TRUST IN THE FAITH GAME

Nobuhiro MIFUNE\textsuperscript{1} and Yang LI\textsuperscript{2}

\textsuperscript{1}Kochi University of Technology, Japan\
\textsuperscript{2}University of Melbourne, Australia

General trust, or trust in people with unknown personal backgrounds and with whom there is no interpersonal relationship, is key for the formation of an effective society. The Trust Game has long been a popular behavioral measurement of general trust; however, other measurements, such as the Faith Game, are increasingly attracting attention. Nevertheless, while the psychological and neural mechanisms of trust in the Trust Game have been revealed in numerous studies, little is known about trust in the Faith Game. In the present study, we exploratorily examined how behavioral trust in both the Trust Game and Faith Game correlates with trust-related individual characteristics (i.e., attitudinal trust, social caution, and social value orientation). The results showed that attitudinal trust is only related to the trust shown in the Trust Game, and that risk aversion and betrayal aversion are related to the trust shown in both games.

Key words: betrayal aversion, Faith Game, general trust, Trust Game, risk aversion

INTRODUCTION

Trust is vital for societal functioning. From the creation of social bonds, to fostering cooperative relationships, to creating and maintaining social systems, trust plays an integral role. For over a century, trust has been a key concept in the social sciences (e.g., Hardin, 1992; Luhmann, 1973; Putnam, 1993; Simmel, 1908). The topic of trust has received academic attention. Specifically, research has focused on individual attributes that promote or prevent trust and the situations in which people trust or distrust others (e.g., Yamagishi et al., 2015).

There are various definitions of trust; however, the present paper focuses on general trust; that is, trust in people with unknown personal backgrounds and with whom there is no interpersonal relationship. The behavioral aspect of general trust is defined as voluntarily putting oneself in a vulnerable situation despite a risk of loss, by allowing another’s decisions to decide an outcome (Hardin, 1992; Yamagishi, 1998/2011). One of the most popular measurements of behavioral trust is an economic game called the “Trust Game” (Berg, Dickhaut, & McCabe, 1995). Typically, the Trust Game is played by two players, a “trustor” and a “trustee.” To begin, the trustor is presented with an endowment, and is asked to choose between two options — receiving a fixed amount of
the endowment ("sure choice") or allowing the trustee to decide how much they both receive ("trust choice"). In the case of the sure choice being selected, the trustee does not need to do anything and the game ends. On the other hand, if the trust choice is selected, the endowment is tripled, and the trustee then decides how to allocate this relatively large amount of money between the two players. Naturally, when the trustee allocates the money equally, the trustor makes a greater profit than if he/she were to have selected the sure choice. However, if the trustee allocates everything to him/herself, the trustor receives much less than he/she would have had he/she chosen the sure choice. When the game is played only once, a rational trustee who is pursuing his/her best interests is expected to keep everything for him/herself; thus, in such a one-shot game, the trustor should always choose the sure choice, which means no trust behavior should be evident in the game. However, as numerous empirical studies have demonstrated, it is not uncommon for players choose the trust choice (Camerer, 2003).

Logically speaking, it is natural to speculate that some psychological mechanisms other than pursuing self-interest drives trustors to choose the trust choice. One possibility is that the trustor is showing altruism and wants to increase the trustee’s income. Alternatively, an expectation of reciprocity, such as “something good must happen when I do something good,” may drive the Trustor to choose the trust choice, as he/she is expecting a fair allocation in response to his/her benevolent decision. However, a more straightforward explanation is that the trustor is willing to risk losing everything in order to gamble on a greater profit. While the pay-off matrix of the Trust Game can suggest various motivations for trust, psychological and neuroscientific research can contribute to identifying the precise mechanisms involved in the game. Considering this, the present paper reviews studies in the fields of psychology, experimental economics, and neuroscience that have examined the Trust Game. After that, the Faith Game, which is a modified version of the Trust Game, is introduced, and explorative analyses regarding the correlation between behavioral trust in these games and individual characteristics are performed.

Studies of the Trust Game in the Fields of Psychology and Experimental Economics

Many studies have emphasized the importance of reciprocity in the Trust Game. For example, Pillutla, Malhotra, and Murnighan (2003) showed that, when the trustor increases the amount they entrust, the trustee gives back a larger amount of money to the trustor, which indicates a reciprocal tendency from the trustee, and contradicts the game theory prediction (see also Malhotra, 2004). Moreover, King-Casas et al. (2005), conducted an investigation where the Trust Game was played over multiple rounds and the trustee and trustor retained the same roles throughout. In this study, reciprocal behavior by the trustee in a prior round was shown to foster greater trust behavior in the trustor in the following round. McCabe, Rigdon, and Smith (2003) demonstrated that such reciprocity arises when the trustee recognizes the trustor’s intentions. The authors analyzed the difference in a trustee’s reaction towards trust decisions made intentionally by a trustor and those made randomly by a computer. They consequently found that the trustee’s frequency of reciprocal reactions to the intentional trust decisions was almost
double that for the randomly selected “trust” decisions. These studies indicate that the 
trustee reciprocates trust from the trustor, and that the trustor shows more trust when the 
trustee reciprocates. However, this does not provide a sufficient explanation regarding 
why trustors can show trust during one-shot games or during the first round of a multiple-
round game.

One possible reason for the presence of such trust in one-shot Trust Games is an 
expectation of reciprocal behavior from the trustee (Berg et al., 1995). In fact, studies 
have found that the likelihood of trust behavior is increased by the expectation of 
reciprocity. For instance, Eckel and Wilson (2004) investigated, after a one-shot Trust 
Game, but before any feedback was given, how trustors’ behaviors could predict trustees’ 
decisions. The study showed that the expectation of reciprocity is the strongest predictor 
of trust behavior. Similar findings have been found in studies employing a sequential 
version of the Prisoner’s Dilemma Game, which has a similar incentive structure, with 
the first player being more likely to cooperate when he/she expects a reciprocal decision 
from the second player. For instance, Watabe, Terai, Hayashi, and Yamagishi (1996) 
reported that the first player in a sequential Prisoner’s Dilemma Game expects the 
counterpart to cooperate when the first player cooperates, and such expectation is 
stronger than the expectation of cooperation in a one-shot simultaneous Prisoner’s 
Dilemma Game. These findings illustrate that the expectation of reciprocity is a key 
factor involved in trust behaviors during the Trust Game.

It is logical to predict a positive correlation between general risk preference and trust 
in the Trust Game because selecting the trust choice is a risky decision. Surprisingly, 
however, empirical findings have shown evidence both for and against this prediction. 
For example, Eckel and Wilson (2004) explored the correlation between the following 
explanatory variables and trust in the Trust Game: 1) the Sensation Seeking Scale 
(Zuckerman, 1994) as a psychological measurement; 2) risk-taking in the form of a 
lottery, where variance of outcome was dependent; and 3) risk-taking tendency in a 
Trust-Game-like situation. However, they did not find a significant correlation. Other 
studies have also found no correlation between behavioral trust in the Trust Game and 
risk-taking in other situations (e.g., Ashraf, Bohnet, & Piankov, 2006; Kanagaretanam, 
Mestelman, Nainar, & Shehata, 2009). However, Bohnet and Zeckhauser (2004) argued 
that the above results are affected by measurement problems, and that risk should be 
considered a factor in the Trust Game. This argument is supported by a study that found 
a correlation between trust and risk-taking when risk attitude was measured using a 
format similar to the Trust Game (Schechter, 2007). Other experiments have also shown 
that when the component of risk in the Trust Game is manipulated, behavioral trust 
changes accordingly (Evans & Krueger, 2011; Snijders & Keren, 1998).

The above studies demonstrate that factors such as reciprocity, expectation of 
reciprocity, and risk attitude play important roles in the behavioral trust present in the 
Trust Game. Further, neuroscientific studies that have used fMRI to explore brain 
activities during the Trust Game have been published in several journals since 
approximately 2000, and this has led to an advance in understanding of behavioral trust 
from neuroscientific perspectives.
Neuroscience Studies of the Trust Game

Studies of the Trust Game using fMRI were initiated by McCabe, Houser, Ryan, Smith, and Trouard (2001). They compared individuals’ responses when they played the Trust Game with a human counterpart to when they played with a computer counterpart. They found that subjects showed a stronger activation of the prefrontal cortex when playing with a human counterpart. According to their argument, the prefrontal cortex contributes in binding a trustor’s attention with the trustee’s joined benefit. This leads to the trustor forgoing pursuing his or her own short-term self-interest. Later, King-Casas et al. (2005) investigated brain activity of both trustor and trustee in a repeated Trust Game featuring the same two participants throughout. In the study, the trustee’s caudate nucleus activities responded to the trustor’s benevolent reciprocity (i.e., awarding more investment relative to the trustee’s previous repayment). The authors reported strong correlations between the trustee’s caudate nucleus and the trustor’s middle cingulate cortex, and between caudate nucleus and anterior cingulate cortex within the trustee’s brain. The authors also found a time shift of the peaks of correlations, indicating a development of model of the counterpart by the trustee. They argued that the head of the caudate nucleus receives or computes information about (i) the fairness of a social partner’s decision and (ii) the intention to repay that decision with trust. Krueger et al. (2007) staged a Trust Game in which the same pair of subjects played repeatedly, but alternated their roles of trustor and trustee. They consequently discovered a correlation between behavioral trust and paracingulate cortex activities, which have been believed to function to infer others’ intentions in predicting their subsequent behavior. Further studies regarding the topic have been conducted (e.g., van den Bos, van Dijk, Westenberg, Rombouts, & Crone, 2009, 2011; see a recent review, Krueger & Meyer-Lindenberg, 2019), and an inclusive meta-analysis paper has recently been published (Bellucci, Chernyak, Goodyear, Eickhoff, & Krueger, 2017). This meta-analysis classified fMRI studies of the Trust Game into six categories, using criteria such as whether the game is a one-shot or repeated game, and whether brain activities are measured during the trust decision-making phase, the allocation by the trustee, or the feedback of the game results. The researchers consequently reported that, in one-shot games, activation in the right anterior insula was observed during the trust decision-making phase. This is noteworthy because the anterior insular is considered to be related to thermal stimulation (Craig, Chen, Bandy, & Reiman, 2000) and time-interval perception (Craig, 2009). In economic game studies the anterior insular has been determined to be related with emotional responses to unfairness (Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003; Tabibnia, Satpute, & Lieberman, 2008) as well as emotional responses to the threat of punishment (Spitzer, Fischbacher, Herrnberger, Grön, & Fehr, 2007). Although these findings suggest the possibility that the behavioral trust observed in the Trust Game reflects emotional responses, it is impossible to identify whether such an emotional response is based on reciprocity and the intentions of the counterpart, risk-aversion (Paulus, Rogalsky, Simmons, Feinstein, & Stein, 2003), or other foundations. One possible means of understanding these brain activities during the Trust Game is experimentally decomposing the trust behavior exhibited in the game.
Decomposing Trust

One popular approach for decomposing the trust behavior present in the Trust Game is used in betrayal aversion studies (Aimone & Houser, 2012; Bohnet, Greig, Herrmann, & Zeckhauser, 2008; Bohnet & Zeckhauser, 2004). Bohnet and Zeckhauser (2004) argued that the trust choice in the Trust Game could be decomposed into preference for risk, preference for increasing another’s income, and betrayal aversion that is defined as an aversion of risk specifically sourced by another person rather than nature (Bohnet et al., 2008). In order to illustrate how these components influence people’s decision, the authors compared a standard Trust Game with two other games with similar game and pay-off structures (i.e., Risky Dictator Game and a Decision Problem). As a reminder, in the Trust Game the trustor chooses between a trust choice and a sure choice, and the payment of the pair is going to be decided by another person, the trustee, if a trust choice has been chosen. In a Risky Dictator Game, the participant (trustor) also chooses between a risky choice and a sure choice to decide the outcome for himself/herself as well as for a counterpart. While the sure choice brings each participant a fixed payment, outcomes when both parties decide upon a risky choice are decided by a lottery with a certain probability. Therefore, the Risky Dictator Game is comparable to the Trust Game except that it involves natural risk rather than social risk. On the other hand, a Decision Problem involves no other people but a decision maker. Similar to the Risky Dictator Game, the participant in the Decision Problem also chooses between a risky choice (a lottery) and a sure choice. However, because it has no impact on other people, the Decision Problem involves only natural risk but not social preference. As a result, if the trust presented in the Trust Game is fully dependent on natural risk, there should be no difference in the level of trust or risky choice exhibited for the three games. If the trust behavior is based on a preference for natural risk or a preference for increasing another’s income, then the Decision Problem should show lower trust than the Risky Dictator Game and the Trust Game, while the latter two should show a similar level of trust. However, if an aversion to the risk of being betrayed by the counterpart is also involved, trust should be lower in the Trust Game than in the Decision Problem, and lower in the Decision Problem than in the Risky Dictator Game.

Laboratory experiments have shown that trust\(^1\) is lower in the Trust Game than in the Decision Problem, while the latter has a similar level of trust to the Risky Dictator Game (Bohnet & Zeckhauser, 2004). Bohnet et al. (2008) performed similar experiments across several countries, reporting that the betrayal-aversion component could not be identified in Brazil, China, and Switzerland, but that it successfully suppresses trust in Oman, Turkey, and the USA.

Previous studies of betrayal aversion (Bohnet et al., 2008; Bohnet & Zeckhauser, 2004) have indicated the possibility that components of behavioral trust can be identified by modifying the structure of the Trust Game. Through a comparison with the Risky Dictator Game, expectation of, or uncertainty regarding, a potential betrayal by a human counterpart has been suggested to induce lower trust in the Trust Game (Bohnet &

\(^1\) More precisely, the paper uses minimal acceptable probabilities (MAP), which increases when trust is suppressed.
Zeckhauser, 2004). An fMRI study further confirmed this hypothesis by revealing stronger anterior insula activity when participants played the Trust Game than when they played the Risky Dictator Game (Aimone, Houser, & Weber, 2014). This finding further suggests that the role that the anterior insula, a brain region that has been reported to be consistently involved in empathy, compassion, and interpersonal phenomena such as fairness and cooperation (Lamm & Singer, 2010), plays a role in the Trust Game (as suggested in the abovementioned meta-analysis; Bellucci et al., 2017) and it is probably related to betrayal aversion rather than natural risk aversion². These examples of betrayal aversion demonstrate that emotion plays a key role in behavioral trust in the Trust Game.

In addition to social risk preference and emotion (Fehr, 2009), a third factor drives behavioral trust, and that is the expectation of the cooperativeness of the counterpart. Although we have stated earlier that a prediction of cooperation from the counterpart is important for behavioral trust (Eckel & Wilson, 2004; Watabe et al., 1996), previous studies have mostly focused on the expectation of reciprocal cooperation in response to cooperation. Nonetheless, another aspect that should not be overlooked is belief in the general trustworthiness or altruism of others. In the Trust Game, trustors must decide whether to trust the trustee, an anonymous counterpart who is not under any outer influence such as a risk of punishment. This situation involves trusting others based on a belief in human nature; in other words, believing in the general trustworthiness and altruism of strangers. This is known as “general trust,” and its average level varies across societies, depending on various social structures (Yamagishi, 1998/2011). General trust measured using psychological scales, which is called “attitudinal trust” in the current study, has been reported to positively correlate with behavioral trust in the Trust Game (Yamagishi et al., 2015). This finding proposes the possibility that a prediction of another’s trustworthiness and altruism underlies the behavioral trust shown in the Trust Game.

Faith Game

The Faith Game (Kiyonari & Yamagishi, 1999) was developed to identify a component of trust based on inferences or beliefs regarding the altruistic propensity of others. In the standard version used in Kiyonari and Yamagishi (1999), featuring an anonymous lab experiment, participants are paired and asked to play the role of either an allocator or a recipient. The allocator receives an endowment of X from the experimenter, and then allocates it between him/herself and the recipient. The recipient, on the other hand, makes a decision, without knowing the result of the allocation, between receiving the amount allocated to them (trust choice) or receiving a fixed amount that is less than half of X (sure choice). It is noteworthy that the allocator is unaware that the counterpart has any choice, which means that the allocator is playing a Dictator Game (Forsythe, Horowitz, Savin, & Sefton, 1994; Kahneman, Knetsch, & Thaler, 1986) in which the allocator is fully capable of deciding his/her own payment, while the recipient is dependent on the allocator’s decision. Since the allocator is

² Lauharatanahirun, Christopoulos, and King-Casas (2012) also reported that the amygdala is activated during instances of social risk, but not during instances of non-social risk.
unaware that that recipient has a choice, no reciprocity should be expected on the part of the allocator, and also almost no fallacy of control should be present. On the other hand, the recipient’s choice has no effect on the allocator’s income; thus, the recipient’s likelihood of making a trust choice should not be affected by altruism toward the allocator. Thus, recipients who expect the randomly assigned allocator (a general other) to show altruism and make a fair decision will choose the trust choice; in contrast, if they believe that people in general are selfish, they will pick the sure choice. As a result, the recipient’s choice can be considered to be based on their level of behavioral trust, which reflects their belief of the level of altruism of people in general.

Studies of the Faith Game are not as common as those of the Trust Game. Kiyonari, Foddy, and Yamagishi (2007) conducted a cross-cultural comparison between Japanese and Australians using the Faith Game and Trust Game, using a minimal group paradigm (Tajfel, Billig, Bundy, & Flament, 1971) to investigate levels of trust towards members from the same group versus members of other groups. For both countries, trust in the Faith Game was higher towards in-group members than out-group members; however, no such differences were found in the Trust Game. Suzuki, Konno, and Yamagishi (2007) introduced a modified version of the Faith Game where participants chose to receive an allocation from either an in-group member or an out-group member. They consequently found that in-group members were only preferred when the receiver knew the allocator was aware the receiver was from the same group. Further, Yamagishi et al. (2013) examined a sample’s behaviors over a series of games, and reported a correlation of .38 between trust in the Trust Game and trust in the Faith Game, which were conducted within nine months of each other. Meanwhile, Oda and Nakajima (2010) explored the correlation between facial memory and trust in the Faith Game, finding that participants remembered the faces of allocators who had not shared money with them during a previous round, and consequently placed less trust in them. Finally, Komiya and Mifune (2015) reported that, for those who are classified as “pro-self” in social value orientation (SVO) measurements, trust in the Faith Game is significantly lower than their risk-taking in a gambling task, when the two tasks share the same pay-off matrix.

Although, theoretically speaking, trust in the Faith Game should reflect an expectation of altruism in general others, there is some inconsistency among previous findings. Kiyonari and Yamagishi (1999) examined whether cultural differences influence trust in the Faith Game, as it has been indicated by previous studies that Americans have higher trust in general others than do Japanese (Yamagishi & Yamagishi, 1994; see also Buchan, Croson, & Dawes, 2002). Surprisingly, however, no significant American-Japan difference was found in regard to trust in a randomly assigned counterpart, while American participants showed more trust than did Japanese in a specific counterpart (Kiyonari & Yamagishi, 1999). Kiyonari, Yamagishi, Cook, and Cheshire (2006) also compared the Faith Game and Trust Game between Japanese and American participants, reporting no statistical difference between the two countries in terms of behavioral trust. Meanwhile, Yamagishi et al. (2015) reported positive correlations between attitudinal trust and behavioral trust in both the Trust Game and Faith Game; however, when pro-sociality, measured using SVO measurements, and age
were controlled, trust in the Trust Game remained positively correlated with the attitudinal trust scale, while the correlation between the Faith Game and attitudinal trust disappeared. These findings highlight further issues to be solved, including whether trust in the Faith Game is produced by a different psychological mechanism than is evident in the Trust Game. If so, the mechanisms the underlie each game must be better understood.

Current Study

The current study aims to explore the psychological mechanisms behind the behavioral trust present in the Faith Game by investigating correlations between trust in the Faith Game and a battery of individual characteristics. This is performed using a dataset collected by Yamagishi et al. (2012, 2013) for the “Games and Culture Project,” where various experimental games and cultural psychological experiments were administered to the same set of participants over three-and-a-half years. Individual characteristics measured include attitudinal trust, pro-sociality (measured using SVO measurements), personality (using the Big Five personality traits), emotional intelligence, social caution, risk aversion, and betrayal aversion.

As is argued above, the behavioral trust shown in both the Trust and Faith Game may partly stem from an expectation or belief regarding the altruistic tendency of general others. Measurement of attitudinal trust can reveal the extent to which people believe that general others are trustworthy, which is expected to correlate with the levels of trust shown in both games.

Meanwhile, pro-sociality concerning SVO reflects personal differences in regard to weighting self-interest, another’s interest, and equality between one’s self and the other (van Lange, de Bruin, Otten, & Joreman, 1997); this is also known to reflect behavioral assimilation (van Lange, 1999), in which cooperation is performed in response to cooperation and defection is performed in response to defection. Such behavioral assimilation is also present in single-round games, through the expectation of a similar level of cooperativeness from others (Yamagishi, Terai, Kiyonari, Mifune, & Kanazawa, 2007). In the Trust Game, trust increases the allocator’s income, and trust is also produced by the expectation that a counterpart should behave fairly in response to receiving trust. Meanwhile, in the Faith Game, an expectation of altruistic allocation also increases the recipient’s trust. Thus, trust in both games is expected to be positively correlated with pro-sociality. These hypotheses have been supported by Yamagishi et al. (2015), and we aim to replicate the results.

Studies involving the Big Five personality traits and attitudinal trust have reported correlations between all five traits and trust (Dinesen, Nørgaard, & Klemmensen, 2014; Hiraishi, Yamagata, Shikishima, & Ando, 2008). In particular, while positive correlations of extraversion, openness, and agreeableness with trust have been consistently observed, both positive (Hiraishi et al., 2008) and negative (Dinesen et al., 2014) correlations have been found for conscientiousness. In contrast, other studies have reported that only agreeableness correlates with trust (Mondak & Halperin, 2008), or that agreeableness is the core Big Five personality trait correlated with trust (Dinesen & Bekkers, 2017). Yamagishi et al. (2010), using data from the same dataset as the current
paper, also claimed that attitudinal trust only correlates with agreeableness. In the current paper, we explore if this pattern can be reproduced in terms of the behavioral trust present in the Trust Game and Faith Game.

In the emancipation theory of trust proposed by Yamagishi (1998/2011), general trust tendency is hypothesized to co-evolve with the ability to detect trustworthiness, because there is no other external guarantee that others will not betray the perceiver. Further, detection of another’s trustworthiness is considered to reflect emotional intelligence, especially in interpersonal abilities (Yamagishi et al., 2010). Yamagishi et al. (2010) confirmed this through identifying a positive correlation with attitudinal trust.

Social caution refers to a tendency to be cautious and to worry about betrayal. People with low attitudinal trust believe that others will betray their trust and have a high level of social caution. In contrast, people with high attitudinal trust believe in the importance of distinguishing trustworthy people from untrustworthy ones (Kikuchi, Watanabe, & Yamagishi, 1997), and highly cautious people and people with low levels of caution can co-exist among high trustors. As a result, a weak correlation is hypothesized to exist between attitudinal trust and social caution (Cheshire, Antin, Cook, & Churchill, 2010). The present study explores whether such a relationship can also be found in regard to behavioral trust.

In the Trust Game, trust is argued to involve risk preference, because choosing to trust can result in a benefit or disadvantage to one’s self-interest (e.g., Ben-Ner & Putterman, 2001; Bohnet & Zeckhauser, 2004). The same argument also applies to the Faith Game, which means that a correlation between behavioral trust in both games and risk preference is likely. Betrayal aversion has also been suggested to be a factor restraining behavioral trust in the Trust Game (Bohnert et al., 2008; Bohnet & Zeckhauser, 2004). Consequently, we predict that there are negative correlations not only between betrayal aversion and trust in the Trust Game, but also between betrayal aversion and trust in the Faith Game; this is because both games are decision-making tasks involving social risk. Although a previous study has considered the difference between the Risky Dictator Game and Trust Game as an indicator of individuals’ betrayal aversion (Aimone, Ball, & King-Casas, 2015), the present study uses an independently created measurement to capture personal differences regarding betrayal aversion.

**Method**

The present paper involves an analysis of data from the Games and Culture Project, which is a project comprising eight waves of studies of non-student samples in Japan, including data regarding various behavioral-based economic games and individual characteristics. Details of the project can also be found in Yamagishi et al. (2012, 2013). The time frame for each study wave was as follows: Wave 1 was conducted between February and March 2008, Wave 2 was conducted between April and May 2008, Wave 3 was conducted in November 2008, Wave 4 was conducted between February and March 2009, Wave 5 was conducted in November 2009, Wave 6 was conducted between July and August 2010, Wave 7 was conducted between February and March 2011, and Wave 8 was conducted between October and November 2011. Long intervals were inserted between the waves in order to minimize potential carry-over effects.
Initially, a set of 108 participants, with an almost equal distribution of gender and evenly spread age range of between 20 and 60 years, were recruited via newspaper advertisement, and this group participated in waves 1 to 6. An additional set of 105 participants with similar criteria was then recruited for Wave 7, in which the findings of key measurements for the previous waves were compared with those of the new sample set. Both sets of samples were included in Wave 8. The research was approved by the ethics committee of the Center for Experimental Studies in Social Sciences, Hokkaido University. All participants read and signed the consent form each time they participated in the study. For all experiments and tasks, anonymity was strictly assured. In the presented paper, we report results from sample set 1, because neither the Trust Game nor the Faith Game was performed using sample set 2. Participants were offered a base participation fee of between 4,000 and 10,000 Yen for half-day to whole-day sessions. In addition, they also received payments in accordance with their performance in game experiments and other incentivized tasks. Hereafter, we describe details relevant to the present paper.

**Trust Game**

The Trust Game was conducted in Wave 5 of the project. In the game, participants were randomly paired, with one playing as the entruster (trustor) and the other playing as the allocator (trustee). The entruster was presented with a sum of 600 Yen (approximately 6 USD) and could then decide how much they would like to entrust. Any amount of money entrusted would be tripled and transferred to the allocator, who could then freely allocate the money between him/herself and the entruster. Meanwhile, allocators decided how much they would allocate to entrusters, ranging from 300 Yen to 1800 Yen (in units of 300 Yen), depending on the amount entrusted (100, 200, ... 600 Yen). After confirming that all participants understood the game rules, each participant played the game twice, once as the entruster and once as the allocator. Each participant was paired with a different counterpart when their role was switched. All the participants had therefore been assigned the role of entruster and that of allocator once, meaning that the Trust game itself was played twice. After all decisions had been made, entruster decisions were paired with allocator decisions, and participants received the corresponding payment. In the present paper, we focus on the decisions of the entrusters; that is, the amount of the endowment that the entrusters decided to entrust to the allocators.

**Faith Game**

The Faith Game was conducted in Wave 6. Participants received a set of instructions reminding them of a Dictator Game they had played in Wave 3, in which an allocator decided how to allocate money between him/herself and a recipient. After the refresher of how the Dictator Game works, participants were given examples of allocations; for example, an allocator allocating all of the endowment to the counterpart, an allocator allocating an equal amount between themselves and the counterpart, and an allocator allocating more to themselves than to their counterpart. Next, they were informed that one of the allocations from the Dictator Game from Wave 3 would be paired with each participant. Participants were each given an endowment of 600 Yen and were asked to decide how much of the endowment they would like to invest to the allocation, with uninvested money paid directly to the participant. Any amount invested would be tripled, and a proportion of the tripled amount would be returned. The proportion of the return would be decided by the allocation from the Dictator Game that was paired with the participant. For example, if the allocator had allocated an equal amount between him/herself and the counterpart, the participant in the Faith Game would receive half of the tripled investment as a return. Thus, if the participant invested all 600 Yen of the endowment, the return would be 50% of 600 * 3, which is 900. However, participants were also told that their counterpart would not receive any payment in the Faith Game, because they had already been paid in the Dictator Game. In the presented paper, we focus on the participants’ decisions regarding investing the 600 Yen endowment.

**Individual Characteristics**

The following measurements of personal characteristics and beliefs were also involved in the analysis of the presented paper.

*Attitudinal trust and social caution.* Attitudual trust and social caution (Yamagishi & Yamagishi, 1994) were measured in waves 1, 3, 4, 5, 6, 7, and 8. We averaged the scores from each wave to construct indicators of attitudinal trust and social caution. Cronbach’s alpha coefficients among the average points of
the 6 waves were .90 and .89 for attitudinal trust and social caution respectively.

*Pro-sociality.* Pro-social orientation was measured using triple-dominance measurement (van Lange et al., 1997) in waves 1, 4, and 7, and using ring measurement (Liebrand & McClintock, 1988) in waves 2 and 8. Both measurements consisted of a few sets of options, each of which illustrated hypothetical payment allocations between the participant and another person. In the triple dominance measurement, participants chose from three options such as, “a) you get 500, the other gets 100; b) you get 500, the other gets 500; c) you get 550, the other gets 300,” while in the ring measure participants chose from two options such as, “a) you get 1450, the other gets 390; b) you get 1500, the other gets 0.” For each measurement, decisions throughout all sets of allocation options were aggregated, and those who showed a stronger preference for increasing others’ payments rather than merely pursuing self-interest were considered to be more pro-social. We aggregated the overall pro-sociality score by averaging the pro-sociality sores for each measurement, which were the mean number of times a person was classified as pro-social in each measurement for each wave. Cronbach’s alpha coefficient among all SVO measurements was .67.

*Big Five personality inventory.* We used the NEO Five-Factor Inventory (NEO-FFI; a short version of the Revised NEO Personality Inventory; Costa & McCrae, 1992) in waves 1 and 7. This was used to determine each participant’s Big Five personality trait.

*Emotional intelligence.* Emotional intelligence, or the emotional quotient (EQ), was measured in waves 1 and 7 using a 65-item scale created by Uchiyama, Shimai, Utsuki, and Otake (2001), in which EQ was construed using three subordinate facets: intrapersonal, interpersonal, and situational. We were particularly interested in the interpersonal facet, which involves sub-factors of empathy, altruism, and interpersonal relationships. The interpersonal facet of EQ indicates that the capability of forming and maintaining interpersonal relationships involves emotional intelligences such as empathies of other’s emotional status, helping those who are in need, and adjusting relationships between self and others as well as among others.

*Betrayal aversion.* Betrayal aversion was measured in waves 5, 6, 7, and 8. The measurements used in waves 5, 6, and 7 were consistent with the Betrayal Aversion game 1 reported by Yamagishi et al. (2015). This game was similar to the Trust Game, with participants being asked to choose between entrusting 1,000 Yen to an anonymous allocator or keeping the 1,000 Yen as a guaranteed amount. The transferred money would be quadrupled instead of tripled, which was designed to simplify the calculations in the instruction. The game was played as a vignette study, with no actual payment of money. One notable point is that participants were informed of the allocator’s history of allocation over the previous 10 rounds played with other participants, which provided a proxy of the probability of the participants receiving a fair allocation from the allocator. Further, participants were asked to imagine playing the game with seven different allocators, in which the probability of a fair return varied between 20% to 80%, in increments of 10%. Logically, the belief aspect of general trust should play no role in the trustor’s choice, because the allocator’s probability of a fair return is known and does not need to be inferred.

The measurement performed in Wave 8 was a similar vignette study, in which participants played a Trust Game. Here, participants were informed that there were 10 Allocators in another room, all of whom had already claimed a fair or a selfish allocation. Participants were informed that one would be chosen as the participant’s counterpart. Participants were informed of the proportion of fair allocators amongst the 10 people, and then decided whether to entrust a randomly chosen counterpart or to retain a guaranteed amount. Like the other version of the Betrayal Aversion game, participants made seven decisions in response to varying probabilities of receiving a fair allocator (between 20% and 80%). Betrayal aversion was indicated by the number of sure choices chosen among the seven decisions. We averaged betrayal aversion scores from each wave to create an indicator of overall betrayal aversion. Cronbach’s alpha coefficient among the three waves was .86.

*Risk aversion.* Risk aversion was measured in waves 5 and 8. This measurement of risk aversion was very similar to the first version of betrayal aversion introduced above, except that the participants chose whether to spend 1,000 Yen playing in a lottery instead of in a Trust Game. Winning the lottery would bring a return of 2,000 Yen, which was equal to the return of a fair allocation in the betrayal aversion measurement. Participants made decisions for seven lottery machines, after being informed of the probability of winning for each machine, which varied between 20% and 80%. We also averaged scores from both waves to create an indicator of overall risk aversion. Cronbach’s alpha coefficient among the two waves was .40.
RESULTS

Table 1 shows correlations among individual attributes. The correlations between attitudinal trust and pro-sociality, Big Five personality trait, and EQ differed slightly from those reported by Yamagishi et al. (2010), which used the same dataset. This is likely because the dataset has been updated to include new waves since the publication of the previous paper. It is noteworthy that attitudinal trust showed a relatively strong negative correlation with social caution, risk aversion showed no correlation other than that with betrayal aversion, and betrayal aversion negatively correlated with pro-sociality.

A high correlation of behavioral trust was identified between the Trust game and the Faith game ($r = .59, p < .01$). Table 2 shows the correlations between behavioral trust in the Trust Game and the Faith Game and individual variables. Betrayal and risk aversions had negative correlations, and pro-sociality had positive correlations with behavioral trust in both games, while positive correlations between behavioral trust and attitudinal trust and interpersonal emotional intelligence were only identified in the Trust Game.

Tables 3 and 4 show partial correlations controlling age and controlling both age and pro-sociality, respectively, which conforms with the analyses of Yamagishi et al. (2015). Regarding behavioral trust in both the Trust and Faith Games, negative correlations were found for both risk aversion and betrayal aversion.

Because correlations with behavioral trust were found in both risk aversion and betrayal aversion, we performed partial correlational analyses to examine the relative strength of the two. When risk aversion was controlled, betrayal aversion was negatively correlated with trust in both games. The partial correlation with trust in the Trust Game was $-0.33 (p < .01)$, while that with trust in the Faith Game was $-0.43 (p < .01)$. On the other hand, when betrayal aversion was controlled, trust in neither the Trust Game ($r = -.15$) or the Faith Game ($r = .01$) showed significant partial correlation with risk aversion.

DISCUSSION

The present paper aims to identify the psychological mechanisms that underlie behavioral trust in economic games by exploring correlations between behavioral trust and individual attributes. To our surprise, the predicted correlations did not appear in many of the individual attributes. The self-reported attitudinal trust, positively correlated with behavioral trust in the Trust Game only, but not in the Faith Game. The scale used to determine attitudinal trust, featuring items such as “most people are trustworthy,” measures the extent to which responders believe that people in general are trustworthy, cooperative, and/or altruistic. Meanwhile, the observed correlation between attitudinal trust and pro-sociality supports the validity of the measurement of attitudinal trust, because pro-sociality measured by SVO has been determined to partly reflect expectations of others’ trustworthiness or cooperativeness (Kelley & Stahelski, 1970; van Lange, 1992). Although, from a theoretical point of view, such an indicator should be
Table 1. Pearson correlation among individual variables

|                        | General Trust | Social Caution | Pro-sociality | B5_Extraversion | B5_Agreeableness | B5_Conscientiousness | B5_Neuroticism | B5_Openness | EQ_Intraperson | EQ_Interperson | EQ_Situation | EQ_overall | Risk Aversion |
|------------------------|---------------|----------------|---------------|-----------------|------------------|----------------------|----------------|-------------|----------------|----------------|--------------|------------|---------------|
| Social Caution         | -0.68**       |                |               |                 |                  |                      |                |             |                |                |              |            |               |
| Pro-sociality          | 0.31**        | -0.25*         |               |                 |                  |                      |                |             |                |                |              |            |               |
| B5_Extraversion        | 0.14          | -0.24*         | 0.16          |                 |                  |                      |                |             |                |                |              |            |               |
| B5_Agreeableness       | 0.51**        | -0.36**        | 0.31**        | 0.37**          |                  |                      |                |             |                |                |              |            |               |
| B5_Conscientiousness   | 0.11          | -0.07          | 0.10          | 0.23*           | 0.33**           |                      |                |             |                |                |              |            |               |
| B5_Neuroticism         | 0.08          | -0.10          | 0.06          | 0.23*           | 0.18             | 0.10                 |                |             |                |                |              |            |               |
| B5_Openness            | 0.12          | 0.01           | 0.10          | 0.29**          | 0.26**           | 0.48**               | 0.21*          |             |                |                |              |            |               |
| EQ_Intraperson         | 0.25*         | -0.03          | 0.17          | 0.21*           | 0.51**           | 0.45**               | 0.08           | 0.55**      |                |                |              |            |               |
| EQ_Interperson         | 0.36**        | -0.18          | 0.29**        | 0.34**          | 0.67**           | 0.27**               | 0.11           | 0.43**      | 0.75**        |                |              |            |               |
| EQ_Situation           | 0.25*         | -0.09          | 0.16          | 0.38**          | 0.40**           | 0.32**               | 0.22*          | 0.64**      | 0.80**        | 0.75**        |              |            |               |
| EQ_overall             | 0.31**        | -0.11          | 0.23*         | 0.35**          | 0.58**           | 0.37**               | 0.16           | 0.59**      | 0.91**        | 0.91**        | 0.93**       |            |               |
| Risk Aversion          | 0.10          | -0.02          | 0.02          | -0.09           | -0.06            | -0.01                | -0.17          | -0.03       | -0.11         | -0.12         | -0.23*       | -0.17      |               |
| Betrayal Aversion      | -0.20         | 0.11           | -0.32**       | -0.09           | -0.14            | 0.07                 | -0.08          | 0.13        | 0.01          | -0.13         | 0.00         | -0.05      | 0.50**        |

Note 1. *p < .05; **p < .01, unadjusted p values.
Note 2. B5: Big-five personality inventory; EQ: emotional intelligence.
Table 2. Pearson correlation of behavioral trust in Trust Game and in Faith Game with individual variables

| Behavioral Trust | General Trust | Social Caution | Pro-sociality | B5_Extraversian | B5_Agreeableness | B5_Conscientiousness | B5_Neuroticism | B5_Openness | EQ_Intraperson | EQ_Interperson | EQ_Situation | EQ_Overall | Risk Aversion | Betrayal Aversion |
|------------------|---------------|----------------|---------------|-----------------|-----------------|---------------------|----------------|-------------|----------------|----------------|--------------|-------------|---------------|------------------|
| Trust Game       | 0.23*         | -0.05          | 0.26*         | 0.08            | 0.20            | -0.03               | 0.16           | 0.06        | 0.08           | 0.24*          | 0.07         | 0.15        | -0.35**       | -0.45**         |
| Faith Game       | 0.14          | -0.04          | 0.26*         | 0.13            | 0.08            | -0.09               | 0.14           | -0.16       | -0.03          | 0.05           | -0.09        | -0.03       | -0.25*        | -0.49**         |

*Note 1. *p < .05; **p < .01, unadjusted p values.
*Note 2. B5: Big-five personality inventory; EQ: emotional intelligence.

Table 3. Partial correlation with behavioral trust in Trust Game and in Faith Game controlling age

| Behavioral Trust | General Trust | Social Caution | Pro-sociality | B5_Extraversian | B5_Agreeableness | B5_Conscientiousness | B5_Neuroticism | B5_Openness | EQ_Intraperson | EQ_Interperson | EQ_Situation | EQ_Overall | Risk Aversion | Betrayal Aversion |
|------------------|---------------|----------------|---------------|-----------------|-----------------|---------------------|----------------|-------------|----------------|----------------|--------------|-------------|---------------|------------------|
| Trust Game       | 0.17          | 0.09           | 0.18          | 0.07            | 0.11            | -0.05               | 0.13           | 0.03        | 0.01           | 0.13           | -0.01        | 0.05        | -0.34**       | -0.41**         |
| Faith Game       | 0.12          | 0.02           | 0.22*         | 0.08            | 0.09            | -0.06               | 0.16           | -0.17       | -0.01          | 0.07           | -0.10        | -0.02       | -0.25*        | -0.47**         |

*Note 1. *p < .05; **p < .01, unadjusted p values.
*Note 2. B5: Big-five personality inventory; EQ: emotional intelligence.

Table 4. Partial correlation with behavioral trust in Trust Game and in Faith Game controlling age and pro-sociality

| Behavioral Trust | General Trust | Social Caution | B5_Extraversian | B5_Agreeableness | B5_Conscientiousness | B5_Neuroticism | B5_Openness | EQ_Intraperson | EQ_Interperson | EQ_Situation | EQ_Overall | Risk Aversion | Betrayal Aversion |
|------------------|---------------|----------------|-----------------|-----------------|---------------------|----------------|-------------|----------------|----------------|--------------|-------------|---------------|------------------|
| Trust Game       | 0.14          | 0.09           | 0.05            | 0.08            | -0.07               | 0.13           | 0.01        | 0.00           | 0.10           | -0.02        | 0.03        | -0.36**       | -0.39**         |
| Faith Game       | 0.09          | 0.02           | 0.05            | 0.05            | -0.09               | 0.15           | -0.20       | -0.02          | 0.02           | -0.12        | -0.05       | -0.27*        | -0.45**         |

*Note 1. *p < .05; **p < .01, unadjusted p values.
*Note 2. B5: Big-five personality inventory; EQ: emotional intelligence.
equally, if not more strongly, correlated with behavioral trust in the Faith Game, a previous empirical study could only find a weak correlation (Yamagishi et al., 2015). Considering evidences from previous and current studies, the possibility arises that behavioral trust in the Faith Game is based on aspects other than “expectations of general others’ trustworthiness.” On the other hand, the correlation between attitudinal trust and behavioral trust in the Trust Game was not strong, either. A lack of correlation between attitudinal trust and behavioral trust in the Trust Game has also been reported in previous work (Watabe, Kato, Monji, Horikawa, & Kanba, 2012). Additionally, Yamagishi et al. (2015) found a weak partial correlation when age and pro-sociality were controlled. The current study also reports no correlation between behavioral trust and a large number of subjectively-reported individual attributes. The above results may indicate the presence of a large inconsistency between behavioral trust and subjective belief and attitudes of trust.

Both risk and betrayal aversion were negatively correlated with behavioral trust in both the Trust and Faith games. Our results are consistent with another study that suggested that risk preference is part of behavioral trust (Bohnet & Zeckhauser, 2004), although its Cronbach’s alpha coefficient (.40) casts doubt on its reliability. However, it is noteworthy that that such correlation might be caused by risk avoidance in social domains, because no partial correlation was found between behavioral trust and risk aversion when betrayal aversion was controlled. Controlling for risk aversion did not have a significantly large impact on correlations between betrayal aversion and behavioral trust. These results indicate a possibility that betrayal aversion is not only a reaction to a contravening of the norm of reciprocity when one’s altruistic behavior is betrayed, but also a reaction towards the selfishness or hostility experienced when one receives mistreatment. According to Aimone et al. (2014), such reaction may reflect negative emotional responses related to insula activities. However, it is beyond the current paper’s scope to discuss whether and what kind of emotion is lying behind betrayal aversion. Future investigations are awaited regarding this topic.

Unlike previous studies using behavioral differences among multiple games as an indicator of betrayal aversion (e.g., Bohnet & Zeckhauser, 2004), the measurements for risk aversion and betrayal aversion that are introduced by our study allow us to directly examine the correlational relationship among behavioral trust, risk and betrayal aversions, and other individual differences. On the one hand, we should be aware that the similar format of aversion measurements and games might have contributed to the correlations among them and hope for future comprehensions. On the other hand, it should be noted that attitudinal trust and betrayal aversion are both correlated with behavioral trust, while no correlations were found between the two. These results indicate that betrayal aversion in this study captures one of the components of behavioral trust. Meanwhile, the finding that attitudinal trust was only correlated with behavioral trust in the Trust Game but not in the Faith Game, along with the above results, indicate that an emotional reaction to selfish people may be differentiated from evaluation of the trustworthiness of others in general. In other words, they are not two sides of the same coin.
In summary, the current study revealed that behavioral trust in the Faith Game might be based on not only evaluations of general others, but also emotional mechanisms, that is particularly suggested by the relationship between trust and betrayal aversion; however, the difference between trust in the Trust Game and that in the Faith Game remains ambiguous. It should not be overlooked that the current work is an explorative analysis using a pre-existing dataset, for which no pre-planning of sample size could be performed, and thus the correlations or non-correlations that this paper reports may have been produced by an inappropriate sample size. Further studies are needed to test the generalizability of our findings in other samples. Also, as stated above, the validity of our findings, especially the strong correlations between aversions and behavioral trust, need further examination by measurements that are less similar to game experiment situation. Finally, although theoretically speaking, behavioral trust in the Faith Game is considered to reflect expectations of the trustworthiness and altruism of general others, further exploration should examine its determinative factors, which will advance understanding of the psychological mechanisms of behavioral trust.

REFERENCES

Aimone, J., Ball, S., & King-Casas, B. (2015). The betrayal aversion elicitation task: An individual level betrayal aversion measure. *PLoS ONE, 10*(9), e0137491.

Aimone, J. A., & Houser, D. (2012). What you don’t know won’t hurt you: A laboratory analysis of betrayal aversion. *Experimental Economics, 15*(4), 571–588.

Aimone, J. A., Houser, D., & Weber, B. (2014). Neural signatures of betrayal aversion: An fMRI study of trust. *Proceedings of the Royal Society B: Biological Sciences, 281*(1782), 20132127.

Ashraf, N., Bohnet, I., & Piankov, N. (2006). Decomposing trust and trustworthiness. *Experimental Economics, 9*(3), 193–208.

Bellucci, G., Chernyak, S. V., Goodyear, K., Eickhoff, S. B., & Krueger, F. (2017). Neural signatures of trust in reciprocity: A coordinate-based meta-analysis. *Human Brain Mapping, 38*(3), 1233–1248.

Ben-Ner, A., & Putnam, L. (2001). Trusting and trustworthiness. *Boston University Law Review, 81*, 523–551.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior, 10*(1), 122–142.

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*(1), 294–310.

Bohnet, I., & Zeckhauser, R. (2004). Trust, risk and betrayal. *Journal of Economic Behavior & Organization, 55*(4), 467–484.

Buchan, N. R., Croson, R. T., & 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*(1), 168–206.

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

Cheshire, C., Antin, J., Cook, K. S., & Churchill, E. (2010). General and familiar trust in websites. *Knowledge, Technology & Policy, 23*(3–4), 311–331.

Costa, P. T., Jr., & McCrae, R. R. (1992). *NEO PI-R professional manual*. Odessa, FL: Psychological Assessment Resources.

Craig, A. D. (2009). Emotional moments across time: A possible neural basis for time perception in the anterior insula. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1525),
1933–1942.
Craig, A. D., Chen, K., Bandy, D., & Reiman, E. M. (2000). Thermosensory activation of insular cortex. *Nature Neuroscience*, 3(2), 184–190.
Dinesen, P. T., & Bekkers, R. (2017). The foundations of individuals’ generalized social trust. In P. A. M. van Lange, B. Rockenbach, & T. Yamagishi (Eds.), *Trust in social dilemmas* (pp. 77–99). New York, NJ: Oxford University Press.
Dinesen, P. T., Nørgaard, A. S., & Klemmensen, R. (2014). The civic personality: Personality and democratic citizenship. *Political Studies*, 62, 134–152.
Eckel, C. C., & Wilson, R. K. (2004). Is trust a risky decision? *Journal of Economic Behavior & Organization*, 55(4), 447–465.
Evans, A. M., & Krueger, J. I. (2011). Elements of trust: Risk and perspective-taking. *Journal of Experimental Social Psychology*, 47(1), 171–177.
Fehr, E. (2009). On the economics and biology of trust. *Journal of the European Economic Association*, 7(2-3), 235–266.
Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior*, 6(3), 347–369.
Hardin, R. (1992). The street-level epistemology of trust. *Analyse & Kritik*, 14, 152–176.
Hiraishi, K., Yamagata, S., Shikishima, C., & Ando, J. (2008). Maintenance of genetic variation in personality through control of mental mechanisms: A test of trust, extraversion, and agreeableness. *Evolution and Human Behavior*, 29(2), 79–85.
Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1986). Fairness and the assumptions of economics. *Journal of Business*, 59, S285–S300.
Kanagaretnam, K., Mestelman, S., Nainar, K., & Shehata, M. (2009). The impact of social value orientation and risk attitudes on trust and reciprocity. *Journal of Economic Psychology*, 30(3), 368–380.
Kelley, H. H., & Stahelski, A. J. (1970). Social interaction basis of cooperators’ and competitors’ beliefs about others. *Journal of Personality and Social Psychology*, 16(1), 66–91.
Kikuchi, M., Watanabe, Y., & Yamagishi, T. (1997). Judgment accuracy of other’s trustworthiness and general trust: An experimental study. *Japanese Journal of Experimental Social Psychology*, 37(1), 23–36.
King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., & Montague, P. R. (2005). Getting to know you: Reputation and trust in a two-person economic exchange. *Science*, 308(5718), 78–83.
Kiyonari, T., Foddy, M., & Yamagishi, T. (2007). Effects of direct and indirect exchange on trust of ingroup members. *Japanese Journal of Psychology*, 77(6), 519–527.
Kiyonari, T., & Yamagishi, T. (1999). A comparative study of trust and trustworthiness using the game of enthronement. *Japanese Journal of Social Psychology*, 15(2), 100–109.
Kiyonari, T., Yamagishi, T., Cook, K. S., & Cheshire, C. (2006). Does trust beget trustworthiness? Trust and trustworthiness in two games and two cultures: A research note. *Social Psychology Quarterly*, 69(3), 270–283.
Komiya, A., & Mifune, N. (2015). An individual difference in betrayal aversion: Prosociality predicts more risky choices in social but not natural domains. *Letters on Evolutionary Behavioral Sciences*, 6(1), 5–8.
Krueger, F., McCabe, K., Moll, J., Kriegeskorte, N., Zahn, R., Stenziok, M., . . . Grafman, J. (2007). Neural correlates of trust. *Proceedings of the National Academy of Sciences*, 104(50), 20084–20089.
Krueger, F., & Meyer-Lindenberg, A. (2019). Toward a model of interpersonal trust drawn from neuroscience, psychology, and economics. *Trends in Neurosciences*, 42(2), 92–101.
Lamm, C., & Singer, T. (2010). The role of anterior insular cortex in social emotions. *Brain Structure and Function*, 214(5-6), 579–591.
Lauharatanahirun, N., Christopoulos, G. I., & King-Casas, B. (2012). Neural computations underlying social risk sensitivity. *Frontiers in Human Neuroscience*, 6, 213.
Liebrand, W. B. G., & McClintock, C. G. (1988). The ring measure of social values: A computerized procedure for assessing individual differences in information processing and social value orientation. *European Journal of Personality*, 2(3), 217–230.
Luhmann, N. (1973). *Vertrauen: Ein Mechanismus der Reduktion sozialer Komplexität* (2 aufl.). Stuttgart, Germany: Ferdinand Enke Verlag.
Malhotra, D. (2004). Trust and reciprocity decisions: The differing perspectives of trustors and trusted parties. *Organizational Behavior and Human Decision Processes, 94*(2), 61–73.
McCabe, K., Houser, D., Ryan, L., Smith, V., & Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences, 98* (20), 11832–11835.
McCabe, K. A., Rigdon, M. L., & Smith, V. L. (2003). Positive reciprocity and intentions in trust games. *Journal of Economic Behavior & Organization, 52*(2), 267–275.
Mondak, J. J., & Halperin, K. D. (2008). A framework for the study of personality and political behaviour. *British Journal of Political Science, 38*(2), 335–362.
Oda, R., & Nakajima, S. (2010). Biased face recognition in the faith game. *Evolution and Human Behavior, 31*(2), 118–122.
Paulus, M. P., Rogalsky, C., Simmons, A., Feinstein, J. S., & Stein, M. B. (2003). Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism. *NeuroImage, 19*(4), 1439–1448.
Pillutla, M. M., Malhotra, D., & Murnighan, J. K. (2003). Attributes of trust and the calculus of reciprocity. *Journal of Experimental Social Psychology, 39*(5), 448–455.
Putnam, R. D. (1993). *Making democracy work: Civic traditions in modern Italy.* Princeton, NJ: Princeton University Press.
Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science, 300*(5626), 1755–1758.
Schechter, L. (2007). Traditional trust measurement and the risk confound: An experiment in rural Paraguay. *Journal of Economic Behavior & Organization, 62*(2), 272–292.
Simmel, G. (1908). *Soziologie: Untersuchungen über die formen der vergesellschaftung.* Berlin, Germany: Duncker & Humblot.
Snijders, C., & Keren, G., (1998). Determinants of trust. In D. V. Budescu, I. Erev, & R. Zwick (Eds.), *Games and human behavior: Essays in honor of Amnon Rapoport* (pp. 355–385). Mahwah, NJ: Lawrence Erlbaum.
Spitzer, M., Fischbacher, U., Herrnberger, B., Grön, G., & Fehr, E. (2007). The neural signature of social norm compliance. *Neuron, 56*(1), 185–196.
Suzuki, N., Konno, Y., & Yamagishi, T. (2007). In-group bias in trusting behavior: A choice of allocator experiment with minimal groups. *Japanese Journal of Psychology, 78*(1), 17–24.
Tabibnia, G., Satpute, A. B., & Lieberman, M. D. (2008). The sunny side of fairness: Preference for fairness activates reward circuitry (and disregarding unfairness activates self-control circuitry). *Psychological Science, 19*(4), 339–347.
Tajfel, H., Billig, M., Bundy, R., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology, 1*(2), 149–178.
Uchiyama, K., Shimai, S., Utsuki, N., & Otake, K. (2001). EQS manyuaru [EQS manual]. Tokyo, Japan: Jitsumukyoiku-Shuppan.
van den Bos, W., van Dijk, E., Westenber, M., Rombouts, S. A., & Crone, E. A. (2009). What motivates repayment? Neural correlates of reciprocity in the trust game. *Social Cognitive and Affective Neuroscience, 4*(3), 294–304.
van den Bos, W., van Dijk, E., Westenber, M., Rombouts, S. A., & Crone, E. A. (2011). Changing brains, changing perspectives: The neurocognitive development of reciprocity. *Psychological Science, 22*(1), 60–70.
van Lange, P. A. (1992). Confidence in expectations: A test of the triangle hypothesis. *European Journal of Personality, 6*(5), 371–379.
van Lange, P. A. (1999). The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation. *Journal of Personality and Social Psychology, 77*(2), 337–349.
van Lange, P. A., de Bruin, E., Otten, W., & Joireman, J. A. (1997). Development of prosocial, individualistic, and competitive orientations: Theory and preliminary evidence. *Journal of Personality and Social Psychology, 73*(4), 733–746.
Watabe, M., Kato, T. A., Monji, A., Horikawa, H., & Kanba, S. (2012). Does minocycline, an antibiotic with inhibitory effects on microglial activation, sharpen a sense of trust in social interaction? *Psychopharmacology, 220*, 551–557.
Watabe, M., Terai, S., Hayashi, N., & Yamagishi, T. (1996). Cooperation in the one-shot prisoner’s dilemma based on expectations of reciprocity. *Japanese Journal of Experimental Social Psychology, 36*(2), 183–196.

Yamagishi, T. (2011). *Trust: The evolutionary game of mind and society.* New York, NY: Springer Science & Business Media. (Original work published 1998)

Yamagishi, T., Akutsu, S., Cho, K., Inoue, Y., Li, Y., & Matsumoto, Y. (2015). Two-component model of general trust: Predicting behavioral trust from attitudinal trust. *Social Cognition, 33*(5), 436–458.

Yamagishi, T., Horita, Y., Mifune, N., Hashimoto, H., Li, Y., Shinada, M., . . . Simunovic, D. (2012). Rejection of unfair offers in the ultimatum game is no evidence of strong reciprocity. *Proceedings of the National Academy of Sciences, 109*(50), 20364–20368.

Yamagishi, T., Mifune, N., Li, Y., Shinada, M., Hashimoto, H., Horita, Y., . . . Simunovic, D. (2013). Is behavioral pro-sociality game-specific? Pro-social preference and expectations of pro-sociality. *Organizational Behavior and Human Decision Processes, 120*(2), 260–271.

Yamagishi, T., Tanida, S., Takahashi, C., Shinada, M., Mifune, N., Inukai, K., . . . Hashimoto, H. (2010). Ippanteki shinrai to joudou chinou [General trust and emotional intelligence] (Working Paper No. 108). Hokkaido University, Center for Experimental Research in Social Sciences.

Yamagishi, T., Terai, S., Kiyonari, T., Mifune, N., & Kanazawa, S. (2007). The social exchange heuristic: Managing errors in social exchange. *Rationality and Society, 19*(3), 259–291.

Yamagishi, T., & Yamagishi, M. (1994). Trust and commitment in the United States and Japan. *Motivation and Emotion, 18*(2), 129–166.

Zuckerman, M. (1994). *Behavioral expressions and biosocial bases of sensation seeking.* New York, NY: Cambridge University Press.

*(Manuscript received 2 February, 2019; Revision accepted 8 March, 2019; Released online in J-STAGE as advance publication 25 December, 2019)*