



Contents lists available at ScienceDirect

Journal of Economic Behavior & Organization

journal homepage: www.elsevier.com/locate/jebo



Betrayal aversion versus principled trustfulness—How to explain risk avoidance and risky choices in trust games

Detlef Fetchenhauer^{a,b}, David Dunning^{c,*}

^a University of Cologne, Department of Economic and Social Psychology, Venloer Str. 151-153 (8. OG), 50672 Cologne, Germany

^b University of Groningen, ICS, Grote Rozenstraat 31, 9712 TG Groningen, The Netherlands

^c Department of Psychology, Cornell University, Ithaca, NY 14853, USA

ARTICLE INFO

Article history:

Received 3 September 2007
Received in revised form 25 June 2011
Accepted 18 July 2011
Available online xxx

JEL classification:

C72
C91
D81

Keywords:

Trust
Risk taking
Economic games

ABSTRACT

Are decisions in a trust game more or less sensitive to changes in risk than decisions in a purely financial, non-social decision-making task? Participants in a binary trust game (they could either keep \$5 for sure or give it to a trustee with the chance of getting \$10 back) were informed that their chance of interacting with a trustworthy person was either 46 percent or 80 percent and then were asked to decide whether to trust that other person. In addition, participants made a decision in a lottery (i.e., whether to gamble \$5 to win \$10) with the same probabilities. In the 46 percent condition, participants were significantly more willing to choose the risky option in the trust game than in the lottery. Overall, the difference in probability of receiving money back had a significantly higher impact on the lottery decision than on the decision to trust. Possible interpretations of the present study and its relation to previous findings are discussed.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

As several scholars have noted, trust is important underpinning for societies in general to succeed and to thrive (Fukuyama, 1995; Sullivan and Transue, 1999), particularly when it comes to their economic health (Kramer, 1998; Putnam, 1993; Knack and Keefer, 1997). But, what is trust? Based on a review of definitions in different social sciences, Rousseau et al. (1998) suggested that trust is best defined as “a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or the behavior of another” (p. 395). Although this definition seems to be widely shared by most scholars of trust it does not answer the question of what constitutes and determines that “psychological state” called trust.

One option is to define trust simply as choosing the risky option in a decision under risk and uncertainty. Such a reductionist definition has been favored by a number of different authors from both, economics (Williamson, 1993) as well as sociology (Coleman, 1990) and philosophy (Hardin, 2002), yet recent research has suggested that the concept of trust is different from risk taking. For example, Eckel and Grossman (1996) showed that different measures of risk attitudes do not predict behavior in trust games. More recently, Houser et al. (2006) replicated this finding by demonstrating that risk attitudes as measured by the method of Holt and Laury (2002) were not related to behavior of trustors in an investment game (Berg et al., 1995). In addition, Houser et al. showed that risk attitudes were significantly related to their participants' behavior in an “investment game” in which people gambled on the responses of a computer (i.e., in which the computer

* Corresponding author.

E-mail addresses: detlef.fetchenhauer@uni-koeln.de (D. Fetchenhauer), dad@cornell.edu, dad6@cornell.edu (D. Dunning).

arbitrarily decided how much it would send back to the “trustor”). General risk attitudes cannot explain trustors’ decisions in trust games or investment games, respectively.

Thus, if people’s decision to trust is not about their general level of risk aversion the question arises: What makes the difference between taking a purely financial and non-social decision and a decision to trust or not to trust another person? In the present paper, we focused on one dimension of an answer to that question. Are people more or less risk taking in trust situations than they are in non-social financial decisions?

There are different possible answers to this question: (1) on average people might be *less* risk taking in trust situations than in non-social situations, or (2) on average people might be *more* risk taking in trust situations than in non-social situations. In principle, it might also be that (3) people’s *average* level of risk taking is similar in both kinds of situations (although their behavior in both situations might not be related to each other). Some data do exist concerning this question. According to the work by Bohnet and Zeckhauser (2004; see also Bohnet et al., 2008), people are less willing to take risks in a trust game than in a gamble. In their study, participants played a trust game in which the trustor could either keep \$10 for sure or hand that \$10 to another person (i.e., the trustee). In that case, the trustee got an amount of \$30. The trustees could either split that money evenly (i.e., keeping \$15 for themselves and giving \$15 to the trustor) or they could keep \$22 for themselves and give \$8 to the trustor. The dependent variable of Bohnet and Zeckhauser was the minimum accepted probability (MAP) of being paired with a trustworthy trustee that participants demanded in order to take the risky option of giving away their \$15. Another group of participants indicated their MAP to participate in a lottery with the same pay-off structure (i.e., to keep \$10 for sure or take the risky option resulting in an outcome of either \$8 or \$15).

Bohnet and Zeckhauser (2004) found that participants demanded substantially higher MAPs in a trust game than in a lottery. Whereas participants on average were willing to participate in the lottery when their chance of winning was at least 37 percent, participants on average demanded a chance of 54 percent of being paired with a trustworthy interaction partner to make them choose the risky option in the trust game (although a recent study by Hong and Bohnet, 2007, could only partly replicate this result). From this finding, Bohnet and Zeckhauser concluded that most people experience “betrayal aversion”. If they lost on a gamble, they just had bad luck—but if they were betrayed in a trust game, they entered a world in which they were exploited by another person, which might feel much worse.

However, a number of studies by Fetschenhauer and Dunning (2006, 2009) and Dunning and Fetschenhauer (2010a,b) discuss data that contradict Bohnet and Zeckhauser (2004), in that participants are actually more risk taking in trust games than they are in non-social gambling situations. For example, Fetschenhauer and Dunning (2009) let participants play trust games in which trustors could either keep an amount of x or give it to a trustee; in Study 1, the amount of x equaled \$5, in Study 2, the amount of x equaled €7.50. If trustors decided to keep the money, trustees would receive nothing. If trustors handed over their money, that amount was quadrupled (i.e., the trustee got an amount of $4x$). The trustees then had two alternatives. They could either keep the money or they could split the money evenly. Thus, the potential outcomes for the trustors were x (when choosing the safe option), $2x$ (when paired with a trustworthy interaction partner) or 0 (when meeting an untrustworthy interaction partner). In addition to making their decision as a trustor, participants were also asked to estimate the number of fellow participants who would split the money evenly in the position of the trustee. Finally, participants also decided whether to participate in a lottery. The pay-off structure was identical to that of the trust game: participants could either keep x for sure or go for the risky option. If they participated in the lottery their outcome was either $2x$ (in case they won) or $0x$ (if they lost).

On average, most participants in Fetschenhauer and Dunning (2009) turned out to be risk averse in the lottery (demanding a chance of about 62 percent to make them choose the risky option). Nonetheless, a majority of about 70 percent of all participants chose the risky option in the trust game. This is the more remarkable as they were rather skeptical about their fellow participants’ trustworthiness: Asked directly to estimate the percentage of peers in the study who would reciprocate trust, participants on average they thought that less than half all trustees would do so. As such, what would seem to be minor differences in experimental paradigms had a significant effect on the results observed. Bohnet and Zeckhauser (2004), focused on providing an MAP to participate in either a trust game or a gamble. In Fetschenhauer and Dunning (2009), participants indicated their MAP for a lottery and had to estimate the likelihood of being matched with a trustworthy trustee.

How can one reconcile these contradictory findings? We argue that both paradigms have their strengths and weaknesses, and that both studies reveal important insights into the study of trust. The essential difference between both sets of studies lies in the fact that in the Bohnet and Zeckhauser (2004) paradigm, participants have to decide the level of risk they would accept in order to *participate* in a trust game, whereas in the Fetschenhauer and Dunning (2009) studies—like in most other studies on trust games and investment games—participants are first inserted into a trust game and then have to decide whether to trust or not. People are averse to being betrayed, and as a consequence they may often choose not to participate in the trust game if they may still opt out—or spell out the conditions (i.e., the level of risk) that would cause them to opt out. However, we argue that once they are placed in a trust game involving a direct yes-or-no decision to trust a specific person, people are reluctant to openly signal their distrust or lack of generosity to another person.

It is important to note, however, that neither research team asked participants to make parallel choices in the two games about whether to gamble/trust after being informed about the probability of receiving a positive outcome. To fill this gap, and to test the validity of our proposed interpretation of the seemingly contradictory findings that have been observed, we conducted a study in which participants were asked to either take a safe option or a risky option in a gamble (i.e., a non-social situation) and in a trust game with a given probability of losing versus winning when taking the risky option. Both situations

had the identical pay-off structure for the trustor and the same expected outcomes for both alternatives (keeping an amount of money versus taking the risky option).

Additionally, we varied the chance of winning versus losing when choosing the risky option. In one condition, participants had an 80 percent chance of doubling their endowment; in another condition that chance was only 46 percent. We hypothesized that this difference would strongly influence participants' decisions in the lottery. However, if our premise is correct and trust decisions are not about risk taking, but about not signaling one's distrust, the chance of winning should only slightly influence participants' decisions in the trust game.

We should note that varying the chance of receiving money back also allowed for a more stringent test of the notion that trust behavior is not about risk-taking. So far in the literature, researchers have argued, with data, that attitudes toward risk fail to predict who will trust (Eckel and Wilson, 2004; Houser et al., 2006). Schechter (2007) found that behavior in a trust game was related to actual risk-taking in a non-social setting, but the correlation between both kinds of behavior was rather weak (roughly a correlation of .25 between taking a non-social bet and giving money to another person in the trust game).¹

However, the data collected so far fail to constitute the most direct test of whether participants give weight to risk when deciding to trust another person. In previous studies, researchers have measured a person's general attitude toward risk, but have not examined the amount of risk people perceive in the trust game. Directly measuring, or influencing, people's specific perceptions of risk in the trust game is important, for it is likely that this perception will influence their decisions to trust far more than more indirect, and diffuse, attitudes toward risk-taking. For example, a person might have a marked disinclination to take risks—but if he or she thinks the chance of receiving money back in the trust game is 90 percent, that might be a gamble worth taking. In a similar vein, a person usually prone to gamble may decide not to trust another individual if he or she thinks the chances of receiving money back is near zero. It is the perception of risk in the trust game that matters—not an attitude toward risk in general.

As a consequence, to examine whether people construe the trust game as a game primarily of risk, one must influence the amount of risk people see in the trust game, rather than measure their more general attitudes toward taking on risks. By informing participants that the chance of receiving money back was either 46 percent or 80 percent, we could see the extent to which people responded to the odds of reward in the trust game relative to a game of chance. We predicted that, much like the Houser et al. (2006) study, participants would give more weight to these odds when choosing whether or not to gamble in the lottery.

2. Experimental design

Participants were 152 students at Cornell University students drawn from a wide variety of classes (e.g., psychology, human development, and economics). They all received extra credit toward course grades for participating. To recruit participants, experimenters either made announcements at the end of regular class sessions, asking anyone interested in participating to stay to complete the experiment. Other participants signed-up for experimental sessions via a department website, coming to group sessions taking place in the social psychology laboratory at Cornell. Regardless of recruitment method, none were informed of the chance of monetary reward until after the experimental session began.

The study was conducted in two phases. The first phase involved obtaining decisions from individuals playing the role of the trustee in the trust game described above. A total of 77 participants were given a questionnaire describing the trust game. That trust game involved a trustor who received \$5 from the experimenter and who then had to choose whether to transfer that \$5 to a trustee. If they declined to transfer the money, the trustee received nothing. If they transferred the money, it was inflated to \$20, and the trustee had to make a decision whether to keep the entire \$20 or transfer \$10 back to the trustor.

Participants were next told that they would be taking part in an actual trust game in which they played the role of the trustee. Their response would be matched at random with an anonymous trustor who would be asked to make his or her decision in a few weeks time, and that the amount of money they received depended on whether the trustor decided to give them their money and whether they decided to reciprocate that trust. Trustees were *not* told that their reactions would be categorized in two groups with a high or low frequency of trustworthiness. Participants then indicated whether they would keep the entire \$20 or give \$10 back to the trustor.

Importantly, they made these decisions anonymously. They placed their decision in an envelope and wrote down a 4-digit code on the front of the envelope. On another sheet of paper, they wrote down their code as well as their name and address, which they then all placed in a manila envelope. Participants were instructed that, to ensure the anonymity of their decisions, one research assistant exposed only to the code numbers would insert the correct amount of money coming to them into the envelope and then seal it. Another research assistant, not knowing the amount of money inside or what decision had been made, would match codes to names and addresses and mail them their money.

Of the 77 participants playing the role of the trustee, 48 (62 percent) indicated that they would give \$10 back to Person A. Thus, three groups were created. The first group of 40 participants was created such that 80 percent of them chose to reciprocate trust; approximately the rate of reciprocation seen in previous studies in our labs (Fetschenhauer and Dunning,

¹ Schechter (2007) did not directly report the correlation between risk-taking in the lottery and in the trust game. However, information in Table 3 of her article allows the correlation to be computed.

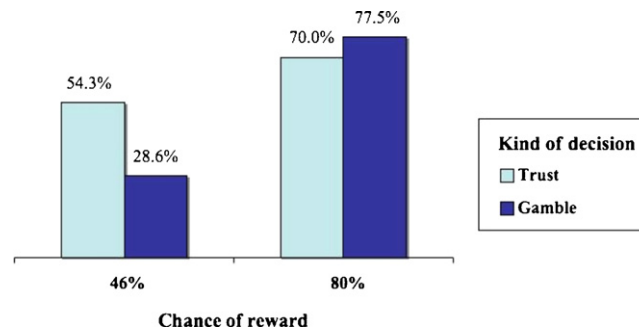


Fig. 1. Rates of participants choosing the risky option dependent on condition and kind of decision.

2009, 2010). The second group of 35 trustees was created such that 46 percent ($n = 16$) reciprocated, which is close to the rate of reciprocation that participants typically anticipated from their peers in our previous studies (Fetschenhauer and Dunning, 2009). A third group of two participants both indicated that they would reciprocate trust. Although they were each paired with a trustor, their data are not discussed further.

The second phase of the study involved collecting data on how participants would respond in the position of the trustor to a trust game as well as a lottery that presented equal chances of reward. Participants in this phase of the study took part in groups of about 10 participants per session. They were given \$5 cash and were asked to make two decisions about what to do with it. They were further told that one of their decisions would be played out for real.

One decision involved whether or not to trust an anonymous trustee. Participants were told they had been paired at random with one of the participants in the first phase of the study. In fact, their questionnaire displayed the 4-digit code number of the trustee they had been matched with. One group, in the high chance condition ($n = 40$), was told that there was an 80 percent chance that their trust would be reciprocated (i.e., that 4 out of every 5 trustees had returned \$10). The other group, the low chance condition ($n = 35$) was told that the chance their trust would be reciprocated was 46 percent. Note that we did not tell the trustors that their personal chance of being paired with a trustworthy other person was derived from the behavior of all trustees. We only informed them about their *personal* chance of meeting a trustworthy trustee and this information was totally valid and honest. The remaining 2 trustees (see above) were paired up with the remaining trustors, which were told they had a 100 percent chance that their trust would be honored. Self-evidently both trustors decided to hand over their money to the trustee in that condition.

The second decision participants made centered on a lottery. They were asked to suppose that an urn contained 100 white and red ping-pong balls. Participants were told they could bet their \$5 that the experimenter would draw a white ball out of the urn. If the ball were white, the participant would win \$10. For participants in the high chance condition, they were told that 80 of the balls were white. In the low chance condition, the probability was described as 46 percent. Participants indicated whether they would keep their \$5 or bet it on the urn.

Thus, participants each made two choices—whether to give their \$5 to another person in a trust game and whether to bet it on the urn. One group made both choices at an 80 percent chance of reward; the other group made both choices at a 46 percent chance of reward. All participants were told that they would make one of these two decisions for real. The order of decision, urn versus the trust game, was counterbalanced across participants.

Participants made their decisions anonymously. After these decisions were made participants were informed that for all of them the choice they made for real was their decision in the trust game (this was the case in all experimental sessions). Their questionnaires, without any identifying mark other than the 4-digit code, were collected, taken out of the experimental room, and then matched with the decision made by their assigned trustee. A research assistant placed the appropriate amount of money into an envelope with the 4-digit code, and these envelopes were then distributed back to the relevant participants. Participants were free, whether to open their envelopes on the spot or to open them in private at a later point of time.

After all trustors had made their decisions, the trustees were mailed their money based on the decisions they and their interaction partner had reached.

3. Results

The main results of the present study are also summarized in Fig. 1. As seen in the figure, participants were quite sensitive to their chance of reward with regard to the lottery. In the high chance condition, 77.5 percent of all participants gambled their \$5, whereas 22.5 percent decided to keep their money. In the low chance condition, only 28.6 percent decided to gamble, whereas 71.4 percent decided to keep their \$5. This difference was highly significant: $\chi^2(1, n = 75) = 18.03, p < .01$. Contrary to the lottery, participants' willingness to go for the risky option in the trust game was not significantly influenced by the chance of reward. In the high chance condition, 70.0 percent of all participants handed their money to the trustee, but so did 54.3 percent of all participants in the low chance condition, $\chi^2(1, n = 75) = 1.97, p = .16$.

Were participants more or less risk averse in the trust game as compared to a non-social financial decision? In the high chance condition, no difference between the lottery and trust game emerged as 70.0 percent of all participants chose the risky option in the trust game and 77.5 percent did so in the lottery (a sign test revealed this difference to be not significant, $p = .61$). However, in the low chance condition, a significant difference was found. Although only 28.6 percent of the participants gambled on the lottery, about twice as many (i.e., 54.3 percent) did so in the trust game ($p < .03$, by binomial sign test).

More comprehensive analyses affirmed that the influence we observed for chance of reward was significantly greater, in the statistical sense, in the lottery scenario than in the trust game. To compare that degree of influence for the lottery and trust game decisions, we sorted each participant into one of three groups. The *trust only* group consisted of those who decided to trust another person yet refused to gamble in the trust game (coded as -1). Another group, the *same decision* group, consisted of those who made equivalent decisions in both scenarios (i.e., either took the risk or refused it; coded as 0). The final *gamble only* group consisted of those who took the gamble in the lottery but refused to trust another person (coded as $+1$).

If chance of reward had a greater impact on behavior in the lottery than in the trust game, people should be significantly more likely to shift away from the trust only group and shift toward the gamble only group as we move from the low chance condition to the high chance one. To assess this, we conducted an ordinal logistic regression analysis examining whether the probability of winning had a significant effect on what group participants fell into. We also included task order (i.e., whether participants made the trust or lottery gamble decision first) as an independent variable to see if it had any influence. The analysis revealed that chance condition had a significant impact, Wald $\chi^2(1, N = 75) = 5.70, p < .02$, but task order did not, Wald $\chi^2(1, N = 75) = 1.72, ns$. These results were driven by the fact that many more participants in the low chance condition fell into the trust only group (31.4 percent) than in the gamble only group (5.7 percent). These percentages were slightly reversed in the high chance condition (15.0 percent vs. 22.5 percent for trust and gamble only groups, respectively).

Finally, at the individual level, did we find any evidence for a relationship between risk-taking in the lottery and in the trust game? We did, indeed, find such a relationship, although it was only moderate in strength and marginally significant. To test whether participants' behavior was consistent across lottery and trust decisions, we conducted a binary logistic regression analysis in which we regressed participants' behavior in the trust game on task order, chance of winning, and most importantly their decision in the lottery scenario. We also inserted four interaction terms into the analysis (referring to the three potential two-way interactions and the one potential three-way interaction of the independent variables, respectively). This analysis revealed only one significant marginal predictor of participants' behavior in the trust game, namely their behavior in the lottery ($\beta = -1.62, p = .07$). Only 50.0 percent of participants who refused to gamble on the lottery decided to take a risk in the trust game. Of participants who gambled in the lottery, 73.2 percent took a risk in the trust scenario. If we look at the zero-order correlation between decisions to gamble and to trust, we find an association that just misses conventional statistical significance, $\phi = .24, p = .055$, but which is close to the size of the relationship found by Schechter (2007).

4. Discussion

The main question of this study was whether people are more or less risk-averse in a trust situation than they are in a purely financial decision situation. To answer this question we confronted our participants with both a trust game and a lottery that presented the same odds and same pay-off structure for the participants. Thus, if subjects are maximizing the expected utility of their own monetary payoffs, one would expect that the behavior of participants in both situations should be identical, both on an aggregate and on an individual level. However, our results revealed large differences between the decisions in both situations.

Participants were sensitive to the probability of reward when deciding to gamble on a non-social, financial outcome. This came as no surprise and is in line with a large body of research (Baron, 2001). In the trust game, however, the likelihood of reward did not significantly influence participants' decision whether to hand their money over to the trustee. In the high chance condition, participants were about as risk taking in the trust game as they were in a lottery: most participants decided to go for the risky option if their chance to double that money was 80 percent, irrespective of whether they took that risk in a social situation (i.e., within a trust game) or within a non-social situation (i.e., within the lottery). In the 46 percent condition, however, participants were much more willing to choose the risky option than they were in the lottery.

We believe this finding provides the most compelling evidence to date that the decision to trust is not just a decision about risk. Even when the direct odds of winning were changed, these changes did not have a significant impact on whether participants gave their money to the trustee—whereas those odds had a stronger and significant impact on decisions to gamble. As such, the odds of winning did not have an equivalent impact on a bet on nature versus a bet on the actions of another. The odds influenced decisions to bet significantly more on the former type of bet than the latter. Thus, we confirm earlier results that trust decisions cannot be regarded as merely decisions under risk and uncertainty (Eckel and Wilson, 2004; Houser et al., 2006)—but do so by putting the notion of trust as risk-taking under the most direct “acid test” to date.

However, the most important finding concerns the differences in both decisions in the low chance condition, where the chance of receiving money back was only 46 percent. We chose that likelihood for two reasons. First, this value mirrors the

estimations of our participants in previous studies about the likelihood that a trustee will split the money evenly (see [Dunning and Fetschenhauer, 2010a,b](#); [Fetschenhauer and Dunning, 2006, 2009](#)). Second, this value leads to a negative expected value for choosing the risky over the safe option (\$4.60 versus \$5.00). Thus, not only risk averse, but also risk neutral participants should have chosen the safe option in that condition. Given that most humans are risk averse with regard to potential gains ([Kahneman and Tversky, 1979](#)) one would expect that most participants would go for the safe option in that condition. This is exactly, what we found in the lottery. Only a minority preferred the risky bet. However, in the trust game a majority of participants gave their money to the trustee.

As noted earlier, [Bohnet and Zeckhauser \(2004\)](#); see also [Bohnet et al., 2008](#)) obtained quite different findings—but perhaps both studies can be seen as two sides of the same coin. Both studies, considered together, might reveal an important insight into the mechanisms that guide trustful or distrustful behavior. To get this insight, one has to realize a seemingly subtle, but important, difference between both studies. In our studies, we put participants into a trust game and then let them make a direct decision whether to trust a specific other person. People had to make a single and direct yes-or-no decision—one that is much like trust decision are made in real life. In [Bohnet and Zeckhauser](#), participants were asked a different question that was removed from a direct trust decision in two ways. First, they were asked a conditional: What specific likelihood of having their trust rewarded did they need to see before they would trust another person. This is not a direct yes-or-no decision, rather it is one step removed, asking participants to consider many different counterfactual conditions (e.g., would you trust with a 10 percent chance of reward? 20 percent? 30 percent?) Second, in stating the chance that would induce them to trust, participants considered not a specific other person when considering these chances but rather the entire aggregate of people they could be assigned to. Both these conditions likely have the effect of making the person feel they are not yet enmeshed in a trust game but had to *consider* the type of trust game they would enmesh themselves in. That is, participants were asked what probability of receiving a reward would induce them to trust. As such, participants may have construed the situation as one in which they named the conditions under which they were willing to volunteer to trust—a situation that is quite different from the single direct question we asked our participants after all conditions had been specified.

We would argue that [Bohnet and Zeckhauser \(2004\)](#) are right in claiming that it is worse to be betrayed in a trust game than it is to simply lose money in a lottery. Therefore, if people can avoid being involved in a trust game, or can name the conditions under which they would participate, they may thankfully choose avoiding to trust the other person. However, if one is *already* inserted into a trust game involving a direct decision to trust another person, the psychological dynamic changes. For example, in another study we had two different conditions ([Dunning and Fetschenhauer, 2010a,b](#)). In both conditions we played a trust game with the same pay-off structure as in the present study, except that we did not inform participants about the likelihood that trust would be reciprocated. In one condition, we told participants that they had already been paired with a specific trustee. In this condition over 60 percent of trustors decided to hand over their money to the trustee (although they thought that fewer than 40 percent of all trustees would reciprocate that trust). In another condition, we told participants that no trustee was assigned to them yet and that this would only be done if they decided to hand over their money. In this condition, the trust rate dropped significantly to roughly 35 percent, presumably because participants felt they could avoid entry into the trust game.

This specific finding echoes similar ones found with different economic games or transactions. People are more likely to act in a pro-social way once inserted into a situation involving a specific other person; they are less likely to act in a pro-social way if not yet inserted into such a situation. For example, people are less likely to cooperate in a prisoner's dilemma game if their partner is not yet assigned ([Yamagishi et al., 2007](#)). If people are assigned a partner and compelled to make a decision in a dictator game, they typically transfer some of their endowment to that person. However, if people are given, instead, a chance to volunteer to play the game that they can decline, they are much more likely to refuse the opportunity to play the game. As a consequence, far less money is transferred to another person ([Lazear et al., 2010](#)). In a similar vein, people are much more likely to donate money to a victim if that victim has been specified rather than when a specific victim has yet to be assigned to them ([Small and Loewenstein, 2003](#); [Small et al., 2007](#)).

To summarize, the assertion is that participants are more likely to act in a pro-social way when they confront a direct decision involving a specific other individual than when they confront a decision that involves volunteering for the decision. By stating the probability at which they would trust others, the participants of [Bohnet and Zeckhauser's \(2004\)](#) study may have been more in the mindset of volunteering than considering a direct decision, yes or no, to trust.

Thus, once confronted with a direct decision involving a specific other person about whether to trust, people give their money to a trustee—even if most of them are sure that most trustees will keep all the money for themselves (see [Fetschenhauer and Dunning, 2009](#)). Thus, humans seem to be averse toward being betrayed, but they seem even more averse to openly signaling their distrust to another person once placed in a relationship with that person. Note that the participants in the present study actually paid money to avoid such a signal in that they were choosing a risky option under circumstances they would never have done in a purely financial decision. Remarkably, this unwillingness to signal one's distrustfulness was found even in one-shot interactions under conditions of total anonymity.

However, there is another possible explanation why our results diverge from those of [Bohnet and Zeckhauser \(2004\)](#). Our instructions stated that trustee would be given no money in the transaction if the trustor did not trust, whereas participants in [Bohnet and Zeckhauser](#) were told that the trustee would be given money in an amount equal to the trustor's own. Perhaps people were more willing to trust in our paradigm because of inequity aversion ([Fehr and Schmidt, 1999](#); [Walster et al., 1978](#)), or just the plain preference to avoid being selfish and depriving the other of monetary reward. [Bohnet and Zeckhauser's](#) paradigm may have avoided these concerns because trustees were left with money regardless of trustor's decision.

We think this is a potentially plausible explanation worthy of further thought and study. However, we should note that extant empirical data already runs against it. Although not designed to test this issue, Bolton and Ockenfels (2010) provide data that speak to it. In their study, they asked participants whether they would rather keep €7 or €9 for sure, depending on the condition, or gamble the money on a 50–50 bet. Their decisions always influenced the possible outcomes of another person. In one condition, for example, winning the bet meant that both participant and the other person won €16; losing it meant they both won nothing. Of key import was what the other person received if the participant kept their money for sure. In one set of conditions, it meant the other person also received €7 or €9 for sure (conditions A/B2, according to Bolton and Ockenfels terminology)—the same amount as the participant. In another set of condition, the other person would be left with nothing if the participant played it safe (conditions A/B4), meaning that participants in this circumstance might be construed as selfish if they chose to keep their money. Did this possible selfishness prompt participants to gamble more? The answer is no: When playing it safe meant the other person received nothing, participants were no more likely to gamble (46 percent) than when playing it safe left the other person with an equal amount of money (48 percent), $\chi^2(1, N = 104) = .04, p = .84$.

Providing a similar comparison were a set of other conditions from Bolton and Ockenfels (2010). In these conditions, taking the bet would mean that participants would have €16 if they won (the other person receiving nothing) and nothing if they lost (but the other participant would gain the €16). Again, two conditions guaranteed that participants and the other person would have the same amount (€7 or €9) if the participant chose not to gamble (A/B5). Two other conditions meant that participants would have money if they played it safe, but that the other person would receive nothing. Did this last circumstance prompt participants to gamble more? The answer appears to be no. The difference in gambling rates between the first and second set of conditions was not significant (38 percent and 52 percent, respectively), $\chi^2(1, N = 104) = 1.90, p = .17$.

In addition, we would also argue that the results of the present study also reveal some basic differences between the so-called investment game by Berg et al. (1995) and a binary trust game as it was used by us or by Bohnet and Zeckhauser (2004). At first glance, the differences seem to be rather minor and technical. Whereas in a binary trust game participants have to make a yes or no decision, in the investment game participants can choose whatever amount of money they want to give to the trustee. We believe that this difference is essential. In an investment game, most participants decide to transfer some money over to the trustee, but most of trustors do not hand over their whole endowment (Berg et al., 1995; Fahr and Irlenbusch, 2000; Snijders and Keren, 2001). Thus, most participants avoid the painful decision that participants must make in a dichotomous trust game: whether to trust completely or not to trust at all. Yet, many choices in real life are such “all or nothing” kind of decisions. For example: Do I buy that used car or not? Do I accept that job offer or not? Do I bid for that Chinese vase (and trust it to be authentic) on e-bay or not?

Data from this study and elsewhere also counter one additional alternative account for our findings. Specifically, one could argue that participants gave away their money in the trust game to “enlarge the pie” and give benefits to another person—even if they ran the risk of being penalized for this generosity. That is, people were sensitive to the fact that deciding to trust injected \$20 of wealth into the greater populace instead of just \$5, and they were not so concerned about who came to possess that greater wealth.

Our data, however, speak against this alternative. If participants found it attractive to “create more wealth,” then they should have found the trust game more attractive than the lottery gamble in both the 46 percent and 80 percent winning conditions, since the trust game created more wealth in both conditions. However, in the 80 percent condition, participants did not favor taking a risk on the trust game rather than the lottery.

In addition, elsewhere we have collected data that speaks against a “greater wealth” argument. In that study, participants faced decisions in three different conditions. One was the trust game, in which people could give 5€ to another person with the prospect of getting 10€ back. Of participants in this condition, 57 percent decided to trust. In another condition, participants were asked whether they wanted to gamble 5€ on a coin flip, with the possibility of winning 10€. Only 25 percent took the gamble. The third condition was the key one. Here, participants were again asked whether they wanted to gamble 5€ on a coin flip, but the gamble included the exact same payoffs for another person as in the trust game. If the participant won the gamble, not only did they receive 10€ but another person in the experiment, randomly chosen, also received 10€. If the participant lost, they received nothing, but some random person in the experiment would receive 20€. Did adding potential benefits to another person increase risk-taking? The answer was no; only 28 percent took the gamble in this final condition, even though participants knew the chance of winning was 50 percent, an probability quite close to the estimate participants gave that they would receive money back in the trust game 53 percent (Mensching et al., 2010).

That all said, an obviously profitable line of future research would be to answer why our results and those of Bohnet and Zeckhauser (2004) directly oppose. What are the conditions that lead people to accept risk more readily with other people rather than with nature? What are the circumstances that leave people to be more reluctant to bet on the behavior of others rather than on the behavior of nature? Regardless of how future research may answer this question, the present research does echo the common psychological finding that the shape of people’s behavior often rests on situational details or circumstances that are, at first, seemingly small or not obvious (Ross and Nisbett, 1991). What this research suggests is that changes in the situational structure of how people may express their trust may turn out to be a major determinant of whether people choose to do so.

Acknowledgement

The authors thank Thomas Schlösser for his commentary on earlier versions of this manuscript.

Please cite this article in press as: Fetschenhauer, D., Dunning, D., Betrayal aversion versus principled trustfulness—How to explain risk avoidance and risky choices in trust games. *J. Econ. Behav. Organ.* (2011), doi:10.1016/j.jebo.2011.07.017

References

- Baron, J., 2001. *Thinking and Deciding*, 3rd ed. Cambridge University Press, Cambridge, UK.
- Berg, J., Dickhaut, J., McCabe, K., 1995. Trust, reciprocity, and social history. *Games and Economic Behavior* 10, 122–142.
- Bohnet, I., Zeckhauser, R., 2004. Trust, risk and betrayal. *Journal of Economic Behavior and Organization* 55, 467–484.
- Bohnet, I., Greig, F., Herrmann, B., Zeckhauser, R., 2008. Betrayal aversion: evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States. *American Economic Review* 98, 294–310.
- Bolton, G.E., Ockenfels, A., 2010. Betrayal aversion: evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States: Comment. *American Economic Review* 100, 628–633.
- Coleman, J., 1990. *Foundations of Social Theory*. Harvard University Press, Cambridge, MA.
- Dunning, D., Fetschenhauer, D., 2010a. Trust as an expressive rather than an instrumental act. In: Thye, S., Lawler, E. (Eds.), *Advances in Group Processes*, vol. 27. Emerald, New York, pp. 97–127.
- Dunning, D., Fetschenhauer, D., 2010b. Understanding the psychology of trust. In: Dunning, D. (Ed.), *Social Motivation*. Psychology Press, New York, pp. 147–170.
- Eckel, C.C., Grossman, P.J., 1996. Altruism in anonymous dictator games. *Games and Economic Behavior* 16, 181–191.
- Eckel, C.C., Wilson, R.K., 2004. Is trust a risky decision? *Journal of Economic Behavior and Organization* 55, 447–465.
- Fahr, R., Irlenbusch, B., 2000. Fairness as a constraint on trust in reciprocity: earned property rights in a reciprocal exchange experiment. *Economics Letters* 66, 275–282.
- Fehr, E., Schmidt, K.M., 1999. A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics* 114, 817–868.
- Fetschenhauer, D., Dunning, D., 2006. Perception of prosociality in self and others. In: Fetschenhauer, D., Flache, A., Buunk, B., Lindenberg, S. (Eds.), *Solidarity and Prosocial Behavior: An Integration of Psychological and Sociological Perspectives*. Kluwer Academic/Plenum Publishers, New York, pp. 61–76.
- Fetschenhauer, D., Dunning, D., 2009. Do people trust too much or too little? *Journal of Economic Psychology* 30, 263–276.
- Fetschenhauer, D., Dunning, D., 2010. Why so cynical? Asymmetric feedback underlies misguided skepticism in the trustworthiness of others. *Psychological Science* 21, 189–193.
- Fukuyama, F., 1995. *Trust: The Social Virtues and the Creation of Prosperity*. Free Press, New York.
- Hardin, R., 2002. *Trust and Trustworthiness*. Russell Sage Foundation, New York.
- Holt, C., Laury, S., 2002. Risk aversion and incentive effects in lottery choices. *American Economic Review* 92, 1644–1655.
- Hong, K., Bohnet, I., 2007. Status and distrust: The relevance of inequality and betrayal aversion. *Journal of Economic Psychology* 28, 197–213.
- Houser, D., Schunk, D., Winter, J., 2006. Trust games measure trust. Unpublished manuscript. Ludwig Maximilian University Munich, Munich, Germany.
- Kahneman, D., Tversky, A., 1979. Prospect theory. An analysis of decision under risk. *Econometrica* 47, 263–292.
- Knack, S., Keefer, P., 1997. Does social capital have an economic payoff? A cross-country investigation. *The Quarterly Journal of Economics* 112, 1251–1288.
- Kramer, R.M., 1998. Paranoid cognition in social systems. Thinking and acting in the shadow of doubt. *Personality and Social Psychology Review* 2, 251–275.
- Lazear, E.P., Malmendier, U., Weber, R.A., 2010. Sorting, prices, and social preferences. Unpublished manuscript, Stanford University.
- Mensching, O., Schlösser, T., Dunning, D., Fetschenhauer, D., 2010. Why do people trust? A multi-perspective approach. Unpublished manuscript, University of Cologne.
- Putnam, R.D., 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton University Press, Princeton, NJ.
- Ross, L., Nisbett, R.E., 1991. *The Person and the Situation: Perspectives of Social Psychology*. McGraw-Hill, New York.
- Rousseau, D.M., Sitkin, S.B., Burt, R., Camerer, C., 1998. Not so different after all: a cross-discipline view of trust. *Academy of Management Review* 23, 393–404.
- Schechter, L., 2007. Traditional trust measurement and the risk confound: an experiment in rural Paraguay. *Journal of Economic Behavior & Organization* 62, 272–292.
- Small, D.A., Loewenstein, G., 2003. Helping “A” victim or helping “THE” victim: Altruism and identifiability. *Journal of Risk and Uncertainty* 26, 5–16.
- Small, D.A., Loewenstein, G., Slovic, P., 2007. Sympathy and callousness: the impact of deliberative thought on donations to identifiable and statistical victims. *Organizational Behavior and Human Decision Processes* 102, 143–153.
- Snijders, C., Keren, G., 2001. Do you trust? Whom do you trust? When do you trust? *Advances in Group Processes*, vol. 18. JAI Press, Stamford, CT, pp. 129–160.
- Sullivan, J.L., Transue, J.E., 1999. The psychological underpinnings of democracy: a selective review of research on political tolerance, interpersonal trust, and social capital. *Annual Review of Psychology* 50, 625–650.
- Walster, E., Walster, G.W., Berscheid, E., 1978. *Equity: Theory and Research*. Allyn & Bacon, New York.
- Williamson, O.E., 1993. Calculativeness, trust, and economic organization. *Journal of Law and Economics* 36, 453–486.
- Yamagishi, T., Terai, S., Kiyonari, T., Mifune, N., Kanazawa, S., 2007. The social exchange heuristic: managing errors in social exchange. *Rationality and Society* 19, 259–291.

Further reading (with label)

- Axelrod, R., Hamilton, W.D., 1981. The evolution of cooperation. *Science* 211, 1390–1396.
- Camerer, C.F., 2003. *Behavioral Game Theory: Experiments on Strategic Interaction*. Princeton University Press, Princeton, NJ.
- Fehr, E., Rockenbach, B., 2002. Detrimental effects of sanctions on human altruism. *Nature* 422, 137–140.
- Glaeser, E.L., Laibson, D., Scheinkman, J.A., Soutter, C.L., 2000. Measuring trust. *Quarterly Journal of Economics* 115, 811–846.
- Loewenstein, G.F., Weber, E.U., Hsee, C.K., Welch, N., 2001. Risk as feelings. *Psychological Bulletin* 127, 267–286.
- Pilutla, M.M., Malhotra, D., Murnighan, J.K., 2003. Attributions of trust and the calculus of reciprocity. *Journal of Experimental Social Psychology* 39, 448–455.
- Schlösser, T., 2006. Against all odds? The role of emotions in the decision to trust or not to trust. Unpublished diploma thesis. University of Cologne.
- Van Winden, F., 2007. Affect and fairness in economics. *Social Justice Research* 20, 35–52.