# Reducing Exposure to Trust-Related Risks to Avoid Self-Blame

Daniel A. Effron

Dale T. Miller

Stanford University

Published in 2011 in *Personality and Social Psychology Bulletin,* Vol 37, pp. 181-192

http://doi.org/10.1177/0146167210393532

## Abstract

Three studies demonstrated that anticipated self-blame elicits more conservative decisions about risks that require trust than about otherwise economically identical risks that do not. Participants were more reluctant to invest money in a company when it risked failure due to fraud versus low consumer demand (Study 1), and to risk points in an economic game when its outcome ostensibly depended on another participant versus chance (Studies 2 and 3). These effects were mediated by anticipated self-blame (Studies 1 and 2). Additionally, participants who actually experienced a loss felt more self-blame when the loss violated their trust, and became even more conservative in subsequent risk decisions relative to participants whose loss did not violate their trust (Study 3). No support emerged for alternative explanations based on either the perceived probability of incurring a loss, or on an aversion to losses that profit others. The motivational power of trust violations is discussed.


Keywords: trust, risk, self-blame, exploitation, sucker effect, decision making, invest, regret, trust game

# Reducing Exposure to Trust-Related Risks in Order to Avoid Self-Blame

In the wake of the 2008 stock market crash, public attention focused on Bernard Madoff, who defrauded investors of billions of dollars (Smith, 2009). Outrage at Madoff stemmed not only from the magnitude of the financial damage he caused, but also from the perception that he had "shattered trust and for that he should never be forgiven" (Gillers, 2009). Reactions like this suggest that people perceive trust violations to be harmful in and of themselves, independent of the financial or material harm they cause. This perception and its consequences for decision-making are the focus of the present paper.

Consistent with other researchers, we define trust as a willingness to make oneself vulnerable to another person based on the expectation that he or she does not have or will not act on harmful or selfish intentions (Rousseau, Sitkin, Burt, & Camerer, 1998; Simpson, 2007). All risks make one vulnerable in some way, but not all risks involve trust. Loaning money that one expects to be repaid involves trust, whereas placing a bet that one expects to win does not. Only in the former case does risking money make one vulnerable to the exploitative intentions of another person. Our central thesis is that people are more motivated to avoid losses that represent trust violations than losses that are otherwise economically identical (i.e., have the same probability of resulting in a material loss of equal magnitude) but do not violate trust. The present studies explore how this motivation manifests itself in the context of risk decisions.

Why would people be especially motivated to avoid losses that represent trust violations? Prior research has suggested that people expect trust violations to add disutility to loss (Bohnet & Zeckhauser, 2004; Koehler & Gershoff, 2003). For example, people predict that trust violations will incite anger, resentment, and punitive sentiment towards the violator (Koehler & Gershoff,

2003). Going beyond prior research, we contend that losses involving trust-violations are especially aversive because they frequently elicit self-blame. Victims of trust violations often feel foolish, think "they should have known better" than to trust, and experience the sting of "kicking themselves" that characterizes regret (Zeelenberg, van Dijk, Manstead, & van der Pligt, 2000). Because they anticipate this self-blame and wish to avoid it, we propose, people will be more averse to risk when it seems to require trust.

The aversive combination of anger and self-blame that trust violations may arouse has been described as *feeling like a sucker* (Vohs, Baumeister, & Chin, 2007). In one study, Vohs and colleagues had participants play multiple rounds of a prisoner's dilemma game (see Camerer, 2003) with a confederate who either defected after cooperating for two rounds, thus causing participants to lose points, or never defected. Participants who were victimized by defection felt more betrayed and spontaneously expressed more self-blame than participants who were not. While this study demonstrates that feelings of betrayal and self-blame sometimes co-occur, it leaves unclear whether or not participants would have experienced equivalent self-blame if they had lost points in a way that did not seem like a betrayal.

Our focus on the self-blame component of feeling like a sucker relates to theory and research regarding the role of anticipated regret in decision-making (e.g., Bell, 1982; Loomes & Sugden, 1982; Miller & Taylor, 1995; Zeelenberg, 1999). How painful people predict a loss will feel, and hence how hesitant they are to risk incurring it, depend in part on properties of the potential loss that are unrelated to its material magnitude. For example, people expect to regret the same negative outcome more when it will be easy to imagine it not having occurred (Gilbert, Morewedge, Risen, & Wilson, 2004; Kahneman & Tversky, 1982; Kahneman & Miller, 1986; Miller & Taylor, 1995), when it will have resulted from a choice that was difficult to make or

justify (Connolly & Zeelenberg, 2002; Simonson, 1992; Zeelenberg & Pieters, 2007), or when they will learn the outcome of an alternative course of action (e.g., Zeelenberg & Beattie, 1997; Zeelenberg & Pieters, 2004). Similarly, we propose that people's willingness to risk is influenced by their belief that losses of the same material magnitude will produce greater self-blame when the losses represent trust violations.

Prior findings support the claim that the fear of "playing the sucker" can have important motivational consequences. One such consequence arises in social dilemmas – situations in which groups benefit from mutual cooperation, but in which individual group members have an incentive to "free ride" on others' cooperation rather than cooperating themselves (e.g., Hardin, 1968). Free-riding arises in these situations, in part, from the fear of being a sucker who enables others to free ride (Orbell & Dawes, 1981; Rapoport & Eshed-Levy, 1989). In a demonstration of this phenomenon, termed the *sucker effect* (Kerr, 1983), participants performed a physically effortful task concurrently with a confederate, knowing that they would both receive a monetary reward for every trial on which at least one of them met a success criterion. The confederate then failed on every trial. Participants responded by reducing their effort on the task when the confederate's performance on a practice trial had suggested that he was capable of succeeding but preferred to free ride on participants' efforts. By contrast, participants did not reduce effort when the confederate's prior performance had suggested that he was incapable of succeeding. One explanation for the reduction of effort, consistent with the present framework, is that participants preferred to fail than to risk feeling exploited again. Yet the extent to which anticipated affect was responsible for this phenomenon is unclear. Rather than stemming from fear of feeling like a sucker, reduced effort could have represented an attempt to pressure the confederate to exert more effort on future trials, given that he seemed capable of doing so. The

nature of participants' experienced affect is also unclear. Did the confederate's poor performance after successful trials merely arouse anger, surprise, or confusion, or did participants experience self-blame as well? Subsequent research on the sucker effect did not examine the nature and role of anticipated or experienced affect either (Jackson & Harkins, 1985; Mulvey & Klein, 1998; Robbins, 1995; Schnake, 1991).

The present research extends prior work on the sucker effect by addressing these issues and answering new questions. Whereas prior work examined the consequences of coactors' expected and actual task performance on individuals' own task performance (Jackson & Harkins, 1985; Kerr, 1983; Mulvey & Klein, 1998; Robbins, 1995; Schnake, 1991), we examined the consequences of potential and actual trust violations on decisions about risk. In this context, we sought to demonstrate the motivational potency of trust violations in three ways. First, we examined whether people make more conservative risk decisions when risk requires trust. In so doing, we went beyond prior work by explicitly assessing the role of anticipated emotion, focusing specifically on anticipated self-blame (Studies 1 and 2). Additionally, whereas other theorists have emphasized that the fear of feeling like a sucker should be especially acute when exploitation seems likely (Vohs et al., 2007), we strove to isolate the effect of anticipated self-blame from the actual and perceived likelihood of incurring a loss.

A second way in which we sought to demonstrate the motivational potency of trust violations was by examining whether people actually do respond with greater self-blame to trust violations than to otherwise economically identical losses. It seemed important to examine experienced affect in addition to anticipated affect given evidence that some common beliefs about which negative outcomes will arouse the most regret are inaccurate (Crawford, McConnell, Lewis, & Sherman, 2002; Fernandez-Duque & Landers, 2008; Gilbert et al., 2004;

Sevdalis & Harvey, 2007). In Study 3, therefore, we examined people's actual reactions to trust violations resulting from risk decisions, taking steps to distinguish self-blame and perceived exploitation from mere surprise (contrast with Kerr, 1983). Furthermore, unlike prior work (Vohs et al., 2007), we compared self-blame after a trust violation to self-blame after an equivalent loss that did not violate trust.

Third, we examined how repeated risk decisions might reveal the motivational potency of trust violations. To the extent that people are highly motivated to avoid trust violations, incurring one should make them especially reluctant to risk incurring another. For example, after a friend reneges on a loan, one might be reluctant to loan money to a different friend, even if one believed that this first trust violation did not make a second one any more probable. By contrast, to the extent that people are relatively less motivated to avoid losses unrelated to trust, incurring a non-trust-violating loss might not have the same impact on future decisions. In Study 3, we tested the prediction that incurring a trust violation would increase aversion to subsequent risks requiring trust relative to risks not requiring trust – a "once bitten, twice shy" effect.

## Study 1: Investing

Study 1 tested the prediction that people would be less willing to invest money in a company when it risked failure due to fraud versus lack of consumer demand, even though the odds of failure were equal in both cases. Risking loss due to fraud requires trust because it makes one vulnerable to the selfish intentions of others, whereas risking loss due to consumer demand does not require trust. To address the possibility that the predicted effect would stem from aversion to others profiting from one's loss rather than from aversion to trust violations, Study 1 also included a condition in which a lost investment resulting from low consumer

demand would benefit others. We predicted that this condition would be insufficient to reduce willingness to invest. Finally, Study 1 tested the mediating role of anticipated self-blame.

## *Method*

### *Participants*

Ninety-seven university students (65% female; *M* age = 19.98, *SD* = 1.51) completed Study 1 in conjunction with unrelated studies for which they received $8. (One additional participant was excluded for providing incomplete data).

### *Procedure*

Participants completed a survey entitled "Financial Decision-Making." We first measured individual differences in risk preferences. All participants imagined that they had $100 to invest, and indicated in increments of $10 how much they would invest in a company whose only risk was that it "could fail because market forces prevent it from realizing its full potential for growth." Ostensibly, there was a 75% chance that the value of an investment would remain unchanged after one year, a 10% chance that it would double, and 15% chance that it would be lost entirely.

Next, participants imagined that they had never heard of the first investment opportunity and were asked to consider an investment in a different company. The "only risk" of this investment was that "the company could fail because consumers do not purchase enough of the company's products" (*trust-not-required* condition) or that "the company could be fake, set up by con artists to steal investors' money" (*trust-required* condition). A third condition (*others profit*) was identical to the *trust-not-required* condition, but also stated that, according to the investment agreement, the owners of the company would "keep your investment as a profit" if the company failed. In all conditions, the following probabilities were ostensibly "entirely

accurate" regarding the value of an investment in the company after one year: an 80% chance that it would not have changed, a 15% chance that it would have doubled, and a 5% chance that it would have been lost completely due to realization of the primary risk (i.e., fraud vs. low consumer demand).

As a measure of anticipated self-blame, participants indicated how they would react to losing a $100 investment in the company: how much they would regret their investment decision, feel upset with themselves, kick themselves, and feel foolish (1 = *Not at all*; 7 = *Extremely*; $\alpha = .88$). Finally, participants indicated in increments of $10 how much of their $100 they would invest in the company.

## *Results*

### *Primary analysis*

A one-way ANCOVA indicated that, controlling for the measure of individual differences in risk preference (i.e., investment in the first company), the amount invested in the second company differed as a function of the manipulation, $F(2, 93) = 4.97$, $p < .001$. Planned orthogonal contrasts revealed the predicted pattern. Participants said they would risk less money when the investment required trust ($M = \$37.35$, $SD = 30.68$) than in the two conditions in which trust was not required, $F(1, 93) = 8.72$, $p < .005$, $d = .61$. Also as predicted, participants said they would invest an equivalent amount in the two conditions that did not require trust, regardless of whether others would profit from a loss ($M = \$56.21$, $SD = 33.85$) or not ($M = \$63.24$, $SD = 27.05$), $F(1, 93) = .96$, *ns*, $d = .20$.

### *Mediation analysis*

We next tested the mediating role of anticipated self-blame, including the mean-centered covariate from the ANCOVA in all analyses (Baron & Kenny, 1986; see Figure 1). A first

regression equation confirmed that participants anticipated more self-blame when the investment required trust (coded as 2) than when it did not (other two conditions each coded as -1), $b = .36$, $t(94) = 3.76$, $p < .001$, $d = .78$.[i] A second equation in which self-blame was mean-centered showed that the more self-blame participants anticipated, the less money they wanted to invest, $b = -7.61$, $t(93) = 3.72$, $p < .001$, $\beta = -.34$, and that controlling for anticipated self-blame reduced the effect of the manipulation on investment amount to non-significance, $b = -3.38$, $t(93) = 1.64$, $p > .10$, $d = .34$ (compare to $b = -6.15$, $d = .61$ when not controlling for self-blame). This reduction was significant by the Sobel test, $z = 2.65$, $p < .01$. In sum, the effect of the manipulation on participants' willingness to invest was significantly mediated by anticipated self-blame.

*Discussion*

Participants in Study 1 predicted that they would blame themselves more for losing an investment due to fraud (a trust violation) versus low consumer demand (not a trust violation). As a result, they were more reluctant to invest in a company that risked failure due to fraud versus low consumer demand, even though the financial parameters of both investments (i.e., the probability and magnitude of losses and gains) were identical. The knowledge that company owners would profit from participants' loss was not sufficient to produce this relative reluctance to invest, suggesting that perceiving the investment as requiring trust was necessary.

A potential limitation of Study 1 is that even though we emphasized in all conditions that the risk of loss was 5%, participants may still have assumed that loss due to fraud was more probable than loss due to consumer demand. An additional limitation is that anticipated self-blame may have influenced decisions only because, in the service of testing for mediation, we

prompted participants to think about it before collecting the dependent measure (Crawford et al., 2002; Hetts, Boninger, Armor, Gleicher, & Nathanson, 2000; Simonson, 1992).

To address these two limitations, we randomly assigned 81 students to either the *trust-required* or the *trust-not-required* conditions of Study 1, with two key modifications. First, participants were not asked to anticipate self-blame. Second, for both the individual difference measure and the dependent measure, participants indicated risk aversion by filling in the blank in a sentence stating that they would invest a fixed sum of money in the company ($500 or $1,000 – a variation that did not affect results) "only if there were at least a __% chance that the company would *not* fail" (9-point scale from 10% to 90%; negative skew corrected by reversing, log-transforming, and re-reversing the scale). We reasoned that this new measure would be sensitive to the perceived psychological cost of loss, but not to the perceived likelihood of loss occurring.

We also made two superficial changes to the Study 1 procedure. First, the *trust-required* condition stated that the only risk of the investment was that "the company could fail because its owners trick investors and steal their money." Second, for the company described before the manipulation (i.e., for the individual difference measure), investors would lose half their money if the company failed and double their money if it did not, whereas for the company described after the manipulation (i.e., for the dependent measure), investors would lose all their money if the company failed and quadruple their money if it did not.

Conceptually replicating the results of Study 1 with a similar effect size, participants indicated greater risk aversion in the trust-required condition than in the trust-not-required condition (log-transformed *M*s = 3.73 and 3.44, respectively; *SD*s = .67 and .69), $F(1, 78) = 6.25$, $p = .01$, $d = .57$, controlling for individual differences in risk preference. This result suggests that anticipated self-blame arose and influenced the risk decisions spontaneously.

## Study 2: Economic Game

Participants in Study 1 expressed greater reluctance to invest money when doing so could result in a trust violation, because they anticipated that a trust-violating loss would spark more self-blame. In Study 2, we sought to conceptually replicate this effect under what we thought would be the minimal conditions sufficient to trigger concerns about trust violations. In addition to shedding light on when anticipated self-blame affects risk decisions, the paradigm used in Study 2 was meant to provide a cleaner manipulation of the potential for trust violations than Study 1 did. Participants made risky decisions in an economic game whose outcome hypothetically depended either on another participant's decision or on a random-number generator. We reasoned that participants would perceive risk as requiring trust if it made them vulnerable to another person, but not if it made them vulnerable to a randomizing device (see also Bohnet & Zeckhauser, 2004). As in Study 1, we predicted that participants would make less risky decisions when trust was at stake than when it was not, and that anticipated self-blame would again mediate this effect.

### *Method*

#### *Participants*

One hundred forty-five participants (66% female; 68% White, 17% Asian, 4% Latino, 4% Black; *M* age = 32.60 years, *SD* = 10.82, range = 18 to 64), recruited from an online subject pool, completed this Web study in exchange for a $5 gift certificate to an online retailer. (An additional 13 participants declined to respond to the primary measures, and one was excluded for taking 6.10 *SD*s above the mean time to complete the study; attrition did not significantly differ between conditions, $\chi^2$ [1] = 1.54, *ns*). Highest level of educational attainment varied: 26% held advanced degrees, 21% held 4-year college degrees, 13% held 2-year college degrees, 14% were

currently enrolled in college, and 24% held a high school-level diploma and were not currently enrolled in college.

*Procedure*

In a "first study," participants imagined choosing between a sure thing of $2 versus a chance to receive either $4 or nothing, and indicated, in increments of 10% ranging from 10 to 100, the worst odds of receiving the $4 at which they would choose the gamble. We used this item as a covariate to control for individual differences in risk preference.

In a "second study," participants learned how to play a two-player trust game (referred to as "The Investment Game" and adapted from Bohnet & Zeckhauser, 2004; see Camerer, 2003). Both players would receive 10 points (worth $.60 each). As Player 1 (P1), participants would first decide whether to "keep" their points, in which case both players would end the game with 10 points, or to "loan" the points to Player 2 (P2), in which case P2 would also receive 10 "bonus points" and the outcome of the game would depend on P2's move. This move would be either an "even split," in which case both players would end the game with 15 points, or an "uneven split," in which case participants would end with 8 points and P2 would end with 22 points. The diagram that summarized these rules for participants is shown in Figure 2.

All participants were told that P2 was another participant, but we manipulated whether P2's move would be determined by his or her own decision (*trust-required* condition) or by a "random number" that was "generated by the website" (*trust-not-required* condition). In the *trust-not-required* condition, we emphasized that P2 had "no control" over how many points P1 would receive.

*Measures*

*Comprehension check*. Seven questions assessed participants' understanding of the rules of the game (i.e., which combination of moves would be best and worse for P1; how many points each player would end the game with for each combination of moves).

*Self-blame*. Participants used the four self-blame items from Study 1 to indicate how they would feel if, after loaning their points, they learned that P2's move was the uneven split (1 = *Not al all*; 9 = *Extremely*; $\alpha = .92$).

(In Studies 2 and 3, participants also indicated how disappointed and upset they would feel about the game's outcome. These items correlated highly with the self-blame scale, and including them in the scale produced analogous results. To maintain consistency with the cleaner self-blame measure used in Study 1, however, Studies 2 and 3 omitted these items from the scale).

*Manipulation checks*. Participants estimated how much they would feel that their trust had been violated and that they had been taken advantage of if the game ended with an uneven split (1 = *Not at all*; 9 = *Extremely*; $\alpha = .94$). Then they indicated how much of a choice P2 would have in determining his or her move (1 = *No choice at all,* 3 = *Some choice*, 5 = *Completely Player 2's choice*).

*Risk aversion*. Finally, participants were asked to imagine that they were playing a single round of the game as P1, and to fill in the blank to make the following sentence true: "I would choose to loan my points to Player 2 only if there were at least a __% chance that [Player 2 would choose / the website would randomly select] an even split."

## *Results*

### *Comprehension and manipulation checks*

Overall, participants understood how to play the game, answering an average of 6.70 out of 7 questions correctly ($SD$ = .95), which did not differ by condition, $t(143) < 1$. As expected, participants perceived Player 2 as having more of a choice in determining his or her move in the *trust-required* condition ($M$ = 4.70, $SD$ = .65) than in the *trust-not-required* condition ($M$ = 1.19, $SD$ = .56), $t(138) = 33.72$, $p < .0001$ (four participants did not respond to this question), and indicated that a loss would make them feel more like they had been taken advantage of and had their trust violated in the *trust-required* condition ($M$ = 6.35, $SD$ = 2.30) than in the *trust-not-required* condition ($M$ = 2.54, $SD$ = 2.10), $t(142) = 10.29$, $p < .0001$.

*Risk aversion*

As predicted, participants were more risk-averse when P2's move would be determined by another participant (*trust-required* condition; $M$ = 65.00%, $SD$ = 22.38) versus by chance (*trust-not-required* condition; $M$ = 58.33%, $SD$ = 16.88), $F(1, 141) = 3.78$, $p = .05$, $d = .33$ in an ANCOVA that controlled for individual differences in risk preference.[ii]

*Mediation analysis*

We predicted that anticipated self-blame would mediate the effect of the manipulation on risk aversion. The regression equations that tested for mediation included a dummy code for condition (*trust-required* = 1, *trust-not-required* = 0; Baron & Kenney, 1986) and the covariate from the ANCOVA (mean-centered). Results are shown in Figure 3. A first equation showed that participants anticipated greater self-blame when risk required trust, $b = 1.17$, $t(141) = 3.23$, $p < .005$, $d = .54$. A second equation showed that greater anticipated self-blame (mean-centered) predicted greater risk aversion, $b = 3.30$, $t(140) = 4.87$, $\beta = .37$, $p < .001$, and that the effect of condition on risk aversion, controlling for anticipated self-blame, dropped to non-significance, $b = 2.25$, $t(140) = .74$, *ns*, $d = .13$ (compare to $b = 6.13$, $d = .33$ when not controlling for self-

blame). This decrease was significant by the Sobel test, $z = 2.69$, $p < .01$. Thus, anticipated self-blame significantly mediated the effect of the manipulation.

*Discussion*

Study 2 produced two main results. First, it showed that participants were less willing to tolerate risk in an economic game when the other player's move was determined by another person than when it was determined by chance (Bohnet & Zeckhauser, 2004). Risk only requires trust when it involves making oneself vulnerable to the behavior of another person. Second, through mediation analysis, Study 2 showed that anticipated self-blame was responsible for the relative aversion to trust-related risk. The paradigm used in Study 2 provided a cleaner demonstration of the link between decision conservativeness and fear of trust violations than Study 1, and the results shed light on the minimal conditions sufficient to spark concerns about trust violations and self-blame.

## Study 3: Once Bitten, Twice Shy

Studies 1 and 2 examined the motivational consequences of anticipated reactions to trust violations. The main purpose of Study 3 was to examine the motivational consequences of *experienced* trust violations. After experiencing a loss due to chance, people often make riskier subsequent decisions because of a motivation to recoup their loss (Leopold, 1978; Thaler & Johnson, 1990). But the experience of a trust-violating loss, we propose, can offset this motivation with the motivation to avoid a second trust violation. We thus hypothesized in Study 3 that a) as in Studies 1 and 2, participants would make more conservative decisions when risk required trust than when it did not, and that b) building on Studies 1 and 2, experiencing a loss that violated trust would increase this relative conservatism – a "once bitten, twice shy" effect. We tested these hypotheses in a two-round trust game whose outcome was ostensibly determined

either by chance (*trust-not-required* condition) or by another participant (*trust-required* condition). A secondary goal of Study 3 was to examine whether losses really do arouse more self-blame when they violate trust, as participants in Studies 1 and 2 expected that they would. We thus measured how much self-blame participants felt after losing points on the first round of the game.

## *Method*

### *Participants*

We recruited 51 participants (57% female; 56% White, 32% Asian; *M* age = 30.42 years, *SD* = 9.31; range = 18 to 54) from a subject pool to complete Study 3 on the Web in exchange for a $5 gift certificate to an online retailer. Highest educational attainment varied: 34% had advanced degrees, 10% had 4-year college degrees, 14% had 2-year college degrees, 20% were currently enrolled in college, 18% held a high school diploma and were not currently enrolled in college, and the remainder held less than a high school diploma.

### *Procedure*

Participants learned that they would be playing "multiple rounds" of "The Investment Game" with other participants. Each point in the game was ostensibly worth $1 and could be redeemed for a gift certificate by the participant who won a drawing (in actuality, because the outcome of the game was rigged, all participants were entered into a drawing for a $50 gift certificate). To increase the plausibility that participants had been partnered with others, we built several delays into the study during which the website ostensibly waited for others to log on or respond to questions. Throughout the study, the instructions emphasized that participants would play each round of the game with a different person.[iii]

The instructions for the game were similar to those in Study 2. Participants, who had ostensibly been randomly assigned to be P1, and P2 would each begin each round of the game with 32 points. Participants would decide how many of their points to invest, and any points invested would double, thus making a profit. Then one of two outcomes would occur. The *favorable* outcome (described as "Outcome A") was that participants would receive their entire investment back, and split the profit with P2. Thus, participants and P2 would end the round with an equal number of points. The *unfavorable* outcome ("Outcome B") was that participants would only receive 80% of their investment back, and P2 would get the entire profit. Thus, P2 would end the round with more points than participants.

*Manipulation*. In the *trust-required* condition, participants learned that P2 would select between the two outcomes. Specifically, after participants chose how many points to "give to Player 2" as their investment, P2 would choose how many to "give back" to them. In the *trust-not-required* condition, by contrast, participants would decide how many points to "designate" as their investment, after which "the website [would] generate a random number" to determine how many points they and Player 2 "each get."

*Comprehension check.* After reading about the two possible outcomes of a hypothetical round in which P1 invested 20 points, participants completed 15 comprehension-check questions (e.g., how many points each outcome would leave them with if they had invested 10 points). Participants were informed of any items they had answered incorrectly, and were urged (but not required) to correct erroneous responses.

*Establishing the odds*. To equate the probability of the unfavorable outcome occurring across conditions, we told participants in the *trust-required* condition that

> Previous research on the Investment Game has determined that, on each round, the probability that Player 2 will choose Outcome B is 30%. That is, 3 out of every 10 people who play this game as Player 2 choose Outcome B.

Participants in the *trust-not-required* condition were told instead that

> The website has been programmed so that, on each round, the probability that Outcome B will be randomly selected is 30%. That is, 3 out of every 10 times the computer plays this game, it randomly selects Outcome B.

To ensure that participants read this message carefully, it remained on their screen for 60 seconds after they had indicated that they were ready to continue, ostensibly while the website waited for P2.

*Investment amount.* Participants chose how many points, between 10 and 32 in increments of 2, to invest. Setting the minimum at 10 allowed us to give all participants feedback that they had lost a non-trivial number of points.

*Subjective probability measure*. Participants indicated how confident they were and how much they trusted that Round 1's outcome would be favorable, as well as how much they suspected that it was unfavorable (1 = *Not at all*, 9 = *Extremely*; last item reversed; $\alpha = .88$).

*Loss feedback.* Participants next learned that the unfavorable outcome had been selected, and viewed the number of points with which they and P2 had each ended the round. Participants always ended the game with fewer points than P2 – a loss from their initial endowment.

*Self-blame measure and manipulation check.* The four-item scale used in Study 2 assessed experienced self-blame ($\alpha = .96$), and the two manipulation check items from Study 2 assessed perceptions of the loss as a trust-violation ($\alpha = .94$).

*Round 2*. Participants were then informed that they would play another round of the game with a different participant playing P2. As in Round 1, participants began with 32 points and chose how much to invest (minimum 10 points).

*Other measures.* Participants provided demographics, responded to free-response suspicion-checks ("What did you think the researchers were hoping to find?" and "Did you think anything was strange or unusual while taking these studies?"), and read a debriefing form.

## *Results*

### *Exclusions*

We excluded participants (five in each condition) who provided incomplete data, who expressed suspicion that P2 did not exist or that the game was rigged, or who took unusually long to complete the study (i.e., one participant who took 4.84 *SD*s above the mean time).

### *Comprehension and manipulation checks*

Participants understood how to play the game, answering on average 14.10 of the 15 questions correctly ($SD = 2.49$), with no differences in comprehension arising between conditions, $t(39) = .26$, *ns*. Indicating that the manipulation had its intended effect, participants experienced the loss as more of a trust violation in the *trust-required* condition ($M = 5.95$, $SD = 2.54$) than in the *trust-not-required* condition ($M = 3.26$, $SD = 2.78$), $F(1, 38) = 10.86$, $p < .005$, controlling for the number of points lost.

### *Investment amount*

Our primary hypotheses were that a) participants would invest fewer points when the round's outcome was ostensibly determined by another person (*trust-required* condition) than when it was determined by chance (*trust-not-required* condition), and that b) this difference would be especially pronounced in Round 2 (i.e., after participants had experienced a loss). As Figure 4 illustrates, the pattern of results fit the hypotheses. A 2 (trust: required vs. not) X 2 (Round: 1 vs. 2) ANOVA with repeated measures on the second factor revealed two significant main effects, indicating that participants risked fewer points when doing so required trust than

when it did not, $F(1, 39) = 8.26$, $p < .01$, and that they risked fewer points in Round 1 than in Round 2, $F(1, 39) = 8.84$, $p = .005$. More importantly, the interaction term was significant, $F(1, 39) = 5.11$, $p < .05$. Consistent with our first hypothesis, participants invested fewer points when trust was required than when it was not both in Round 1 (*M*s = 15.73 and 19.05, *SDs* = 3.92 and 6.48, respectively), $t(39) = 2.02$, $p = .05$ and in Round 2 (*M*s = 16.27 and 23.05, *SD*s = 5.90 and 7.87, respectively), $t(29) = 3.15$, $p < .005$. Consistent with our second hypothesis, this difference was especially pronounced in Round 2.[iv] Examining the interaction from a different perspective shows that when risk did not require trust, participants behaved in accordance with prior research, increasing their exposure to risk after incurring a loss, presumably in an effort to recoup it (Leopold, 1978; Thaler & Johnson, 1990), $t(18) = 3.13$, $p < .01$. By contrast, participants chose not to increase their risk exposure when doing so required trust, $t(21) = .61$, *ns*.

*Experienced self-blame*

Participants' decisions about how many points to invest suggest that, as in Studies 1 and 2, they predicted that a loss would feel worse when another person caused it. Additional results supported our hypothesis that an experienced loss would indeed elicit more self-blame when another person caused it. After incurring a loss in the first round of the game, participants expressed more self-blame in the *trust-required* condition ($M = 5.59$, $SD = 2.49$) than in the *trust-not-required* condition ($M = 4.07$, $SD = 2.20$). Importantly, this condition difference was significant even when controlling for the fact that participants had lost more points in the *trust-required* condition, $F(1, 38) = 6.83$, $p < .05$.[v]

*Subjective probability*

The effects reported above remained significant after controlling for the measure of subjective probability. Indeed, subjective probability did not significantly correlate with investment amount in either Round 1, $r(41) = .09$, *ns*, or Round 2, $r(41) = .22$, $p = .17$. We therefore found no support for the possibility that the pattern of results obtained for either investments or self-blame could be explained by differences in how likely participants felt a negative outcome would be to occur.

*Discussion*

Conceptually replicating the results of Studies 1 and 2, participants in Study 3 were less willing to expose themselves to the risk of loss when doing so required trust. This finding was obtained in a paradigm in which participants made and expected to learn the outcome of non-hypothetical decisions with potential financial consequences. The fact that we obtained this finding without prompting participants to consider how a loss would make them feel suggests, like the follow-up to Study 2, that concerns about self-blame arose and influenced judgment spontaneously.

The primary goal of Study 3 was to build on Studies 1 and 2 by examining the motivational consequences of *experienced* trust violations. Consistent with our hypothesis that incurring a trust-violating loss would motivate more risk aversion than a non-trust-violating loss, the disparity in risk aversion increased after participants experienced a loss. Specifically, whereas a loss that did not violate trust seems to have bolstered participants' willingness to risk a similar loss, presumably in order to recoup their points (Leopold, 1978; Thaler & Johnson, 1990), experiencing a trust violating loss produced no such increase in willingness to risk. Apparently, the motivation to avoid a second trust violation was enough to counteract the motivation to recoup the lost points.

Why did incurring a first trust violation increase the motivation to avoid incurring a second? One possibility is that participants predicted that a second trust violation would feel even worse than the first. This could arise from participants anticipating that allowing themselves to be "suckered" twice would make them seem more responsible for the violations – an intuition captured by the aphorism, "fool me once, shame on you; fool me twice, shame on me." A second possibility is that participants did not predict that a second trust violation would increase self-blame, but rather anticipated that blaming themselves the same amount for two different decisions would represent more cumulative self-blame than they could tolerate. Untangling these two possibilities must await future research. In either case, it appears that experiencing a trust-violating loss has different motivational consequences than experiencing a loss that does not violate trust.

Note that our measures of subjective probability did not support the alternative explanation that the loss in Round 1 made a loss in Round 2 seem more likely to occur. This alternative is rendered even more implausible by the fact that we told participants in both conditions that the probability of loss was 30%, and emphasized to participants in the *trust-required* condition that each round was determined by the actions of a different participant. The outcome of the first round should thus have been equally (un)informative about the outcome of the second round in both conditions.

A secondary goal of Study 3 was to examine whether losses actually would arouse more self-blame when they violated trust. The results revealed that they did, providing an additional demonstration of the motivational potency of trust violations, and suggesting that participants who anticipated greater self-blame in the trust-required conditions of the present studies were right to do so.

Taken together, the results of Study 3 reveal how people's reactions to risks that require trust unfold over time. People predict that they will kick themselves more for a loss when it represents a trust violation. Accordingly, they reduce their exposure to the risk of such a loss. When they experience a loss, they actually do feel more self-blame when the loss represents a trust violation, and then act as if they were particularly motivated to avoid a second trust violation.

## General Discussion

The present research demonstrated, in three ways, the motivational potency of viewing a risk decision as relevant to trust. First, we found evidence that people expect trust violations to result in greater self-blame than otherwise economically identical losses that do not violate trust and that, as a result, people make more conservative decisions about trust-related risks than about otherwise economically identical risks not related to trust. Specifically, Study 1 found that participants were more reluctant to risk incurring a loss caused by fraud (trust required) than one caused by consumer purchasing behavior (trust not required), while Studies 2 and 3 found a greater reluctance to risk incurring a loss caused by another person (trust required) than one caused by chance (trust not required). Studies 1 and 2 obtained direct meditational evidence that this reluctance to risk was explained by anticipated self-blame.

A second manifestation of the motivational potency of trust violations is that losses seem to arouse more self-blame when they violate trust. In Study 3, participants who lost an equivalent number of points in an economic game expressed more self-blame when the loss had been caused by another person as opposed to chance. Given this response to trust violations, people apparently have good reason to be wary of trust-related risks.

Study 3 also provided a third demonstration of the motivational potency of viewing risks as relevant to trust. After incurring a loss, participants seemed eager to take on even greater risk in a subsequent decision – unless the loss had violated trust. This "once bitten, twice shy" effect occurred even though the first loss was designed to be non-diagnostic of the odds that a second loss would occur. Apparently, the motivation to avoid a future loss was greater when a prior loss represented a trust violation.

*Alternative explanations*

When making risky decisions, people consider both the magnitude and the probability of potential losses and gains. We have argued that participants' reluctance to risk losses that would violate trust arose from the greater disutility of such losses (specifically, greater self-blame). Several observations argue against the alternative possibility that this reluctance to risk stemmed from perceptions that such losses were more probable. In Studies 1 and 3, we equated the probability of incurring a loss across conditions. Further, we measured subjective probability of loss in Study 3, and this measure was unable to explain the condition difference we observed. In Study 2, as well as the follow-up to Study 1, participants were willing to tolerate a higher probability of incurring a loss when loss would not violate trust – findings not easily explained by the subjective probability of a loss actually occurring. Finally, the meditational evidence in Studies 1 and 2 suggests that participants' concerns about the magnitude of loss fully accounted for the effects we observed. Thus, it seems highly unlikely that perceived probability of loss can explain our results.

We interpret our results as indicating that participants were more motivated to avoid loss when it would violate trust. A potential alternative is more cognitive than motivational. Framing economic games as social interactions versus games of chance can affect players'

strategies by eliciting different *decision heuristics* – rules of thumb that guide decisions without much thought (Abele, Bless, & Ehrhart, 2004; Weber, Kopelman, & Messick, 2004). In the *trust-not-required* condition of Studies 2 and 3, perhaps the knowledge that a random-number generator would determine the payoff cued a decision heuristic associated with games of chance, which made participants more willing to risk. Our findings cast doubt on this explanation. First, we observed the same pattern of results in the context of investing in stocks when the *trust-not-required* condition involved risking loss due to low consumer demand (Study 1 and its follow-up). It seems unlikely that this manipulation would activate the same decision heuristic as the manipulation in Studies 2 and 3. Second, a purely cognitive explanation would struggle to explain the mediational data we obtained in Studies 1 and 2. These data provide strong evidence for the motivational mechanism we have proposed.

### *Contributions to the literature*

Our studies enrich prior theory and research on trust and on the sucker effect. Whereas prior work is consistent with the idea that people expect trust violations to add disutility to loss (Bohnet & Zeckhauser, 2004; Kerr, 1983), direct assessments of this fear and its role in decision-making have been lacking, allowing for alternative interpretations of the relevant findings. Our data fill this gap and advance a more precise understanding of the nature of the feared disutility, supporting the previously untested idea that a reluctance to trust can arise from anticipated negative affect directed towards the self (Vohs et al., 2007) and not just negative affect directed towards the trust violator (Koehler & Gershoff, 2003). The present research also extends prior literature by examining reactions to the *experience* of incurring a trust-violating loss. The "once bitten, twice shy" effect illustrates that the impact of a loss on subsequent decision-making can depend on whether or not the loss is perceived as a trust violation. Additionally, examining

experienced trust violations allowed us to examine the question, unaddressed by prior research, of whether aversion to risks requiring trust stems from biased predictions of how trust violations would make one feel. We found that participants did feel more self-blame in response to losses that violated trust – in fact, just as much more as other participants had predicted (see Footnote 5).

The accuracy of participants' predictions has implications for the affective forecasting literature, which has focused on systematic errors people make when predicting the emotional impact of different events (see Wilson & Gilbert, 2003, for a review). People seem to have different lay theories about which circumstances will amplify regret. Some of these theories are inaccurate, such as one that assumes that losses feel worse when they almost did not occur (Gilbert et al., 2004; Sevdalis & Harvey, 2007; see also Crawford et al., 2002; Fernandez-Duque & Landers, 2008). But in other domains, the lay theories seem to be correct, as shown by people's accurate estimates of how much regret they will feel in response to losing simple gambles (Mellers, Schwartz, & Ritov, 1999). Our findings suggest that trust is a domain in which the lay theory aligns with reality. This alignment may be coincidental, but it may also speak to the psychological importance of the domain of trust. For example, accuracy might arise from a tendency to ruminate about past experiences with trust violations more than past experiences with other kinds of loss. To explore this possibility, future research should test whether people can accurately predict their reaction to different operationalizations of trust violations. It also remains to be seen whether people overestimate the *duration* of self-blame, as they do with other emotions in other domains (Wilson & Gilbert, 2003).

*When will people trust?*

In the present studies, the risk of a trust violation elicited more conservative decisions, but several processes may sometimes counteract this phenomenon by offsetting the anticipated pain of a trust violation. First, risks that require trust are often accompanied by promises and assurances that may increase confidence that the risk will pay off. Second, social norms prescribing trust may sometimes make the anticipated pain of being too cynical outweigh the anticipated pain of feeling like a sucker (Deutsch, 1973). Third, the anticipated pleasures of trusting wisely – such as affirming one's connectedness to others – may sometimes counteract the anticipated pain of trusting too well. Fourth, people may sometimes trust automatically without considering the costs and benefits of gains and losses (Messick & Kramer, 2001). Each of these processes seems most likely to operate within close relationships in which trust is assumed, as opposed to the kind of interactions with strangers examined in the present studies. We suspect, however, that even in the context of close relationships, the concern about feeling like a sucker can sometimes make people hesitate before trusting.

*Why are trust violations so aversive?*

There could be a number of reasons why people tend to kick themselves more when a loss represents a trust violation. One possibility is that people assume that permitting their own exploitation tends to invite further exploitation. Falling prey to one con artist may make one a desirable target for other con artists, whereas losing money in the stock market does not similarly increase one's vulnerability to future investment losses. A related possibility is that trust violations have especially self-threatening implications. People may fear that having made themselves vulnerable to exploitation may signal to themselves and others that they are the *kind of person* who can be exploited (i.e., a sucker rather than merely someone who has exercised poor judgment). A different kind of explanation stems from the idea that heightened aversion to

trust violations is functional in that it motivates people to avoid falling prey to the exploitative intentions of others (cf. Zeelenberg & Pieters, 2006). This aversion could have resulted from adaptation to evolutionary pressures (Cosmides, 1989; Price, Cosmides, & Tooby, 2002), but could also arise from the experience of navigating comtemporary society, in which it continues to be crucial to trust wisely and not too well.

### *Conclusion*

People are more motivated to avoid material loss when they perceive such loss as a trust violation. Trust-violating losses are painful not only because they affect one's wealth or material resources, or incite anger at others, but also because they arouse self-blame. Anticipating this self-blame, people are more reluctant to expose themselves to risks that require trust than to otherwise identical risks that do not – even when both types of risk are equally likely to end in loss and to profit other people. Incurring a trust violation further magnifies this relative reluctance to risk.

References

Abele, S., Bless, H., & Ehrhart, K. (2004). Social information processing in strategic decision-making: Why timing matters. *Organizational Behavior and Human Decision Processes, 93*, 28-46.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.

Bell, D. E. (1982). Regret in decision-making under uncertainty. *Operations Research, 30*, 961-981.

Bohnet, I., & Zeckhauser, R. (2004). Trust, risk and betrayal. *Journal of Economic Behavior & Organization, 55*, 467-484.

Camerer, C. F. (2003). *Behavioral game theory*. Princeton, NJ: Princeton University Press.

Connolly, T., & Zeelenberg, M. (2002). Regret in decision making. *Current Directions in Psychological Science, 11*, 212-216.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition, 31*, 187-276.

Crawford, M. T., McConnell, A. R., Lewis, A. C., & Sherman, S. J. (2002). Reactance, compliance, and anticipated regret. *Journal of Experimental Social Psychology, 38*, 56-63.

Deutsch, M. (1973). *The resolution of conflict*. New Haven: Yale University Press.

Fernandez-Duque, D., & Landers, J. (2008). "Feeling more regret than I would have imagined:" Self-report and behavioral evidence. *Judgment and Decision Making, 3*, 449-456.

Gilbert, D. T., Morewedge, C. K., Risen, J. L., & Wilson, T. D. (2004). Looking forward to looking backward: The misprediction of regret. *Psychological Science, 15*, 346-350.

Gillers, S. (2009, March 12). The sin lay in the personal betrayals. *The New York Times.* Retrieved February 11, 2010 from http://roomfordebate.blogs.nytimes.com/2009/03/12/is-madoff-wall-streets-greatest-villain/?pagemode=print

Hardin, G. (1968). The tragedy of the commons. *Science, 162*, 1243-1248.

Hetts, J. J., Boninger, D. S., Armor, D. A., Gleicher, F., & Nathanson, A. (2000). The influence of anticipated counterfactual regret on behavior. *Psychology and Marketing, 17*, 345-368.

Jackson, J. M., & Harkins, S. G. (1985). Equity in effort: An explanation of the social loafing effect. *Journal of Personality & Social Psychology, 49*, 1199-1206.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review, 93*, 136-153.

Kahneman, D., & Tversky, A. (1982). The psychology of preferences. *Scientific American, 246*, 160-173.

Kerr, N. L. (1983). Motivation losses in small groups: A social dilemma analysis. *Journal of Personality & Social Psychology, 45*, 819-828.

Koehler, J. J., & Gershoff, A. D. (2003). Betrayal aversion: When agents of protection become agents of harm. *Organizational Behavior and Human Decision Processes, 90*, 244-261.

Leopold, A. (1978). Risk preference in consecutive gambling. *Journal of Experimental Psychology: Human Perception and Performance, 4*, 521-528.

Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal, 92*, 805-824.

Mellers, B. A., Schwartz, A., & Ritov, I. (1999). Emotion-based choice. *Journal of Experimental Psychology: General, 128*, 332-345.

Messick, D. M., & Kramer, R. M. (2001). Trust as a form of shallow morality. In K. S. Cook (Ed.), *Trust in society* (pp. 89-117). New York: Russell Sage Foundation.

Miller, D. T., & Taylor, B. R. (1995). Counterfactual thought, regret, and superstition: How to avoid kicking yourself. In N. J. Roese & J. M. Olson (Eds.), *What might have been: The social psychology of counterfactual thinking* (pp. 305-331). Hillsdale, NJ: Erlbaum.

Mulvey, P. W., & Klein, H. J. (1998). The impact of perceived loafing and collective efficacy on group goal processes and group performance. *Organizational Behavior and Human Decision Processes, 74*, 62-87.

Orbell, J., & Dawes, R. (1981). Social dilemmas. In G. M. Stephenson & J. M. Davis (Eds.), *Progress in applied social psychology* (Vol. 1). Chichester, England: Wiley.

Price, M. E., Cosmides, L., & Tooby, J. (2002). Punitive sentiment as an anti-free rider psychological device. *Evolution and Human Behavior, 23*, 203-231.

Rapoport, A., & Eshed-Levy, D. (1989). Provision of step-level public goods: Effects of greed and fear of being gypped. *Organizational Behavior and Human Decision Processes, 44*, 325-344.

Robbins, T. L. (1995). Social loafing on cognitive tasks: An examination of the "sucker effect." *Journal of Business and Psychology, 9*, 337-342.

Rotter, J. B. (1967). A new scale for the measurement of interpersonal trust. *Journal of Personality, 35*, 651-665.

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

Schnake, M. E. (1991). Equity in effort: The "sucker effect" in co-acting groups. *Journal of Management, 17*, 41-55.

Sevdalis, N., & Harvey, N. (2007). Biased forecasting of postdecisional affect. *Psychological Science, 18*, 678-681.

Simonson, I. (1992). The influence of anticipating regret and responsibility on purchase decisions. *Journal of Consumer Research, 19*, 105-118.

Simpson, J. A. (2007). Foundations of interpersonal trust. In A. W. Kruglanski & E. T. Higgins (Eds.), *Social psychology: A handbook of basic principles* (2nd ed., pp. 587-607). New York: Guilford Press.

Smith, A. (2009, September 28). Madoff Ponzi prosecutors: Still trying to tally loss. *CNN Money.com.* Retrieved February 11, 2010 from http://money.cnn.com/2009/09/24/news/economy/madoff_ponzi/index.htm

Thaler, R. H., & Johnson, E. J. (1990). Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management Science, 36*, 643-660.

Vohs, K. D., Baumeister, R. F., & Chin, J. (2007). Feeling duped: Emotional, motivational, and cognitive aspects of being exploited by others. *Review of General Psychology, 11*, 127-141.

Weber, J. M., Kopelman, S., & Messick, D. M. (2004). A conceptual review of decision making in social dilemmas: Applying a logic of appropriateness. *Personality and Social Psychology Review, 8*, 281-307.

Wilson, T. D., & Gilbert, D. T. (2003). Affective forecasting. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. *35*, pp. 345-411). San Diego, CA: Academic Press.

Zeelenberg, M. (1999). Anticipated regret, expected feedback and behavioral decision making. *Journal of Behavioral Decision Making, 12*, 93-106.

Zeelenberg, M., & Beattie, J. (1997). Consequences of regret aversion 2: Additional evidence for effects of feedback on decision making. *Organizational Behavior and Human Decision Processes, 72*, 63-78.

Zeelenberg, M., & Pieters, R. (2004). Consequences of regret aversion in real life: The case of the Dutch postcode lottery. *Organizational Behavior and Human Decision Processes, 93*, 155-168.

Zeelenberg, M., & Pieters, R. (2006). Feeling is for doing: A pragmatic approach to the study of emotions in economic behavior. In D. De Cremer, M. Zeelenberg & J. K. Murningham (Eds.), *Social psychology and economics* (pp. 117-137). Mahwah, NJ: Erlbaum.

Zeelenberg, M., & Pieters, R. (2007). A theory of regret regulation 1.0. *Journal of Consumer Psychology, 17*, 3-18.

Zeelenberg, M., van Dijk, W. W., Manstead, A. S. R., & van der Pligt, J. (2000). On bad decisions and disconfirmed expectancies: The psychology of regret and disappointment. *Cognition and Emotion, 14*, 521-541.

Author note

Daniel A. Effron, Department of Psychology, Stanford University, and Dale T. Miller, Graduate School of Business, Stanford University.

We thank Max Halvorson for research assistance.

Correspondence may be addressed to the first author via electronic mail to effron@stanford.edu.

Notes

[i] A contrast comparing the *trust-not-required* and the *others profit* conditions, when included in the equation, was non-significant, $b = .05$, $t(94) = .27$, *ns*.

[ii] With equal variance not assumed, $p < .05$.

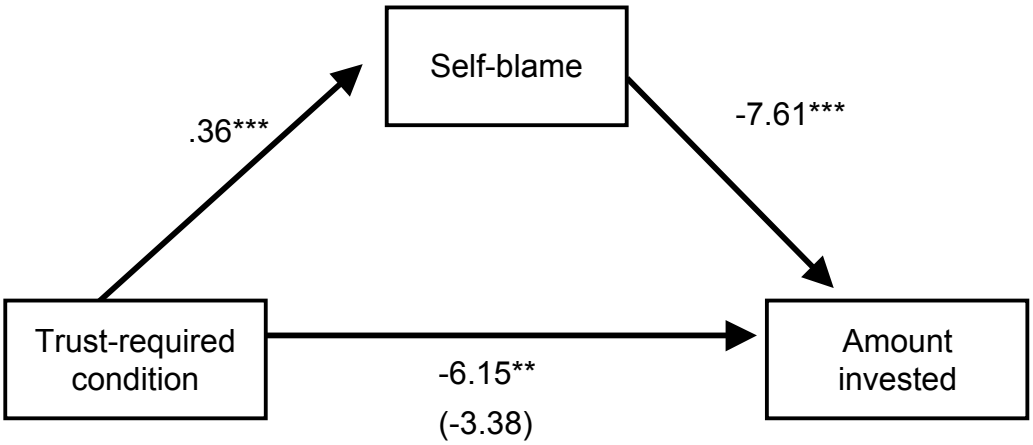
[iii] Before beginning the main part of study, participants completed four items from the Interpersonal Trust Scale (Rotter, 1967). As these items showed poor internal consistency ($\alpha = .56$) and did not correlate with our dependent measures, they are not discussed further.

[iv] With equal variance not assumed, the Round 1 and 2 differences were reliable at $p = .06$ and $p < .005$, respectively.

[v] An additional 41 participants predicted how they would feel after incurring the game's unfavorable outcome. Unbeknownst to them, the number of points that each participant imagined losing was equal to the number of points that one of the 41 participants in Study 3 had lost. As in Studies 1 and 2, predicted self-blame was higher in the *trust-required* condition ($M = 5.48$, $SD = 2.31$) than in the *trust-not-required* condition ($M = 3.66$, $SD = 2.34$), $F(38) = 14.56$, $p < .001$, controlling for the number of points lost. Comparing these predictions to the degree of self-blame actually experienced by Study 3 participants reveals striking accuracy.

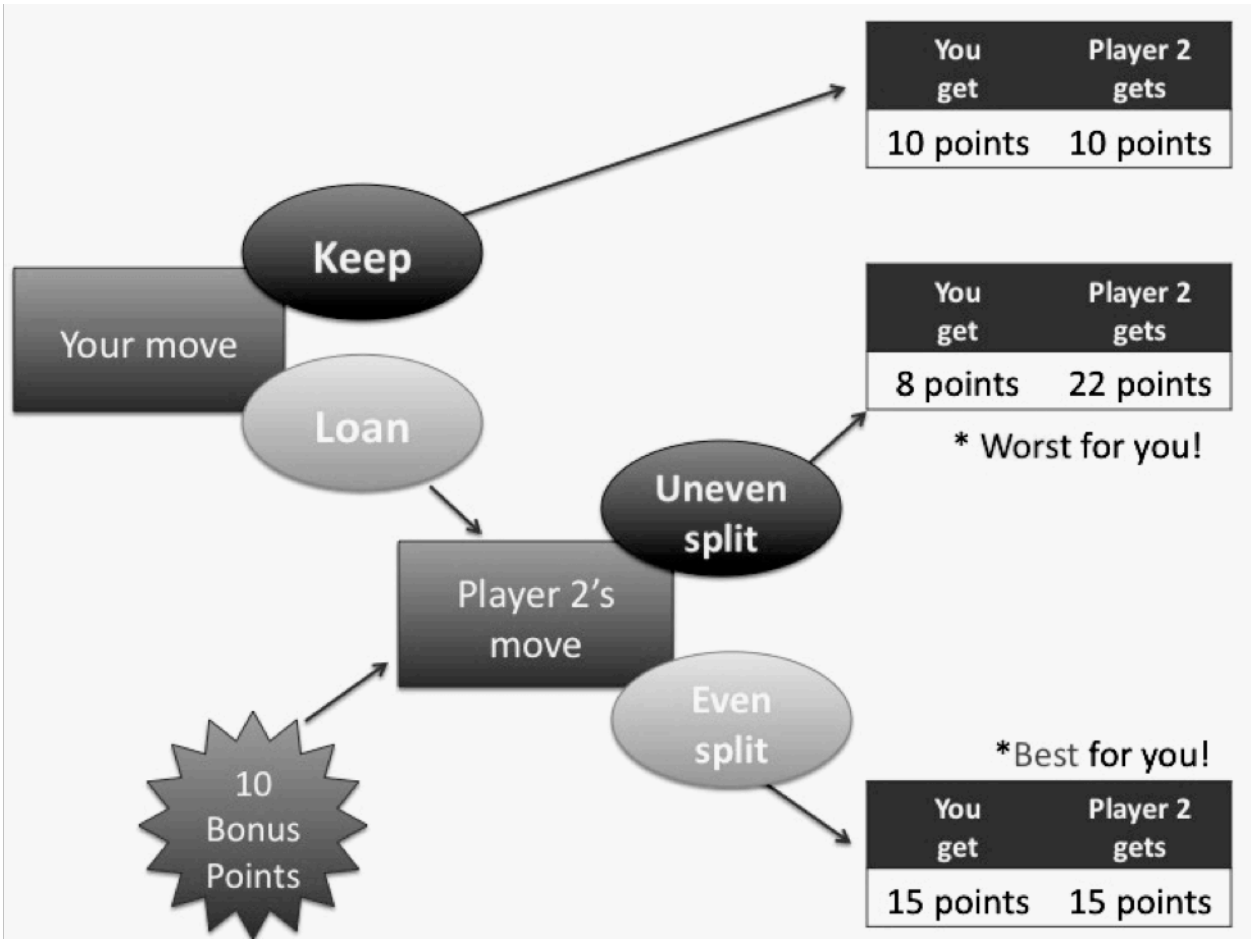
Figures

*Figure 1.* Mediation analysis for Study 1



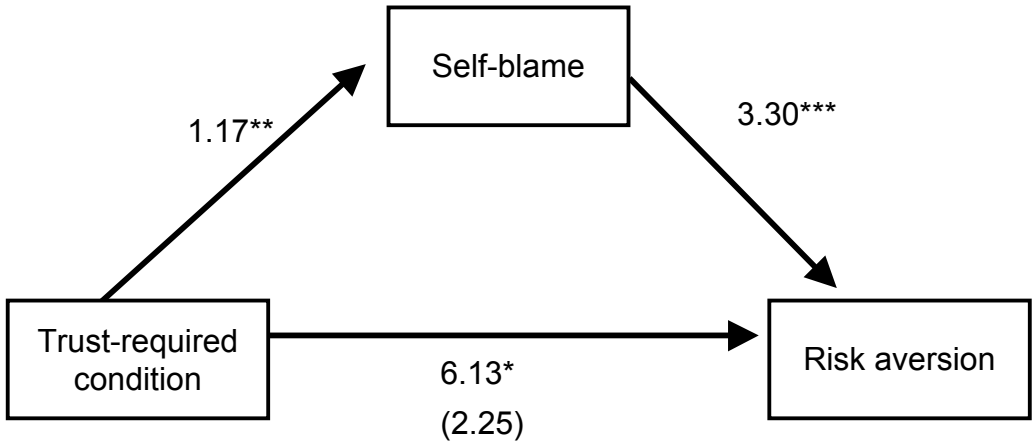

*Notes:* Values are unstandardized regression coefficients. Coding for condition: trust required = 2, other conditions = -1. Parenthetical value is the direct effect of condition on investment controlling for self-blame. ** $p < .01$, *** $p < .001$.

*Figure 2.* Schematic representation of trust game shown to participants in Study 2.




*Note*: Both players begin the game with a 10-point endowment.

*Figure 3.* Mediation analysis for Study 2




*Notes:* Values are unstandardized regression coefficients. Coding for condition: trust required = 1, trust not required = 0. Parenthetical value is the direct effect of condition on risk aversion controlling for self-blame. * $p = .05$, ** $p < .01$, *** $p < .001$.

*Figure 4.* Study 3: Mean number of points invested on each round, by condition (± *SE*)