring. If monitoring systems are accompanied by the prospect of retaliation for non-compliance, monitoring systems may impede trust development in much the same way that contracts do. If trustors make situational rather

than personal attributions for the compliant behavior they observe when they monitor trustees, monitoring systems may fail to build trust. In fact, this mechanism may explain Strickland's (1958) findings. Strickland asked participants to assume the role of a supervisor and evaluate two fictitious subordinates. Both subordinates completed similar tasks, but one subordinate was monitored more than the other. Strickland (1958) found that participants trusted the *infrequently* monitored subordinate more than the frequently monitored one.

We build on Strickland's (1958) findings, and explore the interplay between monitoring systems, behavior, and perceptions. Our work departs from Strickland's (1958) investigation in five important respects. First, rather than asking participants to compare two competing employees, we explore how interpersonal trust develops in dyadic relationships. Second, rather than exploring fictitious outcomes, we study actual behavior with financial incentives. Third, in our investigation we study the behavior of both trustors and the trustees. By studying both trustor and trustee behavior, we can test whether or not trustees engage in strategic behavior and whether or not trustors anticipate strategic trustee behavior. Fourth, the nature of the monitoring we study is quite different from Strickland's investigation. For example, in our studies we exogenously imposed different types of monitoring regimes. Fifth, we study behavior within a repeated trust game with a stochastic ending rule. Taken together, our approach enables us to measure changes in trusting and trustworthy behavior over time.

Like our investigation, both Malhotra and Murnighan (2002) and Strickland (1958) investigated the relationship between compliant behavior and trust. Importantly, both Malhotra and Murnighan (2002) and Strickland (1958) found that observing or monitoring compliant behavior can impede trust development. Malhotra and Murnighan (2002) found that contracts curtailed trust development; when individuals observed compliant behavior, they attributed that behavior to the contract. Strickland (1958) found that when two employees achieved similar outcomes, participants placed greater trust in the one who had been monitored less. In contrast to these findings, we find that trustors over-rely on the compliant behavior they observe. We show that observers fail to appreciate how strategically their counterparts will act when they anticipate that they will be monitored, and as a consequence, systematically misplace their trust. Notably, in contrast to these prior studies, we find that monitoring can boost trust over time.

Trust Game

A substantial literature has used a modified version of Berg, Dickhaut, and McCabe's (1995) trust game to study trust behavior. In this game, an individual (the trustor) starts with an endowment (e.g., \$10) and can either keep the money or pass a portion of it to their partner (the trustee). If the trustor passes \$ x to their partner, the amount of money grows (e.g., triples to \$ $3x$), and the trustee must then decide how to split the resulting sum of money between the trustor and themselves (e.g., by splitting it evenly, \$ $1.5x$ each, or by

keeping the entire \$3x). In a trust game with an initial endowment of \$10 that tripled, Berg et al. (1995) found that trustors passed an average of \$5.16 (out of their initial sum of \$10), and trustees returned an average of \$4.66 to trustors. In this case, on average, trustors who trusted their counterparts and passed money were worse off than those who did not.

In our study, participants play a series of multi-round trust games with the same partner. Repeated trust games have been used to study trust building (Pillutla, Malhotra, and Murnighan 2003; Ho and Weigelt, 2005), trust repair (Schweitzer, Hershey, and Bradlow 2006; Kim, Dirks, Cooper, and Ferrin 2006; Lount, Zhong, Sivanathan and Murnighan 2008), and trust erosion (Haselhuh, et al., 2015). In our studies, we use a repeated trust game to study the influence of monitoring. Trustors in our studies only receive partial feedback about their trustee's behavior. In *monitored* rounds, trustors learn how much their trustee counterparts returned in that particular round; but in *non-monitored* rounds, trustors do not learn information about their trustee counterparts' decisions in that round, until the very end of the experiment. This design enables us to explore the influence of monitoring on trust and trustworthy behavior.

Monitoring Schemes

We consider two dimensions of monitoring schemes: 1) frequency and 2) anticipation. The frequency of monitoring reflects how often trustees are monitored. The anticipation of monitoring reflects whether or not trustees know in advance whether or not they will be monitored in a specific round. That is, anticipation removes trustees' uncertainty with respect to being monitored in a particular round and enables them to behave opportunistically. We focus attention on anticipated monitoring, because it affords trustees the opportunity to act strategically and we can explore how trustees and trustors react to this opportunity. In practice, many monitoring schemes let employees know in advance when they will be monitored (e.g., announced visits or advanced warning of monitoring). Our investigation affords insight into the benefits and costs of anticipated monitoring.

HYPOTHESES

We develop two sets of hypotheses. The first set of hypotheses considers how monitoring influences trustee behavior. Within the trust game, trustees can exhibit cooperative behavior by returning more money than they received, or they can exhibit non-cooperative behavior by returning less money than they received. When they are monitored, trustees may exhibit cooperative behavior either because they are intrinsically trustworthy or because they are keen to manage impressions. By managing impressions, trustees may convince their counterparts to continue to pass them money in non-monitored rounds. In anticipated monitoring rounds, both motivations will promote compliance. In anticipated non-monitored rounds, trustees who are intrinsically trustworthy will exhibit cooperative behavior; trustees who are not trustworthy will

exhibit competitive behavior. In our study, we use the term ‘compliant behavior’ to refer to cooperative behavior in monitored rounds, and we use the term ‘trustworthy behavior’ to refer to cooperative behavior in non-monitored rounds.

We expect trustees to be more likely to return money in rounds in which they anticipate monitoring than in rounds in which they do not. The difference in the incidence of cooperative behavior between the anticipated monitored and anticipated non-monitored rounds enables us to assess the extent of strategic behavior.

In unanticipated monitoring conditions, trustees do not know before they make a decision whether or not their behavior will be monitored. Hence we do not expect to see a difference in cooperative behavior between monitored and non-monitored rounds.

Our second set of hypotheses considers how monitoring influences the decision to trust others. These hypotheses focus on how monitoring changes trustor behavior both when trustors are able to and when they are unable to monitor their counterparts. Trustors’ willingness to trust is measured by whether or not they pass a positive sum of money to their counterparts.

Trustee Behavior

In each round, trustees face the decision to act cooperatively or non-cooperatively, as expressed by the decision to return more or less money than they receive. The rewards for cooperative behavior are greatest when the behavior is observable. Although trustees can reap short-term gains by returning less money in monitored rounds, trustors may punish non-cooperative behavior by not passing money in future rounds. As a consequence, we expect cooperative behavior to be greatest in anticipated monitored rounds and least in anticipated non-monitored rounds.

A substantial literature suggests that monitoring is likely to change how people behave within organizations. Across many social settings, people work hard to create positive impressions (Leary 1996), and individuals pay closer attention to how they behave when others can observe them than when others cannot (Goffman 1959). People also strategically control the information they reveal and the emotions they express to influence others (Goffman 1959; Andrade and Ho, 2007; Andrade and Ho 2009). Frequently, people work to present themselves in positive ways to get others to like them (Schlenker and Leary 1982). Taken together, we expect that when trustees anticipate monitoring, they will behave differently than when they anticipate no monitoring.

Hypothesis 1: *Trustees will exhibit cooperative behavior more frequently in anticipated monitored rounds than in anticipated non-monitored rounds.*

Violations damage trust, and within a repeated trust game trustors will stop passing money when they learn about the actions of an untrustworthy trustee (Ho and Weigelt 2005; Lewicki, Tomlinson, and

Gillespie 2006; Schweitzer et al., 2005). Frequent monitoring affords trustors will many opportunities to observe trustee behavior. Trustees who anticipate the harmful consequences of displaying untrustworthy behavior will be less likely to engage in untrustworthy behavior when monitoring is frequent than when it is infrequent. As a result, we predict that both trustees and trustors will be more trusting and trustworthy when monitoring is more frequent than when it is less frequent:

Hypothesis 2: *Frequent monitoring increases cooperative behavior across anticipated and unanticipated monitoring conditions.*

Trustor Behavior

Trustees are likely to be particularly compliant when they anticipate monitoring. However, when trustees anticipate that they will not be monitored, they may act opportunistically. If trustors anticipate this, they should not pass money to trustees in anticipated non-monitored rounds. In this study, we use the term ‘trusting behavior’ to refer to passing behavior in non-monitored rounds. We explore the thesis that trustors fail to appreciate how strategically trustees will behave. Instead of engaging in perspective taking, trustors may focus on the behavior they observe (Lewicki, Tomlinson, and Gillespie 2006) and rely on available information (e.g., commit the availability bias; Tversky and Kahneman 1974). We postulate that this will be true even when the monitored behavior is not representative of non-monitored behavior. That is, *trustors may misattribute the compliant behavior they observe in monitored rounds to dispositional attributes* (e.g., infer that the trustee is trustworthy; Gilbert and Malone 1995). In particular, we expect trustors to over-rely on the biased sample of behavior they observe in monitored rounds. As a result, trustors may make overly-optimistic inferences of trustworthiness and over-predict trustworthy behavior in non-monitored rounds. That is, we hypothesize that many trustors will pass money in anticipated non-monitored rounds, and as trustors observe compliant behavior in monitored rounds, they will become even more likely to trust and pass money in anticipated non-monitored rounds.

Hypothesis 3a: *Trustors will be less responsive than trustees to the presence and absence of anticipated monitoring.*

Hypothesis 3b: *The more compliance trustors observe in monitored rounds, the more trusting they will become in anticipated non-monitored rounds.*

STUDY 1: MONITORING AND TRUST

In Study 1, we investigate the relationship between monitoring and trust. We randomly and anonymously paired participants in each session. We assigned each participant in a dyad to the role of either the “odd player” or the “even player.” (In our experiment, we used the terms “odd” and “even” instead of

“trustor” and “trustee,” respectively.) We assigned each dyad to a monitoring condition, and we had them play multiple rounds of the trust game with the same partner.

In each round, the odd player started with an endowment of five points. The odd player decided how many of the five points to keep for themselves and how many to pass to their even player counterpart. The even player received a multiple of the amount the odd player passed. In Study 1, the amount the odd player passed was quadrupled. For example, if an odd player passed 2 points (and kept 3 points) then the even player received 8 points (i.e., $4 \times 2 = 8$ points). The even player then decided how many to keep and how many points to return to the odd player. For example, if the odd player passed 2 points (and the even player received 8 points) and the even player returned 4 points, the odd player would earn 7 points (i.e., $3 + 4$) and the even player would earn 4 points (i.e., $8 - 4$). The round ended after the even player made a decision. The same game was repeated multiple times.

Methods

Participants. We recruited 182 participants from a large northeastern university in the U.S. to participate in a one-hour laboratory study for a \$10 show-up fee and the chance to earn additional money. We recruited participants through an online campus subject pool. Participants completed the study in groups that ranged in size from 6 to 14. The average group size was 8. (If an odd number of participants showed up, one participant was randomly selected, removed from the group, paid \$10, and dismissed.)

We told participants that they could earn additional money in the experiment. We explained that the total amount that they earned would depend upon the decisions they made, the decisions their partners made, and upon chance. We told participants that they would earn “points,” and that each point was worth 10 cents.

Design. We manipulated the nature of the feedback players received. Although both players made decisions in every round, they did not always learn what their counterparts chose in a particular round. We randomly assigned pairs of participants to one of four treatment conditions from a 2 (frequency: high vs. low) x 2 (anticipation: anticipated vs. unanticipated) design.

We randomly assigned participants to either a *high-frequency* monitoring condition or a *low-frequency* monitoring condition. Within the experiment, we called a monitored round a feedback round. The chance that any one round would be a feedback round was constant within a condition and independent of whether or not the prior round was a feedback round. That is, we used an *independent* draw to determine whether any given round was a feedback round. In the high-frequency condition, there was an 80% chance that odd players would receive feedback in each round (and learn what their even player counterparts chose for that particular round). In the low-frequency monitoring condition, there was a 40% chance that odd players would receive feedback in each round (and learn what their even player counterparts chose for that particular round). In rounds without monitoring, odd players did not learn what their even player counterparts

chose for that particular round until the experiment ended. In every round, across all conditions, even players always learned what their odd player counterparts passed to them before they made their decisions.

To control for end-game effects, the total number of rounds participants played was randomly determined. Every dyad made 10 rounds of decisions. After the tenth round, participants had an 85% chance of continuing to the eleventh round. If there was an eleventh round, there was an 85% chance of advancing to the twelfth round, and so on. The 85% chance of advancing to the next round was independent of the number of rounds participants had advanced.

We also randomly assigned participants to either an *anticipated* monitoring condition or an *unanticipated* monitoring condition. This condition varied whether or not even players knew before they made their decisions if their current round was a feedback round. In the anticipated condition, even players knew before they made their decisions whether or not their current round was a feedback round (in which case the odd player would learn what they, the even player counterpart, chose). In the unanticipated condition, even players did not know in advance whether or not their current round was a feedback round. Odd players always knew before they made their decisions whether or not the upcoming round was a feedback round. They also knew what their even player counterpart knew about the current round.

As a result, each pair of participants belonged to one of the four monitoring conditions from our 2 (frequency) x 2 (anticipation) between-subject design: high-frequency anticipated, low-frequency anticipated, high-frequency unanticipated, and low-frequency unanticipated. Both odd and even players knew details about their own treatment conditions, but were unaware of the other three treatment conditions. Within each dyad, participants made decisions in monitored and non-monitored rounds. Monitoring is a within-subject factor. We measured both how much the odd player (the trustor) passed and how much the even player (the trustee) returned in each round.

Procedure. After reading the instructions, participants answered a series of comprehension check questions about the nature of the game, their experimental condition, and the payment scheme. The software program returned participants to the instruction page if they made a mistake answering any of the comprehension check questions. Most participants understood both the game instructions and their experimental condition. However, 13 participants made repeated mistakes answering the comprehension check questions. We dismissed these participants and their partners. In total, we dismissed 11 dyads (in two dyads, both pair members made repeated mistakes answering the comprehension check questions). As a result, a total of 160 participants completed the study.

Each pair of participants played the trust game multiple times. In each round, participants remained in the same role and were matched with the same partner. We summed the points players earned for each round and converted their total points into cash. We paid participants in cash before they left the experiment.

Results

A total of 160 participants (80 dyads) completed the study. We randomly and independently assigned each dyad to one of four treatment conditions; a total of 42 participants completed the study in the low-frequency anticipated condition, 44 in the high-frequency anticipated condition, 34 in the low-frequency unanticipated condition, and 40 in the high-frequency unanticipated condition.

In Figure 1a, we depict trustors' passing behavior across the four treatment conditions. Overall, trustors passed in 83% of the rounds in the low-frequency anticipated condition, 97% in the high-frequency anticipated condition, 94% in the low-frequency unanticipated condition, and 80% in the high-frequency unanticipated condition.

In Figure 1b, we depict trustees' cooperative behavior (i.e., when trustees returned more money than they received) across the four treatment conditions. Trustees returned more than they received in 61% of the rounds in the low-frequency anticipated condition, 75% in the high-frequency anticipated condition, 71% in the low-frequency unanticipated condition, and 83% in the high-frequency unanticipated condition. The most striking empirical regularity in the data is that trustees cooperated significantly more frequently in monitored rounds than in non-monitored rounds in the anticipated conditions. Specifically, across the two anticipated monitoring conditions, trustees complied in 88% of the anticipated monitored rounds, but cooperated in only 35% of the anticipated non-monitored rounds (see the left panel in Figure 1b). We next formally test our hypotheses.

Insert Figures 1a-1c about here

Trustee behavior: For each dyad, we have multiple observations in both monitored and non-monitored rounds. For each trustee, we computed two cooperation measures: one measure for monitored rounds and one for non-monitored rounds. Our statistical tests are based on a 2 (frequency) x 2 (anticipation) x 2 (monitoring) ANOVA with repeated measures. Monitoring is the within-subject factor.¹

We first examine whether or not trustees act strategically. According to Hypothesis 1, when monitoring is anticipated, trustees will exhibit more frequent cooperative behavior in monitored than non-monitored rounds. As we observed in Figure 1b, the incidence of cooperation is significantly higher in anticipated monitored rounds than it is in anticipated non-monitored rounds. This difference is significant in both the low-frequency monitoring condition ($M_{\text{Monitored}} = 93\%$, $M_{\text{NonMonitored}} = 33\%$, $F(1, 72) = 47.86$, $p <$

¹ By chance, two dyads in the high-frequency monitoring conditions did not experience non-monitored rounds. In addition, two trustees did not have a chance to make a decision in non-monitored rounds because trustors did not pass any money in these rounds. As a result, we have no observations for these four trustees in non-monitored rounds, and we excluded these trustees from our analyses testing Hypothesis 1.

0.001) and the high-frequency monitoring condition ($M_{\text{Monitored}} = 85\%$, $M_{\text{NonMonitored}} = 39\%$, $F(1, 72) = 37.49$, $p < 0.001$). These findings support Hypothesis 1.

In addition, the amounts trustees returned in monitored rounds were more than double the amounts they returned in non-monitored rounds (low-frequency monitoring conditions: $M_{\text{Monitored}} = 8.13$, $M_{\text{NonMonitored}} = 3.17$, $F(1, 72) = 54.49$, $p < 0.001$; high-frequency monitoring conditions: $M_{\text{Monitored}} = 7.94$, $M_{\text{NonMonitored}} = 3.74$, $F(1, 72) = 47.16$, $p < 0.001$)². These results indicate that trustees behaved strategically in anticipated monitoring conditions.

We examined heterogeneity in behavior among trustees. The right panel of Figure 1c plots trustees' cooperative behavior in non-monitored rounds (y-axis) against their behavior in monitored rounds (x-axis). If trustees exhibit the same level of cooperation in non-monitored rounds as they do in monitored rounds, their points would fall along the 45 degree line. Most, 30 out of 42 (71%) trustees, fell *below* the 45 degree line—these participants behaved strategically. They were more cooperative in monitored rounds than they were in non-monitored rounds.³ Results from 9 out of 42 trustees (21%) fell on the 45 degree line—these participants behaved identically in both monitored and non-monitored rounds; of these 9 trustees, 6 were always cooperative. Surprisingly, 3 trustees were slightly more cooperative in non-monitored than monitored rounds. Overall, these results suggest that a large majority of the trustees behaved strategically.

In this study, the amount trustors passed to their trustee counterparts quadrupled. As a result, the trustors' decision to trust is profitable as long as trustees return at least 25% of what they receive. For instance, if a trustor passed 1 point, the trustee counterpart would receive 4 points. If the trustee counterpart returned 50% of what they received, the trustor would earn 2 points, which is a 100% rate of return. If the trustee returned 1 point (i.e., 25%), the trustor would earn a 0% rate of return. Hence “25%” is the cut-off value of whether or not the decision to trust is profitable in this experiment. As expected, trustees returned the highest percentage in the anticipated monitored rounds and the lowest percentage in the anticipated non-monitored rounds (in the infrequent monitoring condition: $M_{\text{Monitored}} = 46\%$, $M_{\text{NonMonitored}} = 19\%$, $F(1, 72) = 48.12$, $p < 0.001$; in the frequent monitoring condition: $M_{\text{Monitored}} = 45\%$, $M_{\text{NonMonitored}} = 23\%$, $F(1, 72) = 38.18$, $p < 0.001$). As a result, trustors' rates of return were significantly different between the anticipated monitored and non-monitored rounds. In the 80% anticipated monitoring condition, trustors earned an 80.9% (95% CI:

² Similarly, we calculated the average amount returned for each trustee in both monitored and non-monitored rounds. We used the composite measure as the dependent variable in the ANOVA.

³ An individual-level proportion test indicates that 19 out of 30 of these strategic trustees were significantly more compliant in non-monitored rounds than they were in monitored rounds (two-tailed proportion test: $p < 0.05$).

[67%, 95%]]⁴ rate of return in monitored rounds, but a -7.9% (95% CI: [-45%, 39%]) rate of return in the non-monitored rounds. In the 40% anticipated monitoring condition, trustors earned an 84.8% (95% CI: [72%, 98%]) rate of return in monitored rounds, but a -23.9% (95% CI: [-60%, 12%]) rate of return in non-monitored rounds. Overall, these results suggest that trustors enjoyed a significantly positive return in monitored rounds, but a return that is statistically not different from zero in non-monitored rounds.

Hypothesis 2 states that the frequency of monitoring increases compliant behavior. We collapsed the within-subject factor and focused on the between-subject factors. A 2 (frequency) x 2 (anticipation) ANOVA reveals a significant main effect for frequency: when the monitoring frequency was high, trustees had a higher incidence of cooperation ($M_{Low} = 74\%$, $M_{High} = 85\%$, $F(1, 76) = 4.95$, $p < 0.03$). In addition, the frequency of monitoring also increased the amount trustees returned ($M_{Low} = 5.57$, $M_{High} = 7.08$, $F(1, 76) = 5.78$, $p < 0.02$). Taken together, these findings support Hypothesis 2.⁵

Trustor behavior: For each dyad, we have multiple observations in both monitored and non-monitored rounds. For each trustor, we computed the incidence of trusting behavior across their entire decision history to derive a separate trusting measure for monitored and non-monitored rounds. We conducted a 2 (frequency) x 2 (anticipation) x 2 (monitoring) ANOVA with repeated measures. Two dyads did not experience non-monitored rounds, so we used data from the remaining 78 trustors.

As depicted in Figure 1a, trusting rates were uniformly high across the four treatment conditions. The incidence of trusting in anticipated non-monitored rounds was not significantly different from the incidence of trusting in anticipated monitored rounds ($M_{NonMonitored} = 83\%$, $M_{Monitored} = 94\%$, $F(1, 74) = 1.27$, $p > 0.2$). Anticipation of monitoring had no impact on trustor behavior. The main effect of anticipation is insignificant on both the incidence of trusting behavior ($F(1, 76) = 0.34$, $p > 0.5$) and the amount passed ($F(1, 76) = 2.03$, $p > 0.15$).

We explored heterogeneity among trustors. The left panel of Figure 1c plots the incidence of trusting in non-monitored rounds against the incidence of trusting in monitored rounds. As shown, only 8 trustors (19%) trusted less in non-monitored rounds than they did in monitored ones.⁶ A total of 29 of the trustors

⁴ All 95% confidence intervals (CI) reported in the paper are based on the standard errors clustered at the individual level.

⁵ The main effect of anticipation is insignificant on both total incidence of compliance ($F(1, 76) = 0.63$, $p > 0.4$) and the amount returned ($F(1, 76) = 0.34$, $p > 0.5$). Although trustees behaved opportunistically in anticipated monitoring conditions, the total incidence of cooperation across both monitored and non-monitored rounds did not vary across anticipation treatment conditions.

⁶ Note that only 2 out of these 8 trustors made significantly more trusting choices in monitored than non-monitored rounds (two-tailed proportion test: $p < 0.05$).

(69%) chose to trust in all of the monitored and non-monitored rounds. The remaining 5 trustors (12%) trusted more in non-monitored rounds than they did in monitored rounds. These results support Hypothesis 3a; trustors are less sensitive to the absence of monitoring than trustees, and trustors continue to trust and pass money in the anticipated non-monitored rounds.

We found no support for Hypothesis 2. The frequency of monitoring overall neither affected the decision to trust ($M_{\text{Low}} = 88\%$, $M_{\text{High}} = 89\%$, $F(1, 76) = 0.01$, $p > 0.9$) nor the amount that was passed ($M_{\text{Low}} = 3.22$, $M_{\text{High}} = 3.75$, $F(1, 76) = 2.76$, $p = 0.10$). We further investigated the impact of frequency under different anticipation conditions. Interestingly, we observed a significant interaction effect between frequency and anticipation on incidence of passing ($F(1, 76) = 7.36$, $p < 0.01$). In the anticipated monitoring condition, frequency of monitoring increased passing behavior at a marginally significant level ($M_{\text{Low}} = 83\%$, $M_{\text{High}} = 97\%$, $F(1, 76) = 3.26$, $p = 0.08$), which is consistent with Hypothesis 2. In the unanticipated monitoring condition, however, frequency of monitoring significantly decreased passing behavior ($M_{\text{Low}} = 94\%$, $M_{\text{High}} = 80\%$, $F(1, 76) = 4.18$, $p = 0.04$).

The average amount trustees returned was always greater than the average amount trustors passed, except when monitoring was anticipated in the non-monitored rounds. In anticipated non-monitored rounds, passing money did not yield a positive profit ($M_{\text{Low}} = -0.11$, 95% CI: [-1.45, 1.22]; $M_{\text{High}} = -0.31$, 95% CI: [-1.99, 1.37]).

How Trustors React to Observed Compliance. Trustees are frequently compliant in monitored rounds, so we test Hypothesis 3b by exploring how trustors interpret these observations. If trustors made inferences of trustworthiness based on observed compliance, we would expect trustees' history of compliance in the past anticipated monitored round to influence trustors' trust decisions in the subsequent anticipated non-monitored round.

In this analysis, we define $I_i(t)$ as an indicator function: $I_i(t) = 1$ if round t was a monitored round and in that round trustor i passed and trustee i complied; otherwise it is zero. When $I_i(t) = 1$ the trustor observed compliance. In non-monitored rounds or monitored rounds in which the trustor did not pass or the trustee did not comply, the trustor did not observe this signal of compliance. We next define the count of compliance trustor i observes through round t : $C_i(t) = \sum_{\tau=1}^t I_i(\tau)$. Hypothesis 3b implies that as $C_i(t)$ increases, trustors will trust their counterparts more in non-monitored rounds.

Notably, trustors may also observe negative signals: non-compliance behavior in monitored rounds. We define $H_i(t)$ as the indicator function: $H_i(t) = 1$ if round t was a monitored round, trustor i passed money, and the trustee did not comply; otherwise it is zero. The cumulative number of non-compliance signals that trustor i observes through round t is $N_i(t) = \sum_{\tau=1}^t H_i(\tau)$. We conducted regression analysis to test whether trusting behavior (the dependent variable) in non-monitored rounds is positively associated with

$C_i(t)$, which is the cumulative frequency of observed compliance when controlling for $N_i(t)$, which is the cumulative frequency of observed non-compliance.

As shown in Table 1 (column 1 labeled as Study 1 – Trust Decision), the cumulative frequency of observed compliance $C_i(t)$, has a marginally significant ($p < 0.08$) impact on the likelihood of passing: each observation of compliance increased the probability that trustors would pass in a non-monitored round by 1.8%. $C_i(t)$ also has a significant, positive impact on the amount of passing (Table 1, column 2 labeled as Study 1 – Amount Passed). $N_i(t)$ has a negative, but statistically insignificant impact on the likelihood of passing ($p > 0.3$) and the amount of passing ($p > 0.11$). As trustors observed more compliance, they increased the amount of passing from 2.3 points (with no observation of compliance) to more than 4.5 points (when they had more than 13 observations of compliance).⁷

Insert Table 1 about here

Discussion

Monitoring significantly influenced trustee behavior. Trustees were more cooperative when they anticipated that their counterparts would observe their behavior than when they anticipated that their counterparts would not observe their behavior. That is, trustees were opportunistic. Trustors, however, were influenced by the compliance they observed in anticipated monitored rounds, and trusted their counterparts in both anticipated monitored and anticipated non-monitored rounds.

Trustors and trustees were drawn from the same population, and our findings identify an important puzzle. Trustees reacted strategically to anticipated monitoring, but trustors were less sensitive to anticipated monitoring. One possible account for this asymmetric response is a perspective taking failure; trustors may have failed to take their counterparts' perspectives. A second account for the divergent trustor and trustee reactions to anticipated monitoring involves concerns about reciprocity. Although trusting in non-monitored rounds entails risk, it may boost cooperation by triggering reciprocity from trustees in future, non-monitored rounds. That is, trustors may pass money in non-monitored rounds, because they fear signaling distrust and harming future cooperation. We explore the mismatch in strategic behavior and the underlying mechanism in our following studies.

STUDY 2: PERSPECTIVE TAKING

In Study 1, we found that trustors were less responsive to anticipated monitoring than trustees were, and we found that trustors often earned negative profits when trustees anticipated that they would not be

⁷ We used linear regressions for our all analyses (including cases involving 0-1 dependent variables) in order to simplify interpretation. We also ran logistic regressions for cases involving 0-1 dependent variables and we obtained qualitatively similar results. Results are available upon request from authors.

monitored. In this study, we consider whether or not a failure in perspective taking could account for the asymmetry in trustor and trustee behavior when they both knew that a round would not be monitored.

In this study, in addition to making decisions, we asked trustors to predict trustees' behavior. With these data, we can distinguish between failed perspective taking and strategic cooperation (trustors who pass to ensure future cooperation). In this study, we test the following hypotheses:

Hypothesis 4a: *Trustors fail to anticipate how strategically their counterparts behave and misforecast how much trustees will return in non-monitored rounds.*

Hypothesis 4b: *Trustors' forecasts are influenced by the frequency of observed compliance: the more compliant behavior trustors observe in anticipated monitored rounds, the more optimistic trustors become about trustworthy behavior in anticipated non-monitored rounds.*

Method

Participants. We recruited 102 participants from a large public university in Singapore. We offered participants a 5 Singapore dollar (SGD) (about USD 4) show-up fee for a one-hour laboratory study plus the chance to earn additional money. We seated participants at computer cubicles, and participants could only communicate via computer. All of the experimental instructions were presented via computer and an experimenter read the instructions aloud in front of the participants. To ensure that participants fully understood the instructions, we administered a comprehension check before participants could proceed with the experiment. Participants who failed the comprehension check three times were dismissed together with their counterparts. In each session, the group size ranged from 22 to 28.

Design and Procedure. Upon arrival in the behavioral lab, we randomly assigned each participant to a role (odd or even) and to one of two experimental conditions: control or perspective taking. In the control condition, participants played a repeated trust game similar to the repeated trust game participants played in the 40% anticipated monitoring condition in Study 1. In Study 2, however, we simplified the game so that the participants made only binary choices. Specifically, rather than allowing odd players to pass different amounts of points, odd players chose to either pass their entire initial endowment of \$6 or keep all \$6. If odd players passed the \$6, the amount tripled to \$18, and even players chose to either split the \$18 evenly (\$9 for odd, \$9 for even) or keep all of the money for themselves (\$0 for odd, \$18 for even). We believe that by making the players' decisions binary, we increased comprehension without losing the key dynamics of the game.

In Study 2, we asked both odd and even players to make decisions for each round. That is, we asked even players to make a contingent decision to either split or keep the \$18 before they learned whether or not their odd player counterparts had passed them the \$6. This commonly-used "strategy method" enables us to learn even player intentions, irrespective of odd player decisions within a particular round. At the end of the

experiment, we selected one of the rounds at random and we paid participants based upon the outcome of that round.

In the perspective taking condition, participants played the same version of the repeated trust game as the control condition, but we added a series of questions for the odd players. Before making their decisions in each round, we asked odd players the following questions:

If you were the EVEN player in this experiment, what would you have done in this round?

What is your guess of the EVEN player's choice? (A correct guess will earn you \$0.50)

At the end of each monitored round, odd players learned what their counterparts chose. We revealed the choices even players made in non-monitored rounds only at the end of the entire experiment.

After each round, even players learned what the odd players did. As in Study 1, there was an independent 40% chance that each round would be a feedback round. Both players learned whether or not the upcoming round would be a feedback round. If the round was a feedback round, odd players learned what even players chose at the end of that round. If the round was a no feedback round, odd players did not learn what even players chose until the end of the experiment. Participants remained in the same role, matched with the same partner throughout the entire experiment. As in Study 1, we used a stochastic ending. Every participant played a minimum of 14 rounds. After the 14th round, there was an 85% chance of proceeding to round 15, and if there was a 15th round, there was an 85% chance of proceeding to the next round, and so on. Across both conditions, we asked odd players immediately after they made their decisions, “*How much do you trust your counterpart?*” (1: “Not at all” to 7: “Completely”), which is a measure of odd players’ attitudinal trust.

At the end of the study, we collected attitudinal measures to investigate participants’ perceptions of the monitoring system⁸. After completing these measures, participants saw their entire decision history and learned which round would determine their bonus payment. In the perspective taking condition, we paid odd players an additional \$0.50 for every correct forecast they made.

Results

Of the 102 participants who came to the laboratory, 3 dyads were unable to complete the study because one of the partners in the dyad failed the comprehension check three times. We were also unable to collect data in one session because the computer program crashed shortly after the game began. As a consequence, we report results from 84 participants. A total of 18 dyads in the control condition and 24 dyads in the perspective taking condition completed the study. On average, participants completed 19.86 rounds of decisions. We report summary statistics followed by a set of regression results.

⁸ We present analyses of these results in the online appendix.

We report passing and returning decisions in Figure 2a. Overall, trustors passed in 64% of the rounds in the control condition and 63% of the rounds in the perspective taking condition. Trustees returned money in 47% of the rounds in the control condition and 48% of the rounds in the perspective taking condition. On average, the incidence of returning was lower than the incidence of passing in both conditions. Interestingly, we found no difference in players' behavior between the control condition and the perspective taking condition.

Insert Figures 2a and 2b about here

As in Study 1, trustees made strategic decisions in response to monitoring. Trustees were far more likely to return money in anticipated monitored rounds than they were in anticipated non-monitored rounds: in the control condition: $M_{\text{Monitored}} = 82\%$, $M_{\text{NonMonitored}} = 23\%$, $F(1, 40) = 51.29$, $p < 0.001$, and in the perspective taking condition: $M_{\text{Monitored}} = 86\%$, $M_{\text{NonMonitored}} = 23\%$, $F(1, 40) = 65.26$, $p < 0.001$ ⁹. In this study, trustors decided whether or not to pass \$6 to trustees, and trustees decided whether or not to return \$9 to trustors. Therefore, the trustors' decision to trust pays as long as the trustee counterparts return \$9 at least 66.7% of the time. Interestingly, the percentage of rounds that trustees returned money fell below 66.7% in non-monitored rounds in both conditions.

Trustors were responsive to monitoring, but significantly less so than trustees. In the control condition: $M_{\text{Monitored}} = 75\%$, $M_{\text{NonMonitored}} = 57\%$, $F(1, 40) = 10.09$, $p < 0.01$; in the perspective taking condition: $M_{\text{Monitored}} = 77\%$, $M_{\text{NonMonitored}} = 54\%$, $F(1, 40) = 9.1$, $p < 0.01$. We calculated trustors' average earnings conditional on passing across different conditions. As depicted in Figure 2b, although trustors made positive profits in monitored rounds, their average profit in non-monitored rounds was $-\$3.14$ (95% CI: $[-5.6, -0.65]$) in the control condition and $-\$3.44$ (95% CI: $[-5, -1.8]$) in the perspective taking condition. Put differently, for every \$6 passed, trustors received about \$3 in return from trustees. These results demonstrate that trust was exploited when trustees anticipated that they would not be monitored.

Our perspective taking task did not significantly attenuate misplaced trust in anticipated non-monitored rounds. In these rounds, trustors passed similar amounts in the control condition and the perspective taking condition, ($M_{\text{Control}} = 57\%$, $M_{\text{Perspective}} = 54\%$, $F(1, 40) = 0.03$, *NS*).

Trust Level. Trustors' passing decisions are highly correlated with trust ratings across both conditions (Pearson's $\rho > 0.48$). In the control condition, trustors reported the same level of trust in monitored and non-monitored rounds ($M_{\text{Monitored}} = 4.44$, $M_{\text{NonMonitored}} = 4.4$, $F(1, 40) = 0.001$, *NS*). In the perspective

⁹ For each player, we have multiple observations in both monitored and non-monitored rounds, and we computed the average incidence of cooperation in monitored and non-monitored rounds across the entire decision history. We then conducted a 2 (control vs. perspective taking) x 2 (monitoring: monitored vs. non-monitored) ANOVA with repeated measures. Monitoring is the within-subject factor.

taking condition, however, trustors reported that they trusted their trustees somewhat more in monitored rounds than they did in non-monitored rounds ($M_{\text{Monitored}} = 4.53$, $M_{\text{NonMonitored}} = 3.83$, $F(1, 40) = 4.25$, $p < 0.06$).

Response to Monitoring. We performed regression analyses to test whether trustees and trustors responded to monitoring in a symmetric way. We clustered standard errors at the individual level. In the first two columns of Table 2, we report results from linear regression analyses that predict passing/returning decisions for trustors and trustees (1 = passing/returning, 0 = no-passing/no-returning). For independent variables we included role (1 = trustee, 0 = trustor), whether or not a round is a monitored round (1 = monitored round, 0 = non-monitored round), and the interaction between the two. The interaction term is positive and significant across both conditions, demonstrating that trustees were more responsive to monitoring than trustors were. The negative main effect for role reflects that, on average, trustees were less likely to return than trustors were to pass.

Insert Table 2 about here

Perspective Taking. To test Hypothesis 4a, we examine trustors' responses in the perspective taking condition. Trustors could make one of two types of forecasting errors: an *optimistic error*, expecting trustees to return money when trustees chose to keep it, or a *pessimistic error*, expecting trustees to keep the money when trustees chose to return it.

We identify two patterns in trustors' forecast data. First, trustors were significantly more likely to make forecast errors in non-monitored rounds (45%) than they were in monitored rounds (17%) ($F(1, 23) = 11.07$, $p < 0.01$). In non-monitored rounds, trustors' forecasts were only slightly better than chance ($z = 1.7$, $p = 0.08$). Second, we identify a failure of perspective taking in non-monitored rounds. Among the forecast errors in non-monitored rounds, 83% are optimistic errors, far more than pessimistic errors (17%) ($z = 7.24$, $p < 0.001$). In monitored rounds, trustors were slightly, but not significantly more likely to make pessimistic errors (63%) than optimistic errors (37%) ($z = 1.46$, $p = 0.2$). Trustors' forecasts were highly correlated with their passing decisions in both non-monitored (Pearson's $\rho = 0.61$) and monitored (Pearson's $\rho = 0.69$) rounds. Supporting Hypothesis 4a, trustors failed to anticipate how strategically their counterparts would behave and misforecast how much trustees would return in the non-monitored rounds. We depict trustors' forecast errors in Figure 3.

Insert Figure 3 and Figure 4 about here

How Trustors React to Observed Compliance. To test Hypothesis 4b, we explore how trustors react to compliance. As in Study 2, we computed the total number of times trustor i observed compliance up to round t , $C_i(t)$, and the total number of times trustor i observed non-compliance up to round t , $N_i(t)$. We first conducted regression to test whether trusting behavior in the non-monitored round is positively associated

with $C_i(t)$, controlling for $N_i(t)$ (see Table 1, column 3). We found that both $C_i(t)$ and $N_i(t)$ significantly impact passing decisions: each additional observation of compliance increased the probability of passing by 4.3%; each additional observation of non-compliance decreased the probability of passing by 16.7%. Our analyses of trust ratings yielded similar results to the behavioral measures (see Table 1, column 4).

We next examine how trustors' beliefs (which we measured in the perspective-taking condition) react to observed compliance and non-compliance in Table 1 (column 5). We conducted a regression model with trustors' forecast as the dependent variable. We coded the forecast as 1 if the trustor made an optimistic forecast ('trustee will return money') and a 0 otherwise. We found that each additional observation of compliance increased the chance of making an optimistic forecast by 6.1%.

Observed non-compliance, $N_i(t)$, has a significant, negative impact on the likelihood of passing. We plotted the incidence of passing in non-monitored rounds conditional on the cumulative frequency of observed compliance and non-compliance in Figure 4. As trustors observed more compliance, the proportion of trustors who made optimistic forecasts increased from 25% to 75%. Conversely, trustors' beliefs also updated when they observed non-compliance. Trustors who observed two or more rounds of non-compliance, never believed that their counterparts would return money in future rounds.

Discussion

In Study 2, we extended our investigation of monitoring and trust. In this study, we focused on anticipated monitoring and the failure of trustors to appreciate how strategically trustees behave. When trustees anticipated that they would be monitored, they were highly compliant. When trustees anticipated that they would not be monitored, however, they were far less cooperative. Trustors, however, failed to expect this, and as a consequence earned negative returns in non-monitored rounds.

In this study, we also included a perspective taking condition. In this condition, we asked trustors to forecast trustee behavior, and we incentivized accurate forecasts. We found that trustors systematically mispredicted trustee behavior in anticipated non-monitored rounds. These mispredictions were not random. Rather, trustors expected trustees to be trustworthy far more often than they actually were. In this study, we also measured attitudinal trust. We found that attitudinal trust was highly correlated with trustor decisions to pass. Interestingly, we also found that trustors maintained high trust in anticipated non-monitored rounds.

Across both the perspective taking and the control conditions, trustors were heavily influenced by the trustee behavior they observed. When trustees passed in anticipated monitored rounds, trustors expected trustees to be significantly more likely to pass in subsequent anticipated, non-monitored rounds. That is, prior compliance in monitored rounds increased both attitudinal trust and the likelihood that trustors would pass in subsequent non-monitored rounds.

Overall, prompting trustors to forecast their counterparts' behavior did not make them less trusting in non-monitored rounds. It is possible that with greater experience trustors would develop more accurate perceptions. We consider this proposition in our next study, where we expose trustors to more information.

Taken together, our findings document a surprising perspective taking failure. Though trustors and trustees were drawn from the same sample population, the strategic perspectives they adopted were highly disparate. An incentivized perspective taking condition revealed—rather than mitigated—this perspective taking failure.

STUDY 3: LEARNING FROM OTHERS' EXPERIENCES

In Study 3, we extend our investigation of anticipated monitoring and used methods similar to those used in Study 2. In this study, participants played the same binary-choice repeated trust game as they did in Study 2. To moderate the perspective taking failure we identify in Study 2, we include a condition in which trustors observe behavior in other dyads. Specifically, trustors observe how trustees in other dyads responded to monitoring.

By observing trustee behavior in other dyads, trustors witnessed strategic trustee behavior. These observations should mitigate the perspective taking failure. However, recent work suggests that learning from the experiences of others is limited. The same information is weighted more heavily if it reflects personal experience rather than an observation of another's experience (Simonsohn, Karlsson, Loewenstein, and Ariely 2008; Haselhuhn, Pope, Schweitzer, and Fishman 2012). It is possible that trustors discount the information they glean from observing others and instead rely heavily on their own direct, positive experience even when these experiences come from a biased sample of monitored rounds. If participants fail to learn from others' experiences, we would observe persistent over-trusting in non-monitored rounds.

Method

Participants. As in Study 2, we recruited student participants from a university in Singapore to participate in a one-hour laboratory study in exchange for a SGD 5 (about USD 4) show-up fee and the opportunity to earn additional money. We recruited 108 participants and after explaining the procedures, we administered a comprehension quiz. Participants who failed the quiz three times were dismissed together with their counterpart.

Design and Procedure. As in Study 2, we used the binary version of the trust game with 40% anticipated monitoring. Odd players could either keep \$6 or pass it to their even player counterpart. If they passed the \$6, the amount tripled and their counterpart had \$18 to either split evenly or keep for themselves. As in Study 2, we used the strategy method and asked even players to make their decisions before knowing whether or not their odd player counterpart had passed their initial endowment in each round.

In this study, within each session, we assigned half of the participants to the control condition, and half to the observing condition. Odd players in the observing condition played the same repeated trust game with a partner as those in the control condition, but these odd players were also able to observe the behavior of the dyads in the control condition. Specifically, in the observing condition, odd players saw a summary table after each round that depicted the decisions of dyads in the control condition (see Figure 5 for a screenshot of sample feedback for an odd player in the observing condition). The summary table also indicated whether or not the round in the control condition was an anticipated monitored round or an anticipated non-monitored round. Participants in the control condition did not know that their behavior would be observed by other participants, and we did not inform even players in the observing condition that their counterparts were observing the decisions of other groups.

Insert Figure 5 about here

As in Study 2, there was an independent 40% chance that any round would be a feedback round. Participants remained in the same role with the same partner throughout the entire experiment. Similar to the stochastic ending rule we used in Study 2, everyone played 14 rounds, and after round 14, for every subsequent round, there was an 85% chance of proceeding to the next round. For each round, immediately after trustors made their decisions, we asked the trustors “How much do you trust your counterpart?” (1: “Not at all” to 7: “Completely”).

In the observing condition, immediately after odd players indicated how much they trusted their counterpart, odd players saw a summary table of the full set of decisions and whether or not a round was a monitored round for every dyad in the control condition within their session. At the end of the study, we asked post-experiment questions, revealed the entire history of their own game, randomly selected one of the rounds, and paid participants based upon their outcome for that round.

Results

One participant failed the comprehension quiz three times so we dismissed this participant and their counterpart. We report results for the remaining 106 participants; 28 dyads made decisions in the control condition and 25 dyads made decisions in the observing condition. On average, participants completed 20.8 rounds. The number of participants per session ranged from 22 to 28. We report summary statistics followed by a set of regression analyses.

Overall, trustors trusted in 56% of the rounds in the control condition, and in 60% of rounds in the observing condition. Trustees returned in 54% of the rounds in the control condition, and in 50% of rounds in the observing condition. In both conditions, on average, the incidence of returning was lower than the incidence of passing.

As in Studies 1 and 2, trustees behaved strategically in response to anticipated monitoring. Trustees passed significantly more often in monitored rounds than they did in non-monitored rounds (in the control condition: $M_{\text{Monitored}} = 85\%$, $M_{\text{NonMonitored}} = 34\%$, $F(1, 51) = 77.67$, $p < 0.001$; in the observing condition: $M_{\text{Monitored}} = 85\%$, $M_{\text{NonMonitored}} = 26\%$, $F(1, 51) = 82.51$, $p < 0.001$). Similar to Study 2, in non-monitored rounds, the incidence of compliance fell below 66.7%, the break-even point for profitability; on average, trustors lost money passing in non-monitored rounds. In non-monitored rounds, trustors earned: $-\$2.19$ (95% CI: $[-4.3, -0.03]$) in the control condition, and $-\$2.64$ (95% CI: $[-4.9, -0.34]$) in the observing condition. In these rounds, as we found in Studies 1 and 2, trust was misplaced.

Trustors were also responsive to whether or not the round was monitored. Trustors were more likely to pass in monitored rounds than they were in non-monitored rounds (in the control condition: $M_{\text{Monitored}} = 68\%$, $M_{\text{NonMonitored}} = 47\%$, $F(1, 51) = 9.09$, $p < 0.01$; in the observing condition: $M_{\text{Monitored}} = 77\%$, $M_{\text{NonMonitored}} = 49\%$, $F(1, 51) = 22.46$, $p < 0.001$). Trustors in the observing condition passed just as often in the non-monitored rounds as they did in the control condition: $M_{\text{Control}} = 47\%$ vs. $M_{\text{Observing}} = 49\%$, $F(1, 51) = 0.11$, $p > 0.9$. Trustors in the observing condition passed slightly more often in the monitored rounds than they did in the control condition, but the difference was not significant: $M_{\text{Control}} = 68\%$ vs. $M_{\text{Observing}} = 77\%$, $F(1, 51) = 1.98$, $p > 0.16$).

Response to Monitoring. Trustors were less responsive than trustees to whether or not a round was monitored. As a result, in non-monitored rounds trustors passed significantly more than they should have. In Table 2 in the last two columns, we report results from linear regressions that analyze passing decisions. We modeled trustor and trustee passing/returning decisions (1 for passing/returning, and 0 otherwise) as a function of role (1 = trustee, 0 = trustor), monitoring (1 for a monitored round and 0 for non-monitoring), and the interaction between role and monitoring. The interaction is positive and significant in both conditions, indicating that trustees are more sensitive to monitoring than trustors. The main effect for role type in the observing condition is negative, indicating that trustees were less likely to return money than trustors were to pass.

How Trustors React to Observed Compliance. As in Study 1 and Study 2, we investigate the impact of the cumulative frequency of compliance $C_i(t)$ and the cumulative frequency of non-compliance $N_i(t)$. In this study, we further examine whether learning from others' experiences could attenuate the misplaced trust caused by $C_i(t)$. We modeled trustor passing decision as a function of $C_i(t)$, $N_i(t)$, the treatment dummy variable (1 for the observing condition and 0 for the control condition), and the interaction between $C_i(t)$ and the treatment dummy variable. The interaction effect between $C_i(t)$ and the treatment dummy is insignificant ($\beta = -0.003$, $t(52) = -0.18$, $p > 0.8$), suggesting that observing other dyads' behaviors did not attenuate the impact of observed compliance.

Since the observing condition is not significantly different from the control condition, we pooled the data for subsequent analysis. As in Study 1 and Study 2, we that found that both the cumulative frequency of compliance and the cumulative frequency of non-compliance significantly impact the likelihood of passing (Table 1, column 6). Each additional observation of compliance increased the probability that trustors passed by 3.9%; each additional observation of non-compliance decreased the probability that trustors passed by 15.1%. Trustors' perceived trustworthiness, measured by trust rating, was also positively correlated with observed compliance and negatively correlated with non-compliance (Table 1, column 7).

Discussion

In this study, we extend our investigation to explore how learning from others' experiences might improve perspective taking. In our study, trustors in the observing condition learned about the strategic behavior of trustees in other dyads. Although their personal experiences in monitored rounds were less informative than others' experiences in non-monitored rounds, it was trustors' personal experiences that influenced how they behaved in non-monitored rounds. As before, we found a consistent pattern: trustees behaved strategically, exhibiting compliant behavior when they knew they would be monitored and exploiting trust when they knew they would not be monitored. And trustors, failing to anticipate this strategic behavior, placed too much trust in their counterparts when both parties anticipated that they would not be monitored.

STUDY 4: BASE RATES AND SHIFTING BELIEFS

We conducted a final study to explore trustors' beliefs and the attribution process. We gave participants instructions very similar to those we used in Studies 2 and 3, and we asked trustors about their beliefs regarding the trustworthiness of trustees. We solicited initial beliefs and we explored how these beliefs might shift after observing trustworthy behavior in monitored rounds.

Method

Participants. We recruited 149 undergraduate students from a large public university in Singapore for a one-hour laboratory study. We promised each participant SGD 10 (about USD 8) and the chance to earn additional money. We presented participants with the same instructions as we used in Studies 2 and 3, and we asked participants to imagine that they were playing the game as the odd player (trustor). To incentivize participants to read the instructions carefully, we administered a comprehension check quiz and told participants that we would pay them a bonus of SGD 8 if they answered the questions correctly.

Design and Procedure. After passing the comprehension check, we asked participants a series of questions. First, we asked a question about base rates of trustworthy trustees:

Imagine that the EVEN players were randomly selected from the other participants in this session. If you choose to pass, what percentage of EVEN players do you think will return in this first NO FEEDBACK round? %_____

Next, we asked participants to imagine they had observed compliant behavior. Specifically, we asked them to imagine that Round 1 was a feedback round in which they had passed and their partners had returned.

Round 2 is a NO FEEDBACK round. If you choose to pass in Round 2, how likely is it that your partner will return \$9? %_____

In addition to asking how likely their partners were to return \$9, we asked participants to provide a general assessment about their counterparts' character: "1) S/he reacts to incentives in the moment. 2) Somewhat trustworthy person / somewhat reacts to incentives. 3) S/he is a very trustworthy person in general." We then asked participants these same two questions after we asked them to imagine that their partner returned money in two consecutive rounds, and then in three consecutive monitored rounds.

Insert Table 3 about here

Results

We analyzed data from the 117 participants who correctly answered all of the comprehension check questions¹⁰. Without any prior interaction, trustors expected 41% of trustees to return money. This expected 41% rate of returning money is significantly lower than the 66.7% threshold that would make passing money profitable ($t(116) = -10.03, p < 0.001$), but the rate is higher than the 35.1% rate of returning money that we observed in Studies 2 and 3 ($t(116) = 2.3, p < 0.03$). That is, this estimated base rate of 41% is overly-optimistic.

We next examine how observing compliant behavior shifts beliefs. Table 3 shows that the more compliant behavior trustors' observed, the more likely they thought their counterparts were to return money (Column 2) and the more likely they were to characterize their counterpart as a very trustworthy person (Column 3). We conducted two linear regression analyses to test how trustors' beliefs and attributions shifted after observing compliance.

In the first regression analysis, we regressed trustors' subjective beliefs for how likely trustees are to return money on the number of observations of compliance. We found a significant, positive coefficient ($\beta = 0.03, t(116) = 3.04, p < 0.01$), indicating that each additional observation of compliance increased subjective belief by 3%. In the second regression analysis, we used whether or not trustors attributed compliance to trustworthiness as the dependent variable (1 if trustors chose "S/he is a very trustworthy person in general"; 0

¹⁰ Our results are virtually identical when we include the 32 participants who missed one or more comprehension check questions.

otherwise), and the number of observations of compliance as the independent variable. We found that each additional observation of compliance increased the likelihood of characterizing the trustee as trustworthy by 5% ($\beta = 0.05$, $t(116) = 3.51$, $p < 0.01$).

Discussion

In this study, we investigated subjective beliefs about trustworthiness. We find that initial beliefs of trustworthiness are quite high—higher than the actual trustworthiness we observed in Studies 2 and 3. We also document a rise in perceptions of trustworthiness after participants imagine observing compliant behavior. Consistent with our findings in Studies 1, 2, and 3, trustors mistake compliant behavior for a meaningful signal of trustworthiness.

GENERAL DISCUSSION

Monitoring influences behavior. In this investigation, we focus particular attention on how anticipated monitoring influences trusting and trustworthy behavior. We report results from a series of repeated trust game experiments, and we examine two empirical regularities. First, trustees who were monitored engaged in strategic behavior. When they anticipated that they would be monitored, they complied. When they anticipated that they would not be monitored, they defected. Second, trustors failed to anticipate this strategic behavior. Whereas trustees acted as if they were playing two different games, one with monitoring and one without monitoring, trustors acted as if they were playing one game and passed similar amounts regardless of whether or not the round was monitored.

We investigate the underlying mechanism of this failed mental model. Even though trustors and trustees in our studies were drawn from the same population, our studies consistently demonstrate that trustors fail to appreciate how strategically trustees behave. We find that trustors' initial beliefs are biased. They expect the population of trustees to be more trustworthy than they actually are. Worse, we find that trustors rely on the unrepresentative sample of observations they can observe. In monitored rounds, trustees complied and exhibited cooperative behavior. Trustors mistake compliance as a signal of trustworthiness. The more compliance they observe, the more trusting they become. Compliance is not diagnostic of how trustees will actually act in non-monitored rounds. However, after observing compliance, trustors become more likely to pass money in non-monitored rounds.

In Studies 2 and 3, we attempt to de-bias trustors by prompting them to engage in perspective taking and by showing them information about other dyads in which trustees acted strategically. Both attempts failed to close the gap in perspective taking. Trustors mispredicted their counterparts' behavior, and the more compliant behavior they observed, the worse their mispredictions became.

Our findings suggest that managers who observe others when monitoring is anticipated may badly misplace their trust. For example, if employees know when they will be monitored, they are likely to work assiduously during that monitoring period, and shirk when they are not monitored. Managers, however, are likely to over-rely on what they observe and place their trust in employees when they cannot monitor them.

We contrast our findings with those of Malhotra and Murnighan (2002) and Strickland (1958). Whereas prior work identified conditions under which observing compliant behavior harmed trust, we identify conditions under which observing complaint behavior increased trust. We conjecture that the differences we observe in these studies reflect differences in the attributions participants make about the behavior they observe.

Malhotra and Murnighan (2002) found that contracts impeded trust development; when individuals observed compliant behavior, they attributed that behavior to the situation (e.g., the contract influenced behavior) rather than the person (e.g., she is a trustworthy person). In Malhotra and Murnighan (2002), participants *had* to comply with the contract. We postulate that the clear and deterministic nature of this compliance made the situational inference highly salient.

In Strickland's (1958) study, participants compared two employees and placed greater trust in the infrequently monitored employee. Participants in this study, however, also learned that both employees had achieved similar outputs. In this case, the salient difference between the two employees is how frequently they were monitored, and as a result, the salient attribution is that the infrequently monitored participant is intrinsically motivated. In contrast to Strickland's (1958) investigation, participants in our study had no outcome data to inform their assessment of what their counterparts did when they were not monitored. In our studies, participants systematically made incorrect inferences about how targets behaved when they were not monitored.

Our key finding is that participants misattribute the compliant behavior they observe, and fail to appreciate how strategically their counterparts, who were drawn from the same sample, act. That is, participants in our studies failed to appreciate the conditional nature of the compliant behavior they observed. Rather than making a situational attribution (e.g., my counterpart was compliant because she anticipated that I would observe her behavior), participants made a personal attribution (e.g., my counterpart was compliant because she is a trustworthy person). This misattribution process is consistent with prior work that has documented judgment biases related to conditional thinking (e.g., Burns & Wieth, 2004).

One important characteristic of our studies is that we exogenously imposed a monitoring system. That is, trustors did not decide when to monitor trustees. It is possible that monitoring systems in which trustors decide when to monitor will elicit different behavior. They could increase trust, but we suspect that in many cases by choosing to monitor employees, managers may signal distrust, provoke reactance, and perhaps

exacerbate strategic behavior. Future work should explore how endogenous and exogenous monitoring systems differ.

Future work should also extend our investigation to consider the costs of non-compliance. In our experiment the cost of non-compliance was ambiguous. After observing non-compliance, trustors became less likely to pass money. Future work could explore how explicit consequences change behavior and attributions. For example, by making consequences explicit (e.g., you will be fired if you do not show-up to work on time), trustors may become more likely to make situational attributions (e.g., you showed-up for work on time because you feared being fired) rather than dispositional attributions (e.g., you are a trustworthy person) for compliant behavior.

Though technology has expanded the tools managers can use to monitor employees, we know surprisingly little about how monitoring changes the behavior of both those who are monitored and those who do the monitoring. Managers routinely face the challenge of trusting others, and our findings suggest that they may be particularly susceptible to misplacing their trust. Our findings demonstrate that managers should view the behavior they observe skeptically, consider changing the nature of the monitoring system they use, and recognize that what they cannot see may be far more diagnostic of trustworthiness than what they can.

REFERENCES

- American Management Association (2014). The Latest on Workplace Monitoring and Surveillance.
<http://www.amanet.org/training/articles/The-Latest-on-Workplace-Monitoring-and-Surveillance.aspx>
- Association of Certified Fraud Examiners. (2012). Report to the Nations on Occupational Fraud and Abuse,
<http://www.acfe.com/>
- Alder, S., Noel, T., & Ambrose, M. (2006). Clarifying the Effects of Internet Monitoring on Job Attitudes: The Mediating Role of Employee Trust, *Information and Management*, 43, 894-903.
- Andrade, E. & Ho, T. (2007) How is the Boss's Mood Today? I Want a Raise. *Psychological Science*, 18(8), 668-671.
- Andrade, E. & Ho, T. (2009). Gaming Emotions in Social Interactions, *Journal of Consumer Research*, 36, 539-552.
- Berg, J., Dickhaut, J. & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10, 122-142.
- Brooks, A. W., Dai, H., & Schweitzer, M. E. (2014). I'm sorry about the rain! Superfluous apologies demonstrate empathic concern and increase trust. *Social Psychological and Personality Science*, 5(4), 467-474.
- Burns, B. D., & Wieth, M. (2004). The collider principle in causal reasoning: why the Monty Hall dilemma is so hard. *Journal of Experimental Psychology: General*, 133(3), 434.

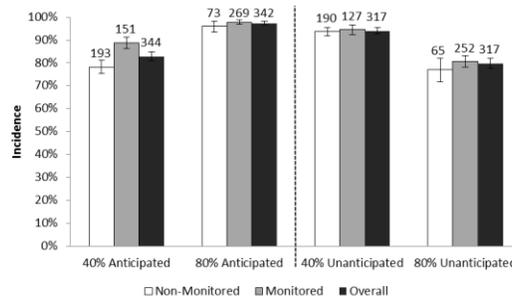
- Cialdini R. (1996). The triple tumor structure of organizational behavior. *In Codes of Conduct*, ed. D. M. Messick, A. E. Tenbrunsel, pp. 44-58. New York: Russell Sage Foundation.
- Dirks, K. T., & Ferrin, D. L. (2001). The role of trust in organizational settings. *Organization Science*, 12, 450-467.
- Dunn, J. & Schweitzer, M. (2005). Feeling and believing: The influence of emotion on trust. *Journal of Personality and Social Psychology*, 88(6), 736-748.
- Enzle M. E., & Anderson S. C. (1993). Surveillant intentions and intrinsic motivation. *Journal of Personality and Social Psychology* 64, 257-266
- Frey, B. (1993). Does monitoring increase work effort? The rivalry with trust and loyalty. *Economic Inquiry*, 4, 663-670.
- Gambetta, D. (1988). Can we trust trust? In D. Gambetta (Ed.), *Trust: Making and breaking cooperative relations*, 213-237. Cambridge, MA: Basil Blackwell.
- Gilbert, D. T., & Malone, P. S. (1995). The Correspondence Bias. *Psychological bulletin*, 117(1), 21.
- Goffman, E. (1959). *The Presentation of Self in Everyday Life*. New York: Doubleday Anchor.
- Haselhuh, M., Pope, D., Schweitzer, M., & Fishman, P. (2012). The Impact of Personal Experience on Behavior: Evidence from Video-Rental Fines, *Management Science*, 58, 52-61.
- Haselhuhn, M., Schweitzer, M. & Wood, A. (2010). How implicit beliefs influence trust recovery, *Psychological Science*, 21 (2010), pp. 645–648.
- Haselhuhn, M., Schweitzer, M. E., Kray, L., & Kennedy, J. A. (2015). Perceptions of High Integrity Can Persist after Deception: How Implicit Beliefs Moderate Trust Erosion. *Journal of Business Ethics*, Forthcoming.
- Ho, T. & Weigelt, K. (2005). Trust Building Among Strangers. *Management Science*, 51, 519-530.
- Kim, P. H., Dirks, K. T., Cooper, C., & Ferrin, D. L. (2006). When more blame is better than less: The implications of internal vs. external attributions for the repair of trust after a competence- vs. integrity-based violation. *Organizational Behavior and Human Decision Processes*, 99, 49-65.
- Kim, P., Ferrin, D., Cooper, C., & Dirks, K. (2004). Removing the shadow of suspicion: The effects of apology versus denial for repairing competence-versus integrity-based trust violations. *Journal of Applied Psychology*, 89, 104-118.
- Kramer, R. (1999). Trust and distrust in organizations: Emerging perspectives, enduring questions. *Annual Review of Psychology*, 50: 569-598.
- Kramer, R. & Lewicki, R. (2010) Repairing and enhancing trust: Approaches to reducing organizational trust deficits, *The Academy of Management Annals*, 4, 245–277.
- Kramer, R. & Tyler, T. (1996) *Trust in Organizations*. Thousand Oaks, CA: Sage.

- Leary, M. (1996). *Self-Presentation: Impression Management and Interpersonal Behavior*. Dubuque, IA: Brown & Benchmark.
- Levine, E. E., & Schweitzer, M. E. (2015). Prosocial lies: When deception breeds trust. *Organizational Behavior and Human Decision Processes*, 126, 88-106.
- Lewicki, R. J., Tomlinson, E. C., & Gillespie, N. (2006). Models of interpersonal trust development: Theoretical approaches, empirical evidence, and future directions. *Journal of Management*, 32: 991–1022.
- Lount, R. (2010). The impact of positive mood on trust in interpersonal and intergroup interactions. *Journal of Personality and Social Psychology*. 98, 420-433.
- Lount R., Zhong C., Sivanathan N., & Murnighan K.(2008). Getting off on the wrong foot: The timing of the breach and the restoration of trust. *Personality and Social Psychology Bulletin*. 34: 1601-1612.
- Malhotra, D. (2004) Trust and reciprocity decisions: The differing perspectives of trustors and trusted parties, *Organizational Behavior and Human Decision Processes*, 94, 61-73.
- Malhotra, D., & Murnighan, J. K. (2002). The effects of formal and informal contracts on interpersonal trust. *Administrative Science Quarterly*, 47, 534–559.
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20, 709-734.
- Niehoff, B. P., & Moorman, R. H. (1993). Justice as a mediator of the relationship between methods of monitoring and organizational citizenship behavior. *Academy of Management Journal*, 36: 527-556.
- Pillutla, M., Malhotra, D., & Murnighan, J. K. (2003). Attributions of trust and the calculus of reciprocity. *Journal of Experimental Social Psychology*, 39, 448–455.
- Shapiro, D., B. H. Sheppard, and L. Cheraskin (1992), Business on a Handshake, *Negotiation Journal*, 8 (4), 365-377.
- Schlenker, B. & Leary, M. (1982). Social anxiety and self-presentation: a conceptualization and model, *Psychological Bulletin*. 92:641-69.
- Schweitzer, M., Hershey, J., & Bradlow, E. (2006). Promises and lies: Restoring violated trust. *Organizational Behavior and Human Decision Processes*, 101(1), 1-19.
- Schweitzer, M. & Ho, T. (2005). Trust but verify: Monitoring in interdependent relationships, in *Experimental and Behavioral Economics* (Ed. J. Morgan), 13, 87-106.
- Simonsohn, U., Karlsson, N., Loewenstein G., & Ariely, D. (2008). The tree of experience in the forest of information: Overweighing experienced relative to observed information, *Games and Economic Behavior*, 62, 263-286.
- Strickland, L. (1958). Surveillance and Trust. *Journal of Personality*, 26, 200–215.

Yip, J. A., & Schweitzer, M. E. (2015). Trust Promotes Unethical Behavior: Excessive Trust, Opportunistic Exploitation, and Strategic Exploitation. *Current Opinion in Psychology*. Forthcoming.

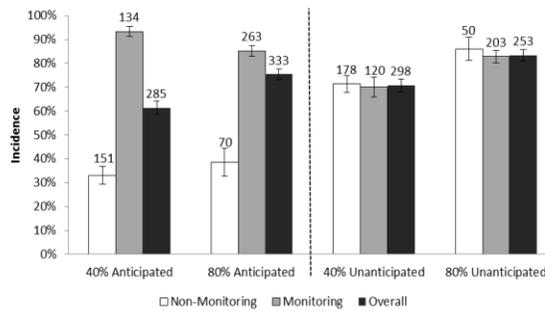
Figures

Figure 1a. Incidence of Passing Behavior (Study 1)



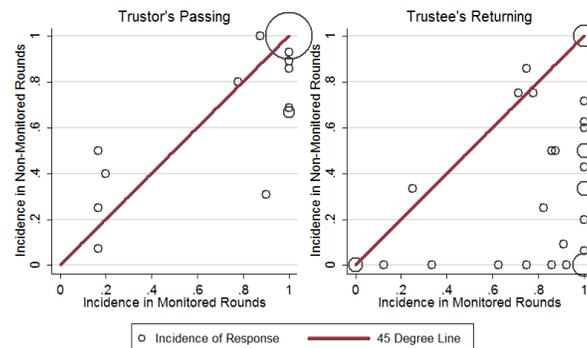
Note: Across all conditions in both monitored and non-monitored rounds, trustors passed most of the time. Error bars represent ± 1 standard error. Numbers above the bars indicate the number of observations.

Figure 1b. Incidence of Returning Behavior Conditional on Positive Passing (Study 1)



Note: Trustees were far more likely to return money in anticipated monitored rounds than they were in anticipated non-monitored rounds. Error bars represent ± 1 standard error. Numbers above the bars indicate the number of observations.

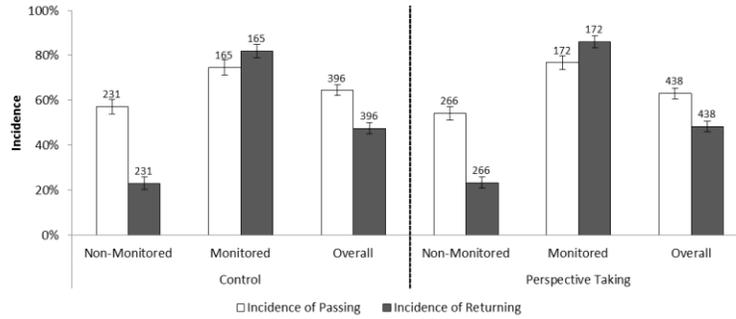
Figure 1c. Individual-level Analysis of Cooperative Behavior in Anticipated Monitored and Anticipated Non-Monitored Rounds (Study 1)



Note: The horizontal axis represents the incidence of cooperative behavior (passing/returning) in anticipated monitored rounds. The vertical axis represents the incidence of cooperative behavior (passing/returning) behavior in anticipated non-monitored rounds. If the incidence of cooperative behavior in anticipated monitored and anticipated non-monitored rounds is equal, the data falls along the 45 degree line. Observations below the 45 degree line represent strategic behavior. The size of the circle indicates the frequency of observations. 71% of trustees returned more money in

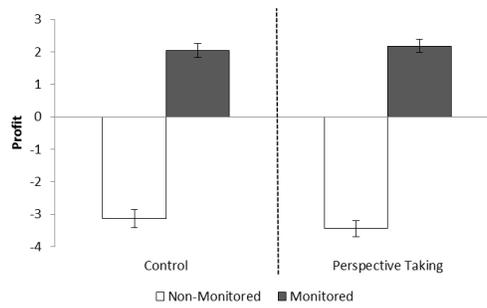
anticipated monitored rounds than they did in anticipated non-monitored rounds; 69% of trustors passed money in all anticipated monitored and anticipated non-monitored rounds.

Figure 2a. Incidence of Passing and Returning by Condition and Type of Round (Study 2)



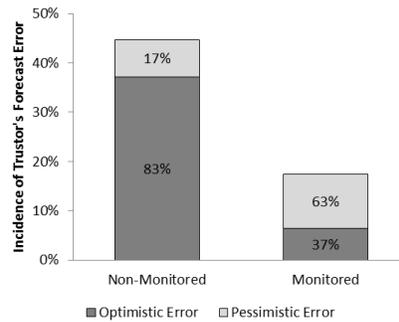
Note: Trustees were more responsive to monitoring than trustors. Error bars represent ± 1 standard error. Numbers above the bars indicate the number of observations.

Figure 2b. Profit for Trustors Conditional on Passing (Study 2)



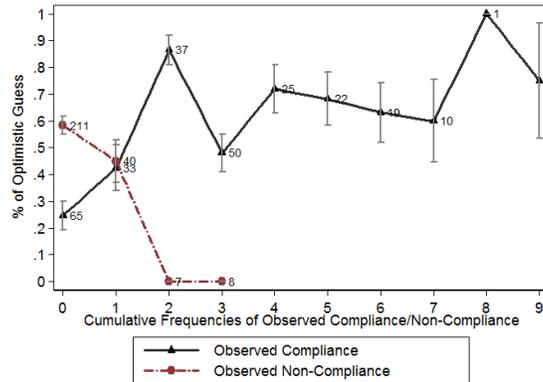
Note: Conditional on passing, the profit from passing was negative in non-monitored rounds. Error bars represent ± 1 standard error.

Figure 3: Incidence of Trustors' Forecast Error (Study 2)



Note: Trustors made twice as many forecast errors in non-monitored rounds as they did in monitored rounds. The vast majority (83%) of forecast errors in non-monitored rounds were optimistic errors.

Figure 4: Trustors' Belief on Trustees' Returning Behavior Following Compliance/Non-Compliance (Study 2)



Note: Trustors made optimistic forecasts when they believed that trustees were going to return money. The vertical line represents the percentage of trustors who made optimistic forecasts. As trustors observed compliant (non-compliant) behavior, they became more (less) likely to believe that their trustee counterparts would return money in non-monitored rounds. Error bars represent ± 1 standard error. Numbers on the side of markers indicate the number of observations.

Figure 5: Sample Screenshot of the Trustor's Observation Table for 3 Dyads after 3 Rounds (Study 3)

End of Round 3

EVEN chose RETURN \$9.

Observation Table

Other matched pairs in the experiment made the following decisions in the prior rounds.

Round	Pair	Feedback?	ODD Chose	EVEN Chose	Odd's Payoff	Even's Payoff	Odd's Total Payoff	Even's Total Payoff
1	1	No	PASS \$6	RETURN \$0	0	18	0	18
2	1	Yes	PASS \$6	RETURN \$9	9	9	9	27
3	1	No	PASS \$6	RETURN \$0	0	18	9	45
1	2	No	TAKE \$6	RETURN \$0	6	0	6	0
2	2	Yes	TAKE \$6	RETURN \$9	6	0	12	0
3	2	Yes	PASS \$6	RETURN \$9	9	9	21	9
1	3	No	PASS \$6	RETURN \$9	9	9	9	9
2	3	Yes	PASS \$6	RETURN \$0	0	18	9	27
3	3	No	TAKE \$6	RETURN \$9	6	0	15	27

Click on the CONTINUE button to proceed.

Note: In the observing condition, trustors received complete information about the decisions other participants made in other dyads. This sample screenshot provides information about 3 other dyads after 3 rounds of decisions.

Tables

Table 1: Trustors' Responses to Observed Compliance/Non-Compliance (Studies 1, 2, and 3)

	Study 1		Study 2			Study 3	
	Trust Decision	Amount Passed	Trust Decision	Trust Rating	Optimistic Guess	Trust Decision	Trust Rating
Cum. Freq. of Compliance	0.018 (0.010)	0.155** (0.048)	0.043* (0.019)	0.272** (0.054)	0.061* (0.027)	0.039** (0.010)	0.191** (0.034)
Cum. Freq. of Non-Compliance	-0.062 (0.059)	-0.450 (0.277)	-0.167* (0.062)	-1.146** (0.196)	-0.208** (0.062)	-0.151** (0.042)	-0.696** (0.217)
Constant	0.770** -0.086	2.387** (0.414)	0.478** (0.074)	3.646** (0.275)	0.432** (0.097)	0.393** (0.066)	3.411** (0.278)
Observations	266	266	497	497	266	657	657
R-squared	0.053	0.128	0.107	0.337	0.145	0.154	0.201

Note: Trustor behavior was significantly influenced by the compliant behavior they observed. This table reports regression results for trustors' responses to observed compliance and non-compliance in non-monitored rounds. In parentheses we report robust standard errors clustered at the individual level. * $p < 0.05$, ** $p < 0.01$ (two-tailed).

Table 2: Passing and Returning Decisions by Role and Monitoring (Studies 2 and 3)

	Study 2		Study 3	
	Control	Perspective Taking	Control	Observing
Role	-0.342** (0.125)	-0.308** (0.109)	-0.134 (0.114)	-0.225* (0.102)
Monitoring	0.174 (0.090)	0.226** (0.052)	0.210** (0.055)	0.292** (0.060)
Role*Monitoring	0.415** (0.128)	0.401** (0.083)	0.298** (0.088)	0.299** (0.093)
Constant	0.571** (0.094)	0.541** (0.088)	0.474** (0.089)	0.487** (0.074)
Observations	792	876	1,204	1,002
R-squared	0.215	0.238	0.147	0.222

Note: The dependent variable is passing and returning for both trustors and trustees. The positive interaction term indicates that trustees were more sensitive to the monitored rounds than trustors were. In parentheses we report robust standard errors clustered at the individual level. * $p < 0.05$, ** $p < 0.01$ (two-tailed).

Table 3: Trustors' Responses to Observed Compliance (Study 4)

Information	Forecast of How Likely Trustees are to Return \$	% of Trustors Choosing "S/he is a very trustworthy person in general"
With No Prior Interaction	41.0% (2.6)	-
After 1 Observation of Compliance	39.4% (2.2)	4.3% (1.9)
After 2 Observations of Compliance	45.4% (2.3)	11.1% (2.9)
After 3 Observations of Compliance	48.9% (2.5)	15.4% (3.4)

Note: The table summarizes trustors' subjective beliefs and attributions following observed compliance ($N = 117$). We report standard errors in parentheses.