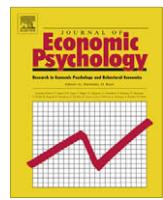




Contents lists available at ScienceDirect

Journal of Economic Psychology

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



Do people trust too much or too little?

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

^a Department of Economic and Social Psychology, University of Cologne, Herbert Lewinstraße 2, 50931 Cologne, Germany

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

^c Department of Psychology, Cornell University, Uris Hall, Ithaca, NY 14853, United States

ARTICLE INFO

Article history:

Received 3 September 2007

Received in revised form 9 April 2008

Accepted 17 April 2008

Available online 11 September 2008

JEL classification:

C72

C91

D81

PsycINFO classification:

3020

Keywords:

Trust

Risk

Behavioral economics

Economic games

Investment game

ABSTRACT

Across two studies, we asked whether people trust too much or too little, relative to what an economic analysis would suggest. In the trust game paradigm, participants decided whether to hand money over to an anonymous individual who could either return more money back or keep all the money. Participants trusted too little, in that they grossly underestimated the proportion of their peers who would return money, prompting them to forgo profitable decisions to trust. However, participants also trusted too much. Given their high levels of cynicism and tolerance for risk, few should have handed money over, yet many still chose to trust. Possible explanations for this paradox of trusting “too little” yet “too much” are discussed.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Trust, but not too much – German proverb.

Trust is an essential feature of social life. It is crucial for the health, harmony, and growth of any organization (Kramer, 1998), as well as maintaining a fully-functioning democracy (Fukuyama, 1995; Sullivan & Transue, 1999). Trust has even been linked to the comparative economic advantage of nations, with those extending and honoring trust growing faster than those that do not (Knack & Keefer, 1997; Putnam, 1993). However, although trust in general is beneficial to both the individual and society at large, trust may be costly in particular cases. Thus, people at times face difficult decisions when mulling over whether to trust another person. For example, should Gary believe that his romantic partner intends to stay faithful? If Sherri sends that check through the mail, will she receive that trinket she bought on eBay?

The present manuscript aims to begin a discussion of how “wisely” people act when faced with decisions of whether to trust another person, asking whether people tend to be too gullible or too skeptical in their dealings with peers. More specifically, it focuses on two questions. The first is whether people possess accurate expectations about the likelihood that their

* Corresponding author. Tel.: +1 607 255 6391; fax: +1 607 255 8433.

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

peers will honor their trust. The second is whether people then act in a way consistent with that expectation and their general tolerance for risk, as assessed by behavior involving non-social economic decisions.

Put differently, we simply ask whether people's decisions to trust can be construed as rational, or whether they trust too much or too little. Across two studies conducted in two different countries, we show that the answer to this simple question is both complicated and a little surprising. We note that the term "rational" can take on many meanings and is difficult to define. What we mean here by the term is whether people choose the "best response" given their expectations and general level of risk tolerance. In its strongest form, we mean that people form accurate beliefs about the trustworthiness of their peers and then make decisions to trust, based on those beliefs, that match what they would do if faced with a simple lottery possessing the same payoff structure and expectation of reward. Therefore, we emphasize the accuracy of participants' beliefs as well as one specific dimension related to rationality, the inter-situational consistency in levels of risk tolerance (Camerer, 1995; von Neumann & Morgenstern, 1944). In a weaker form, we allow people's expectations of their peer's trustworthiness to differ from reality, and then still see if their decisions to trust are consistent with those (mistaken) expectations and their general level of risk tolerance.

Thus, we look to see if people trust too much or too little from two vantage points. The first would be the best response for participants given their risk tolerance and the rates of reward we observe empirically in our experiments. The second would be the best response for participants given their subjective expectations of reward, regardless of whether those expectations are accurate.

2. Defining trust

How is trust to be defined and measured? In different social sciences, one can find a myriad of varying definitions for trust (see McKnight & Chervany, 2001). According to neoclassical economics, people should trust others if and only if it is in their material self-interest to do so. Ultimately, this definition means that people should trust only when it is also in the self-interest of the person being trusted to respond in a mutually rewarding manner – such as when a trustworthiness is forced or when the failure of such a response is subject to punishment (Kramer, 1998; Williamson, 1993). Such a definition of trust has a long philosophical tradition back to Socrates and Plato and can also be found in Machiavelli's *The prince* and Hobbes' *Leviathan* (see Baier, 1986).

However, from a psychological perspective and everyday experience, trust seems to be something more. People trust others even when there is no guarantee that the trustee will respond benevolently. Trust implies an awareness of being vulnerable to and dependent on the trustee – and still taking the risk of being exploited. In line with this reasoning, Rousseau, Sitkin, Burt, and Camerer (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). This definition is designed to best capture the behavior people engage in when they let a friend borrow money or agree to let someone lease their house for the period they will be out of town.

3. Assessing trust

How does one assess people's willingness to trust? One can survey people about their trust behavior (e.g., Glaeser, Laibson, Scheinkman, & Souter, 2000), but often responses to survey questions do not mirror what people chose to do when they make actual decisions of real consequence (e.g., Balceris & Dunning, 2008; Epley & Dunning, 2000, 2006). In recent years, however, behavioral economists have developed a behavioral measure of trust that fits the Rousseau et al. (1998) definition and that can be studied in a laboratory setting via an economic game, variously termed the *trust game* or *investment game* (Berg, Dickhaut, & McCabe, 1995; for a review, see Camerer, 2003).

Consider one variation of this game, which involves two people, a chooser and a responder. Both individuals remain anonymous to each other. Furthermore, both persons are informed that they will interact with each other only once. The chooser receives \$5 from an experimenter and then has to decide between two options: keeping the \$5 or giving it to the responder. If the chooser keeps the money, the interaction is over and the responder receives no money. However, if the chooser gives the money to the responder, the experimenter adds another \$15 so that the responder receives \$20 in total. The responder then has to choose between two options. The first option is to keep \$10 and give \$10 back to the chooser (splitting the money equally). The second option is to keep the entire \$20, leaving nothing for the chooser.

The decision of the chooser in this scenario is a choice about trust, in that giving the \$5 means making one's self open to exploitation by the responder. Under a neoclassical economics analysis, the decision that a rational chooser should reach is rather clear. Given that there is no material incentive for the responder to split the money evenly, the chooser should anticipate that giving the \$5 would result in no money in return. Thus, the chooser should keep the \$5 not to become vulnerable to the responder's certain exploitation (Berg et al., 1995). This result serves as the Nash equilibrium of behavior for this situation.

However, when the trust game is played in the laboratory, people violate the tenets of this neoclassical economic analysis pervasively. Choosers trust responders at substantial rates – even when a favorable response is far from guaranteed, their identities are not known to the responder, and even when their choices are anonymous to both responder and experimenter. Perhaps more vexing for a neoclassical economic analysis is that responders often turn out to be trustworthy – giving money back to the chooser when they have no compelling economic reason to do so, and even when their actions are completely anonymous to outside observers (Berg et al., 1995; Bolle, 1998; Buchan, Croson, & Dawes, 2002; Croson & Buchan, 1999;

Snijders & Keren, 2001). The fact that choosers trust and that responders prove to be trustworthy has led to a good deal of research trying to explain this economic anomaly.

4. Can trust behavior be rational?

Such data suggest that people's decisions to trust fail to follow a rational actor model and violate the predictions of a neoclassical economic analysis. However, there is a way in which decisions to trust may yet still prove to be rational. By rational, we mean that people's decisions are based on accurate expectations about how their peers respond to trust, and that decisions to trust are consistent with other economic choices that present the same expectations of reward.

Let us assume that choosers do not follow the exact economic analysis described above, but recognize instead that some percentage of responders will reciprocate trust out of altruistic motives, moral principles, guilt, or even because they misread the instructions. This knowledge changes the complexion of the trust game for them. This knowledge means that choosers might still maximize their monetary outcomes when they decide to trust. They know that some responders will reciprocate their trust, and if the chances of reciprocation are high enough so that the risk is acceptable, then they should opt to hand their money over.

However, in doing so, an accurate estimate of others' trustworthiness is essential for making the "right" decision. If choosers accurately predict the percentage of their peers who prove trustworthy, and if that percentage is an acceptable level of risk, then choosers can make wise and informed decisions to trust those peers. This analysis places under close scrutiny whether choosers can accurately estimate the likelihood that their peers will honor their trust. However, if choosers overestimate the beneficence of their peers, this could lead them to trust too much, in that they will trust at rates far too high given their general level of risk tolerance or aversion. Alternatively, if choosers underestimate the trustworthiness of their peers, this could lead them to trust too little – foregoing opportunities to trust that they would take if they knew just how trustworthy their peers prove to be.

5. Accuracy of social expectations

When people estimate just how trustworthy their peers will prove to be, there exist three logical outcomes. On average, people might (a) be roughly accurate in their estimates; (b) systematically underestimate the trustworthiness of others, or (c) systematically overestimate the trustworthiness of others. In turn, these possibilities can be labeled the *realism*, *cynicism*, and *gullibility* hypotheses, respectively. To our knowledge, two of these three possibilities are supported by extant empirical research. The one not supported is the *gullibility* hypothesis.

6. Evidence for the realism hypothesis

Extant examples of research exist suggesting that people are largely accurate in their estimates of others. When predicting pro-social behavior, people tend to be largely accurate in their estimates of their peers – especially compared against the biased estimates they provide for their own behavior (Balceris & Dunning, 2008; Epley & Dunning, 2000, 2006). People, for example, are largely accurate in their predictions of the percentage of their peers who will buy a flower for charity, cooperate in a prisoner's dilemma game, vote in a Presidential election, or donate part of their fee for participating in an experiment to charity. Similarly, Fetchenhauer and van der Vegt (2001) used data from the World Value Survey to demonstrate that on a cross-country level the average of people's trust of their fellow citizens was substantially related to actual rates of financial honesty. These examples reaffirm nicely the conclusions of Nisbett and Kunda (1985), who showed that people had largely accurate impressions of the "social distribution" of attitudes, opinions, and behaviors among their peers, such as the percentage of peers who opposed abortion, or who favored MacDonald's hamburgers.

7. Evidence for the cynicism hypothesis

However, a good deal of research also suggests that people might be too skeptical about others' trustworthiness. People tend to overestimate the power of self-interest in the behavior of others (Miller & Ratner, 1998). They overrate the impact of financial incentives on willingness to give blood, and overestimate the impact of having a vested interest on people's attitudes toward drinking and smoking. Miller (1999) explained these results by arguing that people in Western cultures are socialized to accept the selfishness axiom of neoclassical economics as a valid description of human behavior, even though its influence on actual behavior is diluted by other influences. Closer in spirit to the research to be reported here, Fetchenhauer and Buunk (2003) demonstrated that people underestimate the degree to which other people follow fairness norms in economic games, such as in the dictator and ultimatum games.

In short, the literature is rather mixed about whether people either accurately estimate the trustworthy behavior of others or tend to underestimate it. Thus, in framing this research, we decided to pit the realism and cynicism hypotheses against each other. We presented participants with an opportunity to play the trust game, and asked them to estimate the percentage of their peers who would prove to be trustworthy. The realism hypothesis suggests that the average estimate would not systematically differ from the actual values; the cynicism hypothesis suggests that they would offer estimates that were too skeptical – and, thus, trust too little.

8. Consistency with risk tolerance

However, to gauge whether participants trusted at rates that were rational, we also had to measure their decisions relative to one other important variable – their tolerance for risk. We did so by comparing the decisions participants reached with their decisions when playing a non-social financial game that presented information only about risks and payoffs. In both studies, we presented participants with a lottery in which they could use their \$5 to buy a ticket. If they bought a ticket and won, their money was doubled. If they lost, they lost the \$5. However, the chance of winning in the lottery was uncertain. We then asked participants to tell us the minimum chance of winning that they would require to induce them to gamble on the lottery.

By asking this risk tolerance question, we could assess whether people made trust decisions that matched the decisions they would make in a non-social game of risk. There were again three possible results of a comparison between participants' level of risk tolerance in the lottery and in the trust game. First, people might not differ in their willingness to take a risk between the trust game and a lottery. There is a long tradition of regarding trust simply as a risky choice that can be described and explained by a general theory of risk taking (e.g., Williamson, 1993). This is the prediction derived from normative game theory and the one we term rational.

Second, people might be more risk averse in a trust game than in a lottery. From a psychological perspective, such a behavior might be motivated by an aversion toward being betrayed. Losing in a lottery is just bad luck, but losing one's money in a trust game means that one is a chump. Bohnet and Zeckhauser (2004) found some evidence for betrayal aversion. When asking participants to indicate a "minimum accepted probability" (MAP) needed to risk trusting another person versus bet on a lottery, participants demanded an MAP considerably higher in the trust game than in the lottery. However, in Hong and Bohnet (2007) this overall pattern of results could not be replicated. One has to note that in the Bohnet and Zeckhauser paradigm participants did not actually *estimate* the likelihood of being paired with a trustworthy person, but rather had to state the MAP that they would find acceptable to play the trust game.

The third possibility is that participants would be more willing to take a risk in the trust game than in the lottery. Why could that be? Deciding to "play it safe" in a trust game can be perceived as a signal being sent to the responder that he or she is not trusted. Sending such a signal might feel uncomfortable. Much research has shown just how powerfully these types of emotions can influence interpersonal behavior of this sort (see Miller, 2006, for a review).

Beyond exploring these possibilities, we should note again that examining the rationality of decisions to trust really meant examining two different types of rationality. Under *strong rationality*, participants would show both an accurate understanding of the trustworthiness of their peers (i.e., the realism hypothesis would win out) and reach a decision that would be consistent with their general level of risk tolerance. However, there also exists a *weak* form of rationality. Even if participants are wrong in their estimates of their peers (i.e., the cynicism hypothesis wins out), participants could still be rational if their decisions to trust were consistent with those cynical estimates and their general tolerance for risk.

9. Overview of the present research

Across two studies, we examined whether people trust too little, too much, or just as much as they should, given their expectations of trustworthiness and their tolerance for risk. We presented participants with a trust game in which they could hand over money to another person with the possibility that they might get twice that amount in return. We examined whether participants provided accurate or overly cynical estimates of their trustworthiness of their peers, as well as whether these beliefs predicted their decisions to trust.

Beyond this, we also compared subjects' level of risk tolerance expressed in the trust game versus a lottery, examining whether participants trusted their peers only when their estimates of their peers' trustworthiness matched or exceeded the level of risk they would accept in order to gamble in the lottery. In doing so, we could see if participants trusted at a rate that was rational, both in the strong and the weak sense—that is, matched their willingness to gamble on other non-social decisions.

10. Study 1

In Study 1, we examined whether people would be accurate or overly cynical about the proportion of their peers who would prove trustworthy in a trust game. We then looked to see if their subsequent decisions to trust were consistent with their level of risk tolerance that they indicated in the lottery. Study 1 actually comprised three different sub-studies, each with slightly different aims. Study 1a rendered the decisions of the chooser and responder completely anonymous, even to the experimenter. Study 1b varied the speed with which participants made their decisions. Study 1c counterbalanced the order of decisions confronting participants to see if it had an impact on trust decisions.

10.1. Method: study 1a

Participants. Participants in Study 1a consisted of 167 undergraduate students at Cornell University.

Procedure. Each session consisted of 30–50 participants. Participants were asked to make a total of five financial decisions and then were informed that a third of them, at random, would be chosen to play one of these decisions for actual money.

One of the financial decisions involved a version of the trust game. Participants were told to consider a situation with two people (Person A and Person B) who would never know the identity of the other. Person A would receive \$5 from the experimenter and then be asked to choose whether to keep the \$5 or give it to Person B. If Person A kept the money, the interaction was over and Person B received no money. If Person A gave money to Person B, the experimenter would add another \$15 so that Person B received \$20 total. Person B would then have to choose between two options. The first would be to keep \$10 and give \$10 back to Person A (i.e., split the money equally). The second option would be to keep the entire \$20, leaving nothing for Person A. Participants indicated how they would behave as both Person A and Person B. Furthermore, they were asked to estimate the percentage of participants who would keep all \$20 for themselves versus give back \$10 in the position of Person B.

Participants were told that they would make their decisions under conditions of anonymity. Neither their interaction partner nor the experimenter would know which decision they had made. To ensure anonymity, participants were asked to indicate a code word on their questionnaire. They were told that if they were chosen to make their decision in the trust game for real money they would be randomly assigned to be Person A or to be Person B (with the other Person being one of the other participants of the experiment).

Another financial decision that participants had to make was designed to assess their level of risk tolerance. In the scenario, they received \$5 from the experimenter and could choose between keeping that \$5 for sure or using it to buy a lottery ticket. If they won that lottery, they would receive \$10; if they lost, they would receive nothing. Participants were told the lottery was decided by blindly drawing a ball from a basket that was full of red and white balls. If a white ball were drawn they would win; if a red ball were drawn they lost. Participants were asked to indicate how many of the 100 balls in the basket had to be white, at minimum, for them to buy a lottery ticket and participate in the lottery. This number was taken as our measurement of a participant's, level of risk tolerance, that is, the minimum chance they required in order to choose to gamble on the lottery.

After the respondents had finished their questionnaire, the experimenter announced that the trust game was the decision to be played for actual money. A third of participants were randomly chosen, and then randomly assigned to be Person A or Person B and then were randomly paired with another participant. Depending on which decisions were made, they received \$0, \$5, \$10 or \$20, which was put in an envelope with their code word. Participants retrieved their money by anonymously picking up this envelope at the end of the session.

10.2. Method: study 1b

Participants. Participants in Study 1b consisted of 106 undergraduate students at Cornell University.

Procedure. The procedure for Study 1b was largely similar to that above. Participants made a series of five financial decisions, including the trust game and lottery scenarios as in Study 1a. However, participants were run in smaller groups of 5–20. There was also no guarantee that their responses would be unknown to the experimenter. They were told that one of their five economic decisions would be carried out for real. That decision was revealed after participants had completed all their decisions, and involved deciding whether to gamble \$5 on a flip of a coin. If they won, they received \$10; if they lost, they received nothing.

Study 1b also included other experimental manipulations that did not influence the results discussed below and, thus, are not discussed further. Roughly half of the participants were asked to provide an analytical rationale for their decision to trust as Person A (or to give money back as Person B). The other half of the participants was asked to make a rapid "gut" decision.

10.3. Method: study 1c

Participants. Participants in Study 1c were 63 undergraduate students at Cornell University.

Procedure. Participants made a series of three decisions taken from Study 1b: the trust game, the lottery scenarios as above, and a coin flip. Participants were run in two groups of 20 or 43 participants, respectively, and there was also no guarantee that their responses would be unknown to the experimenter. They were told that one of their decisions would be carried out, and were informed after making all decisions that this decision was the coin flip.

Study 1c examined the impact of counterbalancing factors on decisions to trust. Three counterbalancing factors were manipulated. One was whether participants confronted the trust game or the lottery scenario first. The second was whether participants answered as Person A or B first. The third was whether the option to keep the \$5 or give it to the responder was presented first to the participant. We should note that counterbalancing the order of tasks and response options had no significant impact on participants' decisions to trust or to reciprocate trust, and thus this type of variable is discussed no further.

11. Results

Results of the three studies are discussed together. The findings in each individual study are the same unless noted otherwise.

11.1. Estimates of trustworthiness

Our central focus is whether participants accurately predicted the percentage of partners (i.e., Person B) would prove trustworthy, splitting the money and giving \$10 back. Participants turned out to be rather skeptical about the

trustworthiness of their fellow participants. On average, across the three studies, they estimated that only 44.7% of their fellow students would give money back – thus forecasting that a majority of their peers would keep the money in this situation (see Table 1). This skepticism, however, was largely misplaced. Across the three studies, 78.6% of participants stated they would give \$10 back in the role of Person B. Needless to say, participants in all studies significantly underestimated the trustworthiness of their peers, $t(166) = -18.32, p < .001$, in Study 1a and $t(104) = -12.83, p < .001$, in Study 1b, and $t(62) = -11.45, p < .001$ in Study 1c. Indeed, 80% of participants in Study 1a, 94% of those in Study 1b, and 90% in Study 1c underestimated the rate at which their peers chose to hand money back.

11.2. Relation to trust behavior

This undue cynicism about those playing the role of Player B mattered. Participants' decisions to trust across all studies (coded as "1" if they trusted and "0" if they did not) were significantly related to their estimates that Person B would return money back, with those keeping the \$5, relative to those who gave the \$5, thinking that fewer of their peers would turn out to be trustworthy, $Z = 6.25, p < .001$, with $rs(158, 104) = .28$ and $.41$ in Studies 1a and 1b, respectively, both $ps < .001$, and $r(61) = .16, ns$, in Study 1c.

To test, whether this relationship was mainly driven by a false consensus effect (i.e., participants' estimates of their fellow participants' trustworthiness being mainly determined by their own personal behavior as a trustee; Ross, Greene, & House, 1977), we conducted partial correlational analyses. The relationship between participant's decisions to trust and their estimates about others' trustworthiness remained significant even after a participant's own choices as Person B were controlled for, $Z = 3.48, p < .001$, with partial $rs(157, 103) = .19$ and $.27, ps < .01$, for Studies 1a and 1b, respectively, and $.04, ns$, for Study 1c.

11.3. Rationality of trust behavior

Participants' rather cynical views of their peers also prompted them to trust too little, given their stated tolerance for risk and the fact that nearly 80% of their peers chose to reciprocate trust. Analyses suggested that a greater number of participants would have trusted had they known just how many of their peers were willing to reciprocate that trust. Recall that we had asked participants what their chances would have to be in a lottery for them to bet \$5 with the prospect of winning \$10. Most participants revealed a certain degree of risk aversion. On average, they were willing to participate only if their chance to win was at least 62% in each of the three studies.

From the level of risk tolerance expressed by each participant, as well as the rate of trustworthiness we observed empirically (nearly 80%) in each study, we can calculate the proportion of participants who should have rationality chosen to trust, had they known the true rate of trustworthiness among their peers. In each study, we compared the objective rate of trustworthiness we observed with the minimum chance of winning each participant required to gamble in the lottery. If the objective rate of trustworthiness observed was equal or greater than this minimum chance that participant stated, we considered it rational for that participant to have trusted Person B (e.g., if the actual rate of reward in the trust game approached 80% and someone would gamble in the lottery only if the chance to win was higher than 70%). If, however, a participant demanded a winning probability that was higher than the objective rate at which trust was honored, we considered it rational for the participant to keep the money (e.g., if someone was to play in the lottery only if the chance to win was higher than 90%).

Given this analysis, across all studies, 81.2% of participants would have been rationally willing to trust their partner had they accurately estimated that the chance of receiving money back was nearly 80% (see Table 2). The actual rate of trust across all studies, 64.4%, fell significantly below this value across both studies, $p < .0001$ for sign test combining data across studies. However, there is a caveat, in that this result was significant in Study 1b, ($p < .0005$), and Study 1c ($p < .001$) but not in Study 1a, $p = .50$ (see Table 2). Thus, under the assumptions of strong rationality, participants trusted too little.

11.4. Rationality compared with subjective beliefs

However, comparing participants' decisions to trust against their rather cynical subjective beliefs suggested that people, in a dramatic sense, trusted too much. On average, participants thought it was more likely that their peers would keep all the money rather than give some back – yet a clear majority of participants trusted this peer and gave up their initial \$5. Thus, even in a weak sense, their decisions to trust proved not to be rational.

Table 1

Predicted versus actual rates of trustworthiness on the part of Person B (Study 1)

Study	Percentage of others who are trustworthy	
	Predicted	Actual
1A	43.6	77.8
1B	46.4	80.2
1C	44.5	77.8
Overall	44.7	78.6

Table 2

Actual rate of trust compared with rational rate of trust given objective reality or subjective beliefs about others' trustworthiness (Study 1)

Study	Actual rate of trust	Percentage who should rationally trust given	
		Objective reality	Subjective belief
1A	74.7	78.9	30.4*
1B	64.2	86.7*	30.8*
1C	49.2	77.8**	28.6*
Overall	64.4	81.2**	30.2**

Significantly different from actual trust rate: * $p < .005$ ** $p < .0001$.

This observation is affirmed by statistical analysis. For each participant, we calculated whether it would be rational to trust given their level of risk tolerance and subjective estimates of their peers' behavior. We did so by comparing participants' subjective estimates of their peers' trustworthiness to the minimum chance of winning they required to bet on the lottery. If their estimates of trustworthiness were equal to or greater than that minimum chance, we presumed it would be rational to give their money to Person B (e.g., if someone was to play in the lottery only if the chance to win was higher than 50% and they thought that their peers would prove to be trustworthy 60% of the time). If, however, their estimate of trustworthiness was less than the minimum probability demanded in the lottery, we presumed it would be rationale for them to keep their money (e.g., if someone was to play in the lottery only if the chance to win was higher than 64% and they thought that only 40% of their peers would be trustworthy). Based on these calculations, only 30.2% participants across all studies should have rationally given their money to Person B. However, note that the actual rate of trust was nearly 35% higher, a result that was significant across the three studies ($p < .0001$ by combined sign test across studies) as well as in each individual study ($p < .02$ in each study by sign test).

In short, given their rather cynical views of their peers, participants trusted too much. They accepted risks in the trust game that they would not have accepted in the lottery scenario. Thus, their decisions did not display a consistent level of risk tolerance, suggesting that their decisions to trust were not rational even in the weak sense.

12. Discussion

Study 1 was designed to see if people behave in the trust game in a way that normative game theory would have predicted. We found that participants' behavior did not match normative principles in two fundamental, but crosscutting, ways. First, participants underestimated how trustworthy their peers were by 35%, leading them to trust others too little, given their risk preferences and the objective rate of trustworthiness displayed by their peers. Participants' cynicism cost them, in that the vast majority of them would have been rewarded if they had chosen to give their \$5 to the other person they were anonymously paired with.

That said, according to normative tenets, participants trusted way too much given their subjective (albeit mistaken) estimates of trustworthiness coupled with their tolerance for risk. Only around 30% of participants rationally should have chosen to trust their peers given their high level of cynicism coupled with a notable tendency toward risk aversion. However, a clear majority across the three studies decided to hand over their money to an anonymous peer – many of whom doing so even though they thought it was more likely than not that they would never receive any money back. Thus, whether defined in a strong or weak sense, participants' decisions differed from what we would term as rational.

13. Study 2

Study 2 was designed to extend the previous study in several ways. First, we wanted to test the robustness and cross-cultural stability of our findings by presenting the trust situation to individuals from another country, namely, the Netherlands. The behavior in the trust game might well be different in both countries. As Weigel, Hessing, and Elffers (1999) have shown, students from the United States of America score higher on measures of egoism and Machiavellianism than students from the Netherlands.

Second, one might question the validity of the results of Study 1 by pointing to the fact that only one-third of the participants made one of their decisions for actual money in Study 1a and participants in both Study 1b and Study 1c knew that they might not have to act on their decisions in the trust game. The potentially hypothetical nature of these decisions might have driven participants to make their decisions rather capriciously or with a good dose of wishful thinking. Past work shows that people tend to predict that they will act in a far more socially desirable way when considering a hypothetical situation than they face the decision for real (Balceris & Dunning, 2008; Epley & Dunning, 2000, 2006). Thus, in Study 2, we compared decisions to trust when the decision was completely hypothetical versus potentially real.

Third, another issue arising in Study 1 is the rather artificial way in which participants received their money. Participants were not dealing with their own money but rather with funds that had been handed to them by an experimenter. Perhaps participants considered this "play money" that they could gamble or be generous with. In occasional discussions with some of the participants, many of them mentioned as a reason for their trustfulness that "it was only \$5" and that they "did start

with no money at all". There might be some truth in this remark. That said, we should note that participants also dealt with "free money" when they made gambling decisions in the lottery scenario, and their risk averse preferences seemed to belie whether they thought it was "just \$5" in that circumstance.

However, in real life situations, people rarely have to make a decision about money that has just been handed to them. Most of the time, to have money, one must work for it. To test whether people would be so generous with earned money, we asked participants whether they would give money they had just earned, by participating in an hour-long experiment, to another person in a trust game, as well as whether they would risk that money in a lottery.

14. Method

Participants. A total of 118 students from the University of Groningen in the Netherlands participated in this study.

Procedure. Participants were invited to the lab for a study that involved watching videotapes of 25 opposite-sex individuals and rating these them along a number of different dimensions. About 20–30 participants participated in each session, which on average took about one hour. Participants, none of whom were enrolled in psychology classes, received €7.50 for their participation.

After they had rated the videotaped individuals, participants were randomly split into two experimental conditions, a hypothetical versus an actual decision condition. In the hypothetical decision condition, participants were asked to indicate which financial decisions they would make in a number of different situations where they had the opportunity to invest the €7.50 they had just earned. However, participants were informed that their decisions were strictly hypothetical and would have no real consequences for them. Nonetheless, they were explicitly encouraged to imagine they were making their decision with regard to the €7.50 they had just earned. In the actual decision condition, participants were informed that they were to make a number of financial decisions with the €7.50 they had just earned and that one of these decisions would be for real money.

Participants in both conditions confronted a version of the trust game, in which Person A received €7.50 and had the chance either to keep the money or to give it to Person B. If Person A decided to give the money to Person B, then Person B would receive €30 from the experimenter and had to choose between an equal split of the €30, giving €15 back to Person A, or keeping all the money and returning nothing. As in Study 1, participants were asked to indicate their own decisions as both Person A and Person B, and estimated the percentage of their fellow participants who would give money back as Person B. In both conditions, participants were informed that they were to make their decisions under conditions of total anonymity. For this purpose, participants in the actual decision condition were asked to provide a code word on their questionnaire and were informed that they would receive their money in a sealed envelope with their codeword written on it at the end of the session.

Participants' level of risk-tolerance was measured similarly to Study 1. They were told that they could either keep €7.50 for sure or that they could buy a ticket for a lottery with that money. In case their ticket won they would receive €15; in case their ticket lost, they would receive nothing. Like in Study 1, participants were informed that winning or losing would be decided by blind draw of a ball out of a basket with 100 balls, and participants indicated the minimum number of winning balls they would require before to induce them to buy a lottery ticket.

After all participants had completed their questionnaires, they were debriefed about the purpose of the study. In both conditions, participants received €15 Euro for their participation.

15. Results

15.1. Estimates of trustworthiness

As in Study 1, participants were rather skeptical about the trustworthiness of their fellow participants. On average, they estimated that 59.1% of their peers would give money back, an estimate that did not differ across hypothetical or actual conditions, $t(116) < 1$. These estimates proved to be overly cynical, in that a vast majority of 90.4% turned out to be trustworthy and divided the €30 equally between themselves and their partner. This rate was nearly identical in both experimental conditions, $z = .33$, ns. In both conditions, the underestimation of others' trustworthiness was highly significant, $t(68) = -10.72$ in the hypothetical decision condition and $t(48) = -9.12$ in the actual decision condition, $p < .001$. This underestimation of others' trustworthiness was not without consequences. The less people felt others were trustworthy, the less were they willing to give their money to Person B, $r(68) = .32$ in the hypothetical decision condition and $r(48) = .42$ in the actual decision condition, $p < .01$. In the position of Person A, only 45.3% gave their €7.50 to Person B whereas 54.7% kept the money for themselves.

15.2. Rationality of trust decisions

This trust rate was lower than it should have been based on empirically observed rates of trustworthiness and participants' tolerance for risk. As in Study 1, we calculated the percentage of participants who would have rationally trusted

Table 3

Actual rate of trust compared with rational rate of trust given objective reality or subjective beliefs about others' trustworthiness (Study 2)

Decision status	Actual rate of trust	Percentage who should rationally trust given	
		Objective reality	Subjective belief
Hypothetical	37.7	92.6 ^{**}	31.3 ^{**}
Actual	56.3	72.9 ^{**}	29.8 ^{**}
Overall	45.3	84.5	30.7

Significantly different from actual trust rate: ^{**} $p < .01$.

Person A, based on their responses to the lottery scenario, had they known that there was a roughly 90% chance of getting money back in the trust game. Across hypothetical and actual conditions (see Table 3), that percentage was 84.5%, a figure that is significantly higher than the rates of trust we actually observed, $z = 6.26$, $p < .001$. Observed rates of trust were significantly lower than this normative benchmark in both the hypothetical and actual conditions, both $p < .01$.

However, given their level of cynicism, participants once again trusted much too much. As in Study 1, we again calculated the percentage of participants who would have rationally trusted by comparing their level of risk tolerance in the lottery scenario against their subjective expectations they had that their peers would return some of the money back. According to this analysis, only 30.7% of participants across the two conditions should have handed their money over to Person B – a figure significantly lower than observed rates of trust, $z = -2.66$, $p < .01$.

15.3. Real versus hypothetical decisions

Did participants act contrary to normative game theoretic tenets because of the potentially hypothetical nature of their decisions? A final set of analyses definitively ruled this possibility out. Making the decision more real rather than hypothetical influenced participants' preferences and decisions, but not totally in a way that would be expected. When the decision was real, participants were *more*, not less, likely to trust than when the decision was purely hypothetical, $z = 1.96$, $p < .05$ (see Table 3).

Interestingly, people were more willing to trust – and, thus, to take a risk – when the decision was real even though that circumstance made them more risk averse in the lottery scenario. When the lottery decision was potentially real, participants demanded on average a 75.3% likelihood of winning, but when the decision was purely hypothetical they demanded only 66.8%, $t(114) = 2.24$, $p < .05$. This difference was mainly due to the different percentage of participants who decided not to take any risk at all in the lottery (i.e., who demanded a 100% chance of winning). This extreme level of risk aversion was revealed by only 4.4% of all participants in the hypothetical decision condition but by 25.0% of all participants in the actual decision condition. Of those in the actual condition who demanded a 100% certainty in the real condition, 36.4% nonetheless decided to trust Person B (although on average they expected that 40.0% of others would not return any money back to them).

These differences in risk tolerance and trust behavior led to one qualification of our analysis about whether participants trusted too much given their cynical views. In the hypothetical condition, the percentage of participants who trusted did not exceed the percentage who should have trusted given their risk tolerance and beliefs about their peers, $p = .14$. However, in the actual condition, participants' rates of trust significantly exceeded the rate they should have trusted if they had acted rationally, $p < .01$.

16. Discussion

The results of Study 2 largely replicated those of Study 1. Again, a vast majority of participants turned out to be trustworthy, but their peers underestimated this trustworthiness by nearly 35%. Second, participants trusted too little given objective circumstances. Had they known that others were so uniformly trustworthy, roughly 25% more of them should have chosen to trust given their stated tolerance for risk. In short, people trust too little compared with empirically observed reality.

However, in another sense, participants trusted too much given their rather pessimistic views of their contemporaries. Based on their level of risk aversion and their gloomy estimates of their peers' trustworthiness, trust rates were 15% higher than they should have been. Importantly, this discrepancy was not caused by the potentially hypothetical nature of participants' trust decisions. When decisions were purely hypothetical, rates of trust did not exceed the degree suggested by levels of cynicism and tolerance for risk. However, when the decision potentially involved real money, rates of trust rose significantly, despite the fact that participants became more risk-averse with their lottery decisions. In short, these data suggest that, compared against their cynical expectations of others, people trust much too much.

Further, the basic findings of Study 1 – namely that people trust too much and trust too little at the same time – cannot be explained by arguing that participants construed their decisions in Study 1 as purely hypothetical or involving "free" money that had been handed to them by the experimenter to play with. The same counter-normative patterns of trust were found in

Study 2 and they were even stronger when participants were making their choice for real money that had not simply been given to them but which they earned for one hour's work.¹

16.1. General discussion

The main question motivating the research in this article was whether people trust too much or too little, relative to what an economic analysis would suggest. Traditional neoclassical treatments of trust suggest that people should trust others not at all – in that they should never expect their trust to be reciprocated. We, however, noted that trust is often reciprocated, leaving open the possibility that people might be making rational decisions to trust. People would be making such rational decisions if two conditions were met: they accurately anticipated the rate at which their peers would honor their trust and they, thus, made decisions that were consistent with their general level of risk tolerance. However, across our studies, we found that participants' decisions were not rational under this decision. They trusted both too much and too little, a complicated answer to our original question and somewhat of a paradox. Let us consider both sides of the paradox in turn.

17. People trust too little

On one hand, participants consistently displayed too little trust in their peers. In reality, an overwhelming majority of participants, sometimes as high as 90%, chose to honor trust and return money to someone who had invested money in them. Participants, however, underestimated this rate by some 30–35%, depending on the study. As a consequence, participants trusted others at a rate that was significantly lower than the rate suggested by their tolerance for risk. A large majority of participants indicated that they were quite willing to gamble on a lottery in which the likelihood of winning was 80–90%. However, because participants thought the chances that their trust would be honored was far lower than that rate, they passed up chances to trust that they otherwise would have taken – and profited from.

18. Costs of cynicism

It is worth mentioning that the undue skepticism our participants expressed about their peers was costly. The lower participants estimated others' trustworthiness, the less often they gave their money to Person B and the less money they took home. Given that the vast majority of Persons B gave money back, the money maximizing strategy for Persons A was to hand their money over to Person B. For example, in Study 2, the expected value for taking the safe option was €7.50, whereas the expected value for trusting Person B was €13.56 (as 90.4% of all Persons B split the €30 equally).

Findings from other work suggest that the cynicism we observed might underlie other more subtle yet important costs. In Pilutla, Malhotra, and Murnigham (2003), participants played a version of the trust game in which Persons A were free to choose how much money they wanted to hand over to Person B. Similarly, Persons B were free to choose how much money they wanted to give back to Person A. The results revealed that Persons A were more richly rewarded when they gave all their money to Person B rather than keeping everything less for themselves. However, presumably driven by their cynicism, which was not measured, many participants gave only a part of their money to Person B. Unfortunately, such behavior was perceived by Persons B as a signal of suspicion – and their reaction in turn was to withhold giving much money back to Persons A.

Still other evidence for possible hidden costs of cynicism stems from research by Fehr and Rockenbach (2002). Participants had to choose whether to trust another person, but with a twist. Participants could pay a premium to sanction the behavior of others if those others proved inadequately trustworthy. A large number of participants chose this option. However, such a choice had an unintended consequence of undermining the prosocial motivation of Persons B, who returned much less money when under threat of sanction (for similar results see Malhotra & Murnigham, 2002). In a similar vein, Mulder, Van Dijk, De Cremer, and Wilke (2006) demonstrated that sanctioning systems undermine people's intrinsic motivation to act in a trustworthy manner.

19. Accounting for cynicism

The two studies presented in this manuscript provide a good deal of information about how people approach decisions to trust. However, they also leave open many questions that future research could profitably address. First, future research

¹ It is interesting to compare results of Study 2 with a study by Fahr and Irlenbusch (2000). In their study, either Person A (the trustor) or Person B (the trustee) or both had to work to get into their position. Fahr and Irlenbusch found that trustees were more trustworthy when the trustor, but not they themselves, had to work for participating in the trust game. In doing so, they honored the effort of the trustor. However, trustors gave less money to the trustee in the same situation – thus depriving themselves of the increased generosity the trustees would have displayed.

Thus, whereas in our Study 2 of participants were more trusting if they had to work for their initial endowment, Fahr and Irlenbusch (2000) found that their participants were less trusting if they had to work for their initial endowment. How can we explain the discrepancy? We would note that participants in Fahr and Irlenbusch played another version of the trust game, the so-called investment game, in which trustors could make a choice on what percentage of their money to send to the trustee, whereas participants in our study faced an all-or-nothing choice. Given this difference, it would be worthwhile to replicate their study with a binary, all-or-nothing trust game to see if rates of trust are influenced differently depending on whether trust can be partial versus an all-or-nothing choice.

could focus on why people are so overly cynical about their peers, particularly when it is costly to do so. To date, there are three different explanations for such cynicism (for a discussion, see [Fetchenhauer & Dunning, 2006](#)).

Motivated Reasoning. A first explanation would argue that such a bias arises from motivated reasoning processes. People have a strong desire to see themselves as especially ethical, pro-social, and fair individuals among their peers (e.g., [Allison, Messick, & Goethals, 1989](#); [van Lange & Sedikides, 1998](#); see [Dunning, 2005](#), for a review). In line with this theorizing, Dunning and colleagues ([Balceris & Dunning, 2008](#); [Epley & Dunning, 2000, 2006](#)) have shown across a number of circumstances that people perceive themselves to be “holier than thou”, being overly optimistic about how pro-social (e.g., donating to a charitable organization) their future behavior will be, showing no comparable bias about their peers.

Cultural Norms. A second explanation centers on the notion of cultural norms prevalent in Western industrialized countries, which teach people the neo-classical economic assumption that people tend to act out of material self-interest ([Miller, 1999](#)). Consistent with this view, [Miller and Ratner \(1998\)](#) have shown that people hold strong beliefs about the power of self-interest, beliefs that emphasize self-interest far beyond its actual power. People, for example, assume that financial incentives will have a bigger impact on whether people choose to donate blood than those incentives actually do, that the relevance of an alcohol policy to one's personal lifestyle has a stronger impact on one's attitude toward it than actually turned out to be true, and that having a smoking habit will produce more opposition to a cigarette tax than is really the case. It stands to reason that such cultural beliefs about self-interest would also influence people's estimates of whether their peers will honor trust or act in a more selfish manner.

Error management theory. Based on the error management theory by [Haselton and Buss \(2000\)](#); ([Haselton, Nettle, & Andrews, 2005](#)), [Fetchenhauer and Buunk \(2003\)](#) argue that systematic cynicism about peers is an adapted mechanism to avoid being exploited by others. When having to trust another person, people can engage in two different kinds of mistakes. They can either distrust a person who is actually trustworthy or they can trust another person is later turns out to be untrustworthy. They argue that throughout human evolution the maximal negative consequences of each type of mistake has proven to be very different. Being overly suspicious of trustworthy people simply leads to a missed chance of cooperation. Being too trusting of a malevolent individual, however, may lead to rather extreme negative consequences (e.g., being killed when turning one's back). Therefore, human evolution has resulted in a stable tendency to underestimate others' trustworthiness to avoid being costly mistakes of exploitation.

Flawed Measures of Rationality. Beside these three theoretical accounts, another way to explain the undue cynicism we observed would be to argue that our measures of rationality were flawed. The estimates that participants provided of their peers' trustworthiness may not have reflected their true beliefs, in that participants were motivated to provide estimates that were overly skeptical of their peers. Recall that we asked all participants what they would do in the role of Person B. Those who said they would keep the money as Person B might realize that such a behavior harms rules of reciprocity and distributive justice. Such a behavior seems less unfair if one claims to the experimenter that most others would have done the same. People who give money back as Person B, by underestimating percentage of their peers who would do the same, can exalt the moral quality of their own personal behavior. Thus, all participants might have been motivated to report rather pessimistic estimates of others' trustworthiness that differs from their private beliefs (see [Dunning & Cohen, 1992](#), for a similar argument).

The standard cure of experimental economics for such problems is to “incentivize” participants to give valid estimates. Following this reasoning, [Fetchenhauer \(2008\)](#) showed short videotapes of 56 individuals to participants who estimated what those individuals would do as Person B in a trust game. Consistent with our findings, participants heavily underestimated trustworthiness. Indeed, when given monetary incentives for accuracy, their estimates became more, not less, skeptical. Thus, we would argue that the underestimation of others' trustworthiness is much more than a methodological artifact.

20. People trust too much

On the other hand, participants also showed evidence of trusting too much given their skepticism about others' trustworthiness. A large number were willing to place their trust in Person B than were willing to bet on a lottery offering the same odds and payoffs. In short, the decision to trust was not based just on the rewards offered and the odds of achieving them, as an economic analysis would suggest. Many participants chose to trust even when they thought the chances were greater that they would lose their money rather than gain a profit.

Although our studies did not explain why people trust so much, we feel that the beginnings of an answer might be identifiable throughout the various patterns of data we observed, although future research is required to test this answer more directly. Specifically, the results of Study 2 reveal a first hint to what might be producing high rates of trust. Recall that participants in this study made their choices either purely hypothetically or potentially for a real €7.50 that they had earned. One might have expected that participants in the hypothetical decision condition would have given their money more often to Person B than in the actual decision condition because there was no price tag attached to any generosity. However, the results of Study 2 were exactly the other way around. This increase in trust arose even though participants became more risk averse when considering a potentially real versus a purely hypothetical chance to gamble on a lottery ticket.

These apparently counterintuitive findings can be explained if one presumes that the decision to trust has an emotional dimension attached to it. For example, perhaps deciding not to trust the other individual involves a touch of guilt or shame that arises only as the decision becomes more real and less hypothetical. There is precedent for this line of reasoning: The

level of risk that people take often diverges between hypothetical and actual decisions because people in a hypothetical situation are not able to foresee the emotions they would feel if their decision were made for real, as well as what impact those emotions will have on their thoughts and preferences (e.g., Van Boven, Loewenstein, & Dunning, 2005; Van Boven, Loewenstein, Welch, & Dunning, 2008). Indeed, emotional influences can enter trust decisions in two different ways (Loewenstein, Weber, Hsee, & Welch, 2001). People's decisions can be driven by anticipated emotions (e.g., how do I anticipate I will feel when it turns out that the other person returns no money back to me) as well as by immediate emotions (i.e., how I feel right now about saying I will give the money).

Thus, it seems plausible that participants in the actual condition of Study 2 were less risk-seeking in the lottery relative to those in the hypothetical condition because they felt the emotions associated with gambling and potentially losing. If the same logic applies to the trust game, one could argue that Persons A in the actual condition felt immediate emotions that pushed them toward giving their money to Person B rather than keeping it for themselves. There are several candidates for what this emotion might be. People might feel embarrassed when they have to signal their distrust to another person. Indeed, such an unwillingness to communicate one's distrust might be so pervasive that it works even in a one-shot game under total anonymity.

This reluctance might also explain why our results at first glance seem to contradict the results of Bohnet and Zeckhauser (2004), who found that people are less, not more, willing to trust than to play a lottery. Although the differences in paradigms between their study and our study seem to be only minor, these differences appear to have an enormous impact on participants' behavior. Our interpretation of these differences would focus on whether participants construed themselves as already placed in a trust game or not. In our studies, participants were assigned to a trust game and we gave them an option to trust another person. In Bohnet and Zeckhauser, participants were not yet participating in a trust game, but rather had the option to choose whether to participate at all. Given the high level of social cynicism revealed in our studies, it is not surprising that the participants of Bohnet and Zeckhauser set their MAP (minimum accepted probability) higher in the trust game than in a lottery.

That is, our analysis suggests that if people are not yet placed in a trust game with another person, they avoid making themselves vulnerable. However, once in a trust game, other considerations become salient, such as wishing to avoid sending a signal to the other person that he or she is not trusted – even if that signal is totally anonymous in a one-shot game. That said, more research is obviously needed to explain the differences between our studies and Bohnet and Zeckhauser (see also Fetchenhauer & Dunning, 2007).

Generally speaking, our proposals about the role of emotion in trust behavior exist solely as speculation that requires further research to affirm, refute, or refine. However, given the potentially complex role that emotions can play in the decision to trust, any future researchers would have to be sure to explore the many possibilities about how emotions might shape the behavior of those encountering the types of decisions we presented here.²

21. Implications for the study of trust

Taken as a whole, these studies also have direct implications for the future study of trust. The fact that participants expressed too little trust on the "cognitive level" (i.e., overpredicting the percentage of their peers who would respond to trust in a greedy way), but exhibited too much trust on the behavioral level (i.e., handing over their money to a stranger even when they believed it was quite likely that the stranger would pocket it), carries many implications for how trust should be studied by economists and psychologists alike. Remarkably, to our best knowledge, we are the first ones to have directly measured the expectations of participants about their fellow participants' trustworthiness. Instead, researchers usually infer participants' expectations from their behavior (much like old-school behaviorists used to assume an organism's representation of a stimulus by its overt responses toward it). That is, if Ned hands Stacey his \$5, that must mean that he thinks there is a high likelihood that Stacey will respond by returning that money with a profit. After all, why else would Ned turn over that money?

Such a procedure assumes that decisions to trust are responsive to only two factors: The value of the outcomes involved and the probabilities associated with those outcomes. The studies presented in this paper show that the issue of trust is a much more complicated matter than just simple calculations involving v (value) and p (probability). A good number of our participants believed that it was more likely than not that they would receive no money back from the stranger they were assigned to, yet they handed their money over – defying an analysis based simply on p and v : In addition, participants chose to trust a stranger even though, given the same p and v , they refused to participate in a comparable lottery. This suggests that researchers investigating trust behavior should not merely assume that they are able to determine the expecta-

² One might argue that our way of eliciting participants' level of risk aversion in both, the trust game and the lottery, not only measured risk aversion but also willingness to engage in bets that involve ambiguity, that is, uncertainty about the true level of risk the lottery entailed. Participants did not know the exact percentage of their peers who would reciprocate trust (an ambiguous situation), yet could specify an exact probability of winning in the lottery (an unambiguous situation). Could this difference have produced the discrepant behavior we observed between trust and lottery scenarios? We think not. We found that people are more likely to take risks in the trust game (the ambiguous situation), whereas past research has shown that people are more adverse to ambiguous risks than to well-specified ones (Anand, 1993; Ellsberg, 1961), counter to our results. However, the results we present herein are very similar to another study, in which participants were informed about their exact chance of winning and losing in both a lottery and a trust game (Fetchenhauer & Dunning, 2007).

tions people have of their peers from merely observing their behavior. Instead, those expectations must be directly measured, not merely inferred.

22. Concluding remarks

Where does this leave us? Basically, we would argue that we have identified two fascinating phenomena. People trust too little and trust too much at the same time. However, the studies presented in this paper have not yet been able to provide a full explanation for these phenomena, and so there is obvious room for future work to be done.

There are two clear lines to follow in future research. First, we can further investigate why people so systematically underestimate the trustworthiness of others. Such research should try to disentangle the motivational and cognitive processes that lead to such a bias and should try to answer the question whether this bias can be regarded as another example of motivated reasoning or whether such a bias can better be explained by its functional adaptiveness. Second, we can investigate further what drives people to trust others on the behavioral level, although they are so skeptical about others' trustworthiness, examining further how much the decision to trust is based on moral concerns and emotional reactions, among other factors.

Possibly, the paradoxical pattern of trusting too much and too little might not, in the end, be the worst approach to dealing with one's social environment. Indeed, these two biases might be construed as a good way to adhere to the old German proverb that begins this report: trust on a behavioral level, but be wary so as not to be exploited. Whether this combination proves to be an optimal or beneficial way to conduct one's life depends on a close scrutiny that future research might be able to provide. We place our trust in that future research.

Acknowledgement

The research described herein was supported financially by the University of Groningen and the National Institute of Mental Health Grant RO1-56072 awarded to David Dunning.

References

- Allison, S. T., Messick, D. M., & Goethals, G. R. (1989). On being better but not smarter than others: The Muhammad Ali effect. *Social Cognition*, 7, 275–296.
- Anand, P. (1993). *Foundations of rational choice under risk*. Oxford University Press: Oxford.
- Baier, A. (1986). Trust and antitrust. *Ethics*, 96, 231–260.
- Balceris, E., & Dunning, D. (2008). A mile in moccasins: How situational experience diminishes disposition in social inference. *Personality and Social Psychology Bulletin*, 34, 102–114.
- 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.
- Bolle, F. (1998). Rewarding trust: An experimental study. *Theory and Decision*, 45, 83–98.
- Buchan, N. R., Croson, R. T. A., & Dawes, R. M. (2002). Swift neighbors and persistent strangers: A cross-cultural investigation of trust and reciprocity in social exchange. *American Journal of Sociology*, 108, 168–206.
- Camerer, C. F. (1995). Individual decision making. In J. Kagel & A. Roth (Eds.), *The handbook of experimental economics* (pp. 587–703). Princeton, NJ: Princeton University Press.
- Camerer, C. F. (2003). *Behavioral game theory: Experiments on strategic interaction*. Princeton: Princeton University Press.
- Croson, R. T. A., & Buchan, N. R. (1999). Gender and culture: International experimental evidence from trust games. *American Economic Review*, 89, 386–391.
- Dunning, D. (2005). *Self-insight: Roadblocks and detours on the path to knowing thyself*. New York: Psychology Press.
- Dunning, D., & Cohen, G. L. (1992). Egocentric definitions of traits and abilities in social judgment. *Journal of Personality and Social Psychology*, 63, 341–355.
- Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *Quarterly Journal of Economics*, 75, 643–669.
- Epley, N., & Dunning, D. (2000). Feeling "holier than thou": Are self-serving assessments produced by errors in self- or in social prediction? *Journal of Personality and Social Psychology*, 79, 861–875.
- Epley, N., & Dunning, D. (2006). The mixed blessings of self-knowledge in behavioral prediction: Enhanced discrimination but exacerbated bias. *Personality and Social Psychology Bulletin*, 32, 641–655.
- 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., & Rockenbach, B. (2002). Detrimental effects of sanctions on human altruism. *Nature*, 422, 137–140.
- Fetchenhauer, D. (2008). Impact of incentives on estimates of trustworthiness. Unpublished data, University of Cologne.
- Fetchenhauer, D., & Buunk, A. P. (2003). How do people estimate others' prosociality and others' revengefulness? Unpublished manuscript, University of Groningen, The Netherlands.
- Fetchenhauer, D., & Dunning, D. (2006). Perception of prosociality in self and others. In D. Fetchenhauer, A. Flache, B. Buunk, & S. Lindenberg (Eds.), *Solidarity and prosocial behavior: An integration of psychological and sociological perspectives* (pp. 61–76). New York: Kluwer Academic/Plenum Publishers.
- Fetchenhauer, D., & Dunning, D. (2007). *Betrayal aversion versus principle trustfulness. How to explain risk avoidance and risky choices in trust games*. Cornell University: Unpublished Manuscript.
- Fetchenhauer, D., & van der Vegt, G. (2001). Honesty, trust and economic growth. A cross-cultural comparison of Western industrialized countries. *Zeitschrift für Sozialpsychologie*, 32, 189–200.
- Fukuyama, F. (1995). *Trust: The social virtues and the creation of prosperity*. New York: Free Press.
- Glaeser, E. L., Laibson, D., Scheinkman, J. A., & Souter, C. L. (2000). Measuring trust. *Quarterly Journal of Economics*, 65, 811–846.
- Haselton, M., Nettle, D., & Andrews, P. W. (2005). The evolution of cognitive bias. In D. Buss (Ed.), *The evolutionary psychology handbook* (pp. 724–746). New York: Wiley.
- Haselton, M. G., & Buss, D. M. (2000). Error management theory: A new perspective on biases in cross-sex mind reading. *Journal of Personality and Social Psychology*, 78, 81–91.
- Hong, K., & Bohnet, I. (2007). Status and distrust: The relevance of inequality and betrayal aversion. *Journal of Economic Psychology*, 28, 197–213.

- 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.
- Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, 127, 267–286.
- Malhotra, D., & Murnighan, J. K. (2002). The effects of contracts on interpersonal trust. *Administrative Science Quarterly*, 47, 534–559.
- McKnight, D., & Chervany, N. (2001). Trust and distrust definitions: One bite at a time. In R. Falcone, M. Singh, & Y.-H. Tan (Eds.), *Trust in cyber-societies: Integrating the human and artificial perspectives* (pp. 27–54). Berlin: Springer-Verlag.
- Miller, D. T. (1999). The norm of self-interest. *American Psychologist*, 54, 1053–1060.
- Miller, D. T. (2006). *An invitation to social psychology: Expressing and censoring the self*. Belmont, CA: Wadsworth-Thomson.
- Miller, D. T., & Ratner, R. K. (1998). The disparity between the actual and assumed power of self-interest. *Journal of Personality and Social Psychology*, 74, 53–62.
- Mulder, L. B., Van Dijk, E., De Cremer, D., & Wilke, H. A. M. (2006). Undermining trust and cooperation: The paradox of sanctioning systems in social dilemmas. *Journal of Experimental Social Psychology*, 42, 147–162.
- Nisbett, R. E., & Kunda, Z. (1985). Perception of social distributions. *Journal of Personality and Social Psychology*, 48, 297–311.
- 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.
- Putnam, R. D. (1993). *Making democracy work: Civic traditions in modern Italy*. Princeton, NJ: University Press.
- Ross, L., Greene, D., & House, P. (1977). The “false consensus effect”: An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, 13, 279–301.
- 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.
- Snijders, C., & Keren, G. (2001). Do you trust? Whom do you trust? When do you trust? *Advances in Group Processes*, 18, 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.
- Van Boven, L., Loewenstein, G., & Dunning, D. (2005). The illusion of courage in social prediction: Underestimating the impact of fear of embarrassment on other people. *Organizational Behavior and Human Decision Processes*, 96, 130–141.
- Van Boven, L., Loewenstein, G., Welch, N., & Dunning, D. (2008). The illusion of courage: Underestimating the impact of fear of embarrassment on the self. Unpublished manuscript, University of Colorado, Boulder.
- Van Lange, P. A. M., & Sedikides, C. (1998). Being more honest but not necessarily more intelligent than other: generality and explanations for the Muhammad Ali effect. *European Journal of Social Psychology*, 28, 675–680.
- von Neumann, J., & Morgenstern, O. (1944). *The theory of games and economic behavior*. Princeton, New York: Princeton University Press.
- Weigel, R. H., Hessing, D. J., & Elffers, H. (1999). Egoism: Concept, measurement and implications for deviance. *Psychology, Crime and Law*, 4, 349–378.
- Williamson, O. E. (1993). Calculativeness, trust, and economic organization. *Journal of Law and Economics*, 36, 453–486.