# Behavioral characteristics in general trust: an exploratory laboratory-based analysis using the ultimatum game

**Daisuke Hamada (taisa1415@gmail.com)**
Ritsumeikan Global Innovation Research Organaization, Ritsumeikan University
2-150 Iwakura-cho, Ibaraki, Osaka, 567-8570 Japan

**Kouyou Otsu (k-otsu@fc.ritsumei.ac.jp)**
College of Information Science and Engineering, Ritsumeikan University
1-1-1 Noji-higashi, Kusatsu, Shiga, 525-8577 Japan

**Yugo Hayashi (yhayashi@fc.ritsumei.ac.jp)**
College of Comprehensive Psychology, Ritsumeikan University
2-150 Iwakura-cho, Ibaraki, Osaka, 567-8570 Japan

**Abstract**

This exploratory study investigated how the combination of top-down and bottom-up processing influences decision-making for high and low trusters using the ultimatum game against a computer agent. We designed an experiment wherein (1) participants expected their partners to be humans or agents (top-down processes) and (2) agents used one of four different types of algorithmic behavior (bottom-up processes) to propose and respond. We found that high trusters made fairer decisions in the human condition than in the agent condition in the proposal phase, when opponents' behaviors were not ambiguous but intentional. In the response phase, the higher the level of trust, the more likely they were to avoid unfairness for an opponent that proposed a distribution amount approved by the participant. Results suggest that, in interpersonal communication, high trusters flexibly use both types of cognitive processing to economically process information when developing representations of others and deciding on a response.

**Keywords:** general trust; decision making; top-down/bottom-up processes; social interaction; ultimatum game

## Introduction

Trust is a core element of everyday interactions and is crucial for social functioning for individuals, groups and nations (Helliwell, 2006; Putnam, 2000). General trust is defined as the default value of the trustworthiness of others when there is insufficient information regarding a particular partner (Rotter, 1967; Yamagishi, 2011; Yamagishi & Yamagishi, 1994). General trust promotes cooperation among people and enables them to find solutions to interpersonal problems; it contributes to the development and stability of democracy (Hooghe & Stolle, 2003; Putnam, Leonardi, & Nanetti, 1992; Uslaner, 1999). People with high levels of general trust (high trusters) actively seek relationships with new partners, while people with low levels of general trust (low trusters) try to maintain only existing human relationships (Yamagishi & Yamagishi, 1994). To clarify the impact of such differences in behavioral tendencies on social functioning, it is important to understand the differences in the behavioral characteristics of both types of truster.

Behaviors that invest cognitive resources are necessary to more accurately evaluate the trustworthiness of others and gain benefit. Cognitive resource investment behavior is behaving cautiously toward others and paying attention to information that suggests a lack of trust in those with whom they interact (Kikuchi, Watanabe & Yamagishi, 1997). High trusters appear to be just good-natured people, but they are more sensitive to the trustworthiness of others than low trusters are (Yamagishi & Yamagishi, 1994). For example, high trusters are not perceived as gullible by others (Rotter, 1967). However, once high trusters are deceived, the difference disappears between how high and low trusters perceive the trustworthiness of others (Hamsher, 1969; Wright, 1973). This difference also disappears when high trusters are informed in advance that an opponent is a liar (Geller, 1966). This suggests that individuals with high levels of trust do not always trust others. In addition, Kikuchi, Watanabe & Yamagishi (1997) used the prisoner's dilemma task, in which the opponent was unknown, to show that high trusters more accurately predicted their opponent's cooperative and/or non-cooperative behavior than low trusters did. In a dilemma task in which one can adjust one's gain, high trusters tend to decrease the gain if the partner is uncooperative and increase the gain if the partner is cooperative; however, low trusters showed a tendency to continue increasing gains and tended not to increase gains even after phases of noncooperation followed by cooperation. These findings suggest that high trusters are cautious people who are more sensitive to information that suggests another's trustworthiness.

### Influence of top-down and bottom-up processing in an ultimatum game

Previous studies using economic games report that high trusters are more sensitive to information that indicates the reliability of their partners than low trusters are. However, previous studies traditionally have focused on only one of top-down and bottom-up in cognitive processes, and differences between high and low trusters in integrative information processing are unknown. Cognition regarding



In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society.* 

others, such as responses to and impression formulation about others, is determined by two different styles of processing: top-down and bottom-up. Top-down processing is based on the socialized knowledge of others, such as interpersonal schemas (Cohen, 1981; Fisk & Taylor, 1991) or stereotypes (Dion, 1972). This processing is essential for developing representations of others in the initial stage of interaction and can be used as supplemental information when representations are difficult to develop based on others' behaviors. Bottom-up processing refers to evidence-based processing based on the actual actions of others. It is important to note that, in real interpersonal communication between humans, people flexibly use both types of cognitive processing to economically process information when developing representations of others and deciding on a response. However, it is unclear how the combination of top-down and bottom-up processing influences decision making for both high and low trusters.

The ultimatum game has been used to investigate how both types of cognitive processing influence decision-making. This game offers a reliable and valid method for studying interpersonal strategies such as fairness, selfishness, and reciprocity (Güth, 1995; Güth & Tietz, 1990). Two players play the ultimatum game: a proposer and a responder. First, the proposer receives a sum of money from the experimenter and makes a proposal to the responder concerning how to share the money with them. Second, the responder has two options for responding: to reject or accept the proposal. If the proposal is accepted by the responder, both players receive money according to the proposal; however, if the responder rejects the proposal, neither receives any money.

Recent studies on human–agent interaction have pointed out the importance of top-down and bottom-up cognitive processing (Miwa & Terai, 2006; 2012). To investigate top-down processing, previous studies have explored the impact of two types of opponents (human and agent conditions), by telling the participant that the partner was either human or a computer agent, but the opponent was the same agent in both conditions. Comparing human and agent conditions helps extract the behavioral characteristics of humans in the ultimatum game. Unfair proposals are more likely to be rejected in the human condition than in the agent condition (Sanfey, Rilling, Aronson, Nystrom & Cohen, 2003), suggesting that people distribute in response to the other person's intentions and that the larger this influence, the more top-down processing works according to the other party's attributes. For bottom-up processing, the actual behavior during interactions is used to update the representation of others, and there are different strategies. The previous study used a multi-period version of the ultimatum game and some representative strategy programs (e.g., random, adaptive, egocentric, and exocentric) in addition to comparing human and agent conditions (Hayashi & Okada, 2017). If there were behavioral changes in participants according to the programs, they were considered a bottom-up influence. If there was a change in participants' behavior between the human and agent conditions, it was considered a top-down influence. In the random program, the agent selected fair/unfair proposals randomly with equal probability. The researchers revealed that in the random condition, proposals made by an apparently human opponent were rejected more often than those made by a computer opponent (Hayashi & Okada, 2017). This approach helps with investigating the differences in behavioral characteristics between high and low trusters in terms of integrative cognitive processes.

## The current study

Previous studies on general trust have considered only human opponents, and considered either top-down or bottom-up processes. In the present study, we explored how high trusters obtained resources as proposers and responders compared with low trusters by adapting the integrative approach to the combination of expectations in top-down (human and agent conditions) and actual bottom-up behavior (random, adaptive, egocentric, and exocentric programs). We designed the instruction (human and agent) as a within-participants factor, and the programs as a between-participants factor. For the order of proposal and response phases, we focused on how high trusters allocated resources (points) when they had the initiative. Therefore, for all participants, first the proposal phase and then the response phase was performed to avoid inducing an order effect on the performance of the former.

This study had two goals. The first goal (Study 1) was to explore how the combination of intentions and actual behavior impacted cognitive processing in decision-making for high and low trusters. If high trusters are more sensitive to others' trustworthiness than low trusters are, we expected that the combination of intentions and actual behavior will differ between high and low trusters. The second goal (Study 2) was to investigate how high trusters processed information about the other's point acquisition status in the response phase. In the version of the ultimatum game we used, performing the response phase after the proposal phase induced a different points status for each participant when they switched to the response phase. That is, we could also explore whether high trusters adjusted their allocation after considering the information on others' point status in the response phase. For example, if a high truster received more points than an opponent in the proposal phase, would they reciprocally assign points to the opponent in the response phase?

# Methods

## Participants

Seventy-eight Japanese university students (47 women, 31 men, mean age [Mage]:19.14, standard deviation [SD]:1.13) were recruited to participate in the experiment. Each participant was assigned to one of four opponent strategy algorithms (random, adaptive, egocentric, or exocentric). All participants signed an informed consent form. This study was approved by the ethics committee of our university (approval number kinugasa-human-2021-36).

## Materials

The opponents of the human and agent conditions were same computer programs. The game was played in the Virtual Reality space to provide participants with a more real experience and opponents. The experimental software was developed using Unity software (Unity Technologies) and participants used a head-mounted display (Oculus Rift S). For the avatar of the opponent's agent, we used a humanoid model released on VRoidHub, a character model-sharing service operated by pixiv. The VR environment consisted of only three elements: a white floor, an agent, and a display screen. Participants selected items using the buttons on the controller held in their hands while referring to the description of the game content and options displayed on the screen.

## The questionnaire

The level of general trust each participant displayed was measured using the General Trust Scale (Yamagishi, & Yamagishi, 1994). This scale consists of six items: "Most people are basically honest," "Most people are trustworthy," "Most people are basically good and kind," "Most people are trustful of others," "I am trustful," and "Most people will respond in kind when they are trusted by others." Participants responded to each item on a 6-point scale (1 = strongly disagree to 6 = strongly agree). The score obtained by averaging the answers for each item for each individual was used as the general trust score for that participant. Higher scores corresponded to higher levels of trust.

## Behavior of the agent

Table 1 shows the rejection rates of the programs as responders in the proposal phase. As proposers in the response phase, all four programs offered a fair proposal (r4:500-500) in the first round. For subsequent rounds, the adaptive, egocentric, and exocentric conditions based the proposal on the participant's response decisions.

In the random condition, the agent responded with accept or reject, with a probability of 50%. As the proposer, the agent proposed one of the options at random, regardless of the content of the previous proposal and its approval or rejection. This allowed ambiguous behaviors to be investigated; that is, bottom-up processing was restricted.

In the adaptive condition, the agent as responder tended to accept proposals that were relatively favorable to the proposer; as proposer, the agent repeated the proposal if it was accepted and proposed a completely opposite monetary strategy (i.e., fair versus unfair) if it was not accepted. This strategy was based on the Pavlov strategy in social games, wherein the basic rules are "win–stay" and "lose–shift" (Nowak & Sigmund, 1992). The adaptive program allowed for the investigation of fairness and reciprocity behaviors because the participants could adjust their points relative to each other. The egocentric and exocentric conditions were derived from the adaptive conditions.

In the egocentric program, the agent, as the responder, accepted proposals that were relatively favorable to themselves. As proposer, the agent offered a proposal that was advantageous to itself but reacted economically, for example, proposing r3 if the participant continued accepting the proposal. The r3 is the option with the least loss for the agent among the options that are favorable to participants. This allowed us to investigate selfish behavior.

In the exocentric condition, the agent as a responder accepted proposals that were relatively favorable to participants, and as a proposer offered a proposal that was advantageous to participants. Schematics of all four algorithms as proposers are provided in the Appendix.

## Procedure

Two participants were placed in a psychological laboratory and instructed on how to play the ultimatum game. They were instructed to score more points throughout the game and were told that 1000 points corresponded to 1,000 Japanese yen (approximately 8 US dollars). Then, they were informed that they would participate in two separate conditions: one with a human partner (human condition) and the other with a computer agent partner (agent condition), and that they would play the ultimatum game in the human condition. The order of the human and agent conditions was counterbalanced across the participants. After receiving the instructions, one participant was moved to an adjoining room. The participants underwent 15 rounds of the proposal phase and 15 rounds of the response phase. In the proposal phase, in which the program acted as the responder, participants were told that 1,000 points would be provided to the proposer.

The proposer selected one of seven options, r1 to r7 (Table 1), for splitting the money. In the response phase, the program acted as the proposer and proposed the money allocation; participants either accepted or rejected the proposal. After concluding the 30-round task, each participant completed a General Trust Scale questionnaire. Finally, the experimenter conducted a debriefing and informed them that both opponents in the human and agent conditions were the same computer agents.

Table 1: Options for proposals and rejection rates corresponding to the proposals by the four types of agents.

| Proposals | Random | Adaptive | Ego centric | Exo centric |
|---|---|---|---|---|
| r1: 900-100 | 50% | 30% | 95% | 5% |
| r2: 800-200 | 50% | 20% | 80% | 20% |
| r3: 700-300 | 50% | 10% | 80% | 20% |
| r4: 500-500 | 50% | 0% | 100% | 0% |
| r5: 300-700 | 50% | 40% | 20% | 80% |
| r6: 200-800 | 50% | 70% | 20% | 80% |
| r7: 100-900 | 50% | 70% | 5% | 95% |

Notes: left values show points for a participant and right values show points for an agent in options for proposals.

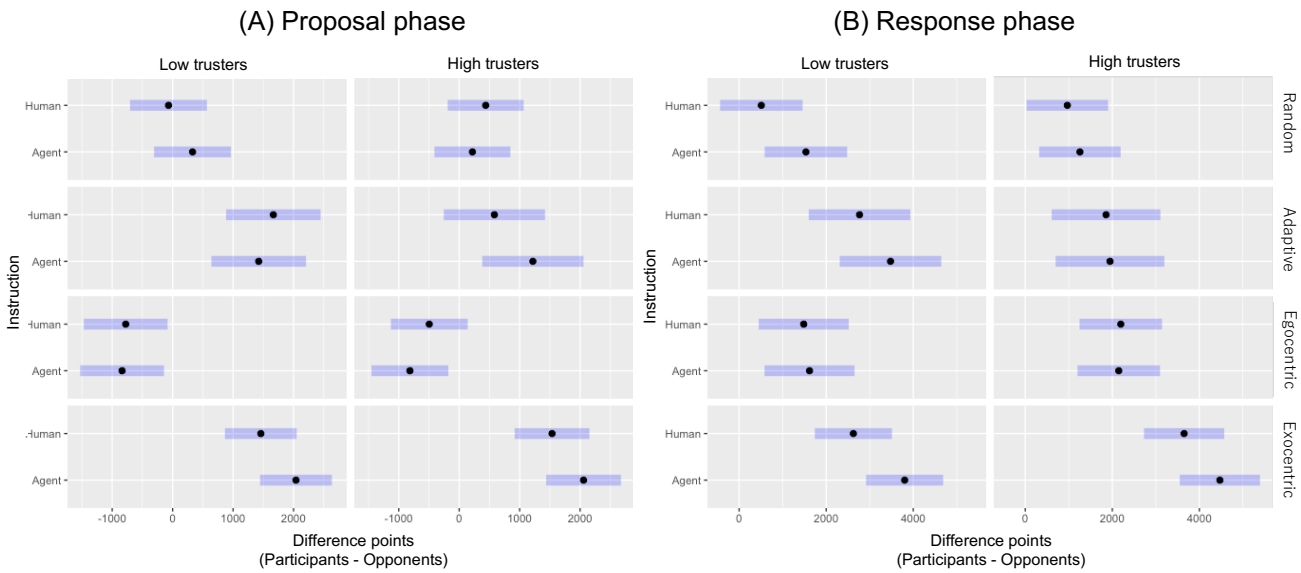


Figure 1. Difference points between participants and opponents according to instruction and general trust for each program. Higher positive values on the x-axis indicate more selfish distribution for participants, and negative values indicate more altruistic distributions. Black points indicate predicted values for participants who scored high or low on the general trust scale (high trusters: 1 SD above the mean; low trusters: 1 SD below the mean). Blue error bars represent 95% confidence intervals.

## Analysis

We used R (R Core Team, 2012) and the lme4 function (Bates, Mächler, Bolker & Walker, 2014) to perform generalized linear mixed-effects model analysis. Each phase (proposal and response) was modeled separately. We used the difference in points between participants and opponents as the dependent variable in all models. This was calculated by subtracting the points of opponents from that of the participants for each round. That is, the higher the positive difference point value, the more selfish the distribution for participants, and the higher the negative difference point value, the more altruistic the distribution. Therefore, this measure helped us compare selfishness and altruistic behavior of the participants (Kaneuchi, Miura & Karasawa, 2021).

The first goal was to investigate how the combination of expectations and actual behavior influences cognitive processing. We tested difference points as a dependent and the instruction (human or agent), general trust, and their interaction as independent variables (i.e., fixed effects). In all the models, we included the participants as random intercepts. Specifically, we focused on the two-way interaction (instruction × general trust). Multiple comparison tests using estimated marginal means were performed for interactions; emmeans (Lenth, Singmann, Love, Buerkner & Herve, 2018) are packages available in R (version 3.5.1; www.r-project.org). The Bonferroni method was used to adjust the p-values for multiple comparisons.

The second goal was to investigate how high trusters processed information on the acquisition status of each other's points during the response phase. We added the independent variable of the total points obtained in the proposal phase for each participant and focused on a three-way interaction (instruction × general trust × total points got in the proposal phase). If the three-way interaction was significant, we conducted post hoc estimation and comparisons of the slopes for trust for each condition using emtrends (emmeans package).

## Results

Twelve participants were excluded from the data analysis because they found out that their partners were not human (random: 2; adaptive: 2; egocentric: 4; exocentric: 4). Twenty-nine participants were female and 37 were male. The mean age was 19.04 years, with a standard deviation of 1.11. The random, adaptive, egocentric, and exocentric programs had 18, 16, 16, and 16 participants, respectively.

We investigated a three-way interaction (instruction × general trust × programs) (Study 1). In the proposal phase, there was a three-way interaction among instruction, general trust, and programs ($F = (3, 1943) = 6.36$; $p < .001$; Figure 1). For the random program, low trusters scored more difference points in the agent condition than in the human condition ($p = .004$). However, there was no significant effect for high trusters ($p = .11$). For the adaptive program, high trusters scored more difference points in the agent condition than in the human condition ($p < .001$). However, there was no significant effect for low trusters ($p = .16$). For the egocentric program, high trusters scored less difference points in the agent condition than in the human condition ($p = .02$). However, there was no significant effect for low trusters ($p = .69$). For the exocentric program, high and low trusters scored more difference points in the agent condition than in

the human condition ($p < .001$). In the response phase, there was no interaction among instruction, general trust, and programs ($F = (3, 1933) = .45$; $p = .71$).

We investigated a three-way interaction (instruction × general trust × total points scored in the proposal phase) (Study 2). Results showed that, in the adaptive program, there was a three-way interaction between instruction, general trust, and the total points scored in the proposal phase ($F = (1, 474) = 8.35$; $p = .004$; Figure 2). In the agent condition, the slope of trust was significant for the group with high total points ($B = -3030$; $SE = 1482$; 95% confidence level [$CI$; $-5993$, $-67.6$]). However, it was not significant for the group with low total scores ($B = 487$; $SE = 1428$; 95% $CI$ [$-2379$, 3353]). There was no significant difference between the general trust slopes for the groups that scored high and low total points in the proposal phase ($p = .06$). In the human condition, the slope of general trust was significant for both the high total points ($B = -2618$; $SE = 1196$; 95% $CI$ [$-5062$, $-174$]) and low total points ($B = 5780$; $SE = 1771$; 95% $CI$ [2263, 9296.4]) groups. There was a significant difference between the general trust slopes for the groups that scored high and low total points in the proposal phase ($p < .001$). In programs other than the adaptive program, there was no interaction among instruction, general trust, total points scored in the proposal phase (Random: $F = (1, 537) = 1.68$; $p = .20$; egocentric: $F = (1, 305) = 1.49$; $p = .22$; exocentric: $F = (1, 510) = .92$; $p = .33$).

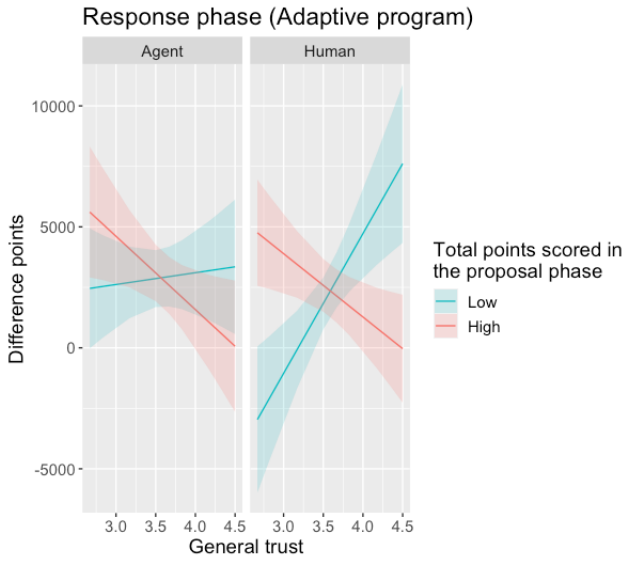


Figure 2. Slopes of trust for each of conditions. Difference points between participants and opponents according to general trust and instruction for each program. Red and blue lines indicate predicted values for participants who scored high or low in the proposal phase, respectively (high: 1 SD above the mean; low: 1 SD below the mean). Error bands represent 95% confidence intervals.

## Discussion

We investigated whether impacts of the combination of intentions and actual behavior on decision-making differed between high and low trusters. Statistical differences were found in the proposal phase. High trusters behaved more fairly in the human condition than in the agent condition for programs with intention in its actions other than the random program. While, low trusters behaved more fairly in the human condition than in the agent condition for the random program without intention in its actions and the exocentric program. The exocentric program was more altruistic than the adaptive one. Interestingly, both high and low trusters behaved more fairly in the human condition than in the agent condition in this program. However, there was no difference according to instruction among low trusters in the adaptive program. This indicates that high trusters are more sensitive to the reciprocity of others than low trusters are.

Then we investigated how high trusters processed information about the other's point acquisition status in the response phase. Results showed that, for the adaptive program, the higher the level of trust, the more the points were adjusted to avoid inequity according to those in the proposal phase, even though participants were instructed to score more points. And, the difference between the impacts of general trust for the groups that scored high and low total points in the proposal phase were confirmed in not the agent condition but the human condition. This suggests that high trusters have a social norm to reciprocally behave to human opponents that allow them to adjust their points relative to each other.

Our results on the difference of reciprocity between individuals with high and low trust can be explained by long-term interests according to the error management theory and the effect of reputation. The error management theory posits that psychological mechanisms are designed to exhibit predictable biases when judgments are made under uncertainty and the costs of false-positive (Type I) and false-negative (Type II) errors are unequal (Haselton & Buss, 2000). For example, in terms of the sensitivity of warning devices such as fire alarms, the cost of decreasing sensitivity is annoyance of misinformation (Type I error), whereas the cost of increasing sensitivity is missing a real fire (Type II error). In this case, the cost of missed detections (Type II errors) is higher, leading to a bias towards false positives (Type I errors). In the case of the ultimatum game, people are asked to choose behaviors that appeared kind or selfish to others. The cost of the former was the loss of money and the cost of the latter was being perceived as selfish by others. Reputation is an important factor in social interactions (Nowak, 2012), and it promotes cooperative behavior (Milinski, Semmann, & Krambeck, 2002). People care about what others think; their reward system responds positively to good evaluation and reputation (Izuma, Saito, & Sadato, 2008; Lieberman, 2013), while their anterior cingulate cortex and anterior insula respond in a similar way as in physical pain when they are isolated from the group (Lieberman, 2013). That is, choosing actions that appear to be kind to

others and sharing a certain amount of money with the other person reduces the cost of losing reputation and serves one's long-term interests.

We conjecture that, even in the case of the ultimatum game, the decision-making of high trusters depends more on such long-term interests than that of low trusters, according to the reciprocity of others. The adaptive program was somewhat altruistic as a responder and depended on the behavior of participants as proposers. In other words, the participants were given room for collaborative relationships. Because to cooperate with an adaptive opponent was rational based on long-term interests, high trusters behaved more fairly in the human condition than in the agent condition both in proposal and response phases. However, the egocentric program behaved selfishly, and participants were not given room for cooperative relationships. In this case, the participants were encouraged to reduce the cost of the loss of money. Therefore, high trusters were fairer in the human condition than in the agent condition. For the exocentric program, both high and low trusters behaved more fairly in the human condition than in the agent condition. Because the exocentric program was highly altruistic, the highly reciprocity also would influence low trusters. A random program does not intend to make decisions in terms of the bottom-up behavior. Stable cooperative relationships cannot be established without mutual intention. Therefore, there was no difference according to instruction in high trusters in the random program. In contrast, low trusters behaved more fairly in the human condition than in the agent condition. This indicates that low trusters tend to select a conservative strategy when interacting with human opponents in situations where there is insufficient information regarding a particular partner.

Our findings have implications for the comprehension of social interactions. What impact do the behavioral characteristics of high trusters have on social interaction in the real world? First, environments where trust is necessary are those with high social uncertainty (Yamagishi & Yamagishi, 1994; Yamagishi, 2011). High social uncertainty refers to situations where there is little information to evaluate the trustworthiness of the other person. In such cases, it appears to be more adaptive to only associate with a specific, trusted individual, as it reduces social uncertainty and creates a secure environment. However, in environments with high social uncertainty, there is a risk of increased opportunity costs. Opportunity cost refers to the benefit that could have been gained by investing the cost or time spent on one action into another action. Opportunity costs would be increasing in modern society, due to rapid networking. Therefore, in order to efficiently gain benefits at both the individual and societal levels in such environments, it is considered important to trust and engage with new others, rather than only associating with a specific individual. Previous studies have emphasized that high trusters are not easily deceived. The results of this study suggest the possibility that high trusters behave reciprocally towards intentional partners and can quickly form reciprocity with cooperative partners. The behavioral characteristics facilitate collaboration among individuals, enabling them to resolve interpersonal issues, and it also contributes to the growth and sustainability of democratic societies. The cultivation of trust is an extremely important challenge in modern society. Previous studies showed that the more a society engages in buying and selling goods in the market, the more likely individuals are to behave fairly towards unfamiliar others (Baldassarri, 2020; Henrich et al., 2004; Henrich et al., 2010). This indicates that trust is cultivated through interaction with the social environment. In future research, we should consider not only interpersonal communication but also the social environment in order to understand how trust is formed.

The implications of our study extend beyond the domain of human–agent interaction in terms of individual differences. High trusters promotes the treatment of agents as social actors. A previous study showed that participants tried to avoid inequity even when they believed that their opponent was a computer agent (Hayashi & Okada, 2017). The present study revealed that, when the adaptive program was used, participants with higher general trust gave points to the agent. This indicates that high trusters are more likely than low trusters to treat their counterparts as social actors. With the development of AI, interaction with computers is expected to increase even further in the future. It is important to understand how individuals perceive computer agents in such a scenario. This study suggests that general trust may be an important factor in understanding how both people and computers are perceived.

This study had two important limitations. First, we did not investigate a scenario in which the order of the response and proposal phases were changed because we focused on performance in the proposal phase, which was not influenced by the order effect. However, if participants were to act first as responders and then as proposers, it is possible that they would change the proposal according to the total points obtained in the response phase. Second, the condition of the four types of algorithmic behavior was not within, but between-participant factors. The present study confirms that unlike low trusters, high trusters can process both top-down and bottom-up information to adaptative opponents at the between-participants level. However, it is unknown whether the same individual changes strategy according to different programs at the within-participant level. Investigations on these aspects are necessary to clarify the behavioral characteristics of individuals with high levels of trust.

In summary, focusing on the long-overlooked issue, we investigated how the combination of top-down and bottom-up processing influences decision-making for high and low trusters using the ultimatum game. Our analyses suggest that high trusters, compared with low trusters, make decisions flexibly using both types of cognitive processing in interpersonal communication. Our findings have implications for understanding individual differences in reciprocity and perceiving an agent as a social actor. Therefore, our study of behavioral characteristics in general trust would contribute to both methodological and theoretical approaches to social interactions.

## Acknowledgments

This research was supported by Grant-in-Aid for Challenging Research (Exploratory) (JSPS KAKENHI Grant Number 20K20873).

## References

Chalnick, A., & Billman, D. (1988). Unsupervised learning of correlational structure. *Proceedings of the tenth annual conference of the cognitive science society* (pp. 510-516). Hillsdale, NJ: Lawrence Erlbaum Associates.

Feigenbaum, E. A. (1963). The simulation of verbal learning behavior. In E. A. Feigenbaum & J. Feldman (Eds.), *Computers and thought.* New York: McGraw-Hill.

Hill, J. A. C. (1983). A computational model of language acquisition in the two-year old. *Cognition and Brain Theory*, *6*, 287-317.

Matlock, T. (2001). *How real is fictive motion?* Doctoral dissertation, Psychology Department, University of California, Santa Cruz.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Ohlsson, S., & Langley, P. (1985). *Identifying solution paths in cognitive diagnosis* (Tech. Rep. CMU-RI-TR-85-2). Pittsburgh, PA: Carnegie Mellon University, The Robotics Institute.

Shrager, J., & Langley, P. (Eds.) (1990). *Computational models of scientific discovery and theory formation*. San Mateo, CA: Morgan Kaufmann.

Baldassarri, D. (2020). Market integration accounts for local variation in generalized altruism in a nationwide lost-letter experiment. *Proceedings of the National Academy of Sciences*, *117*(6), 2858-2863.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv,* arXiv:1406.5823.

Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences*, *4*, 317-331.

Dion, K. K. (1972). Physical attractiveness and evaluation of children's transgressions. *Journal of Personality and Social Psychology*, *24*, 207-213.

Fisk, T. S., & Taylor, E. S. (1991). *Social cognition*. New York: McGraw-Hill.

Geller, S. (1966). Magnetic behavior of substituted ferrimagnetic garnets. *Journal of Applied Physics*, *37*, 1408-1415.

Güth, W. (1995). On ultimatum bargaining experiments—a personal review. *Journal of Economic Behavior & Organization*, *27*, 329-344.

Güth, W., & Tietz, R. (1990). Ultimatum bargaining behavior: a survey and comparison of experimental results. *Journal of Economic Psychology*, *11*, 417-449.

Hamsher, J. H., Jr. (1969). Validity of personality inventories as a function of disguise of purpose. *Dissertation Abstracts International, 29*(8-B), 3086–3087.

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*(1), 81-91.

Hayashi, Y., & Okada, R. (2017). Compound effects of expectations and actual behaviors in human-agent interaction: Experimental investigation using the Ultimatum Game. *Proceedings of the 39th Annual Conference of the Cognitive Science Society (CogSci2017),* 2168-2173.

Helliwell, J. F. (2006). Well-being, social capital and public policy: what's new?. *The economic journal*, *116*(510), 34-45.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2004). Overview and synthesis. *Foundations of human sociality*, 8-54.

Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., ... & Ziker, J. (2010). Markets, religion, community size, and the evolution of fairness and punishment. *science*, *327*(5972), 1480-1484.

Hooghe, M., & Stolle, D. (Eds.). (2003). *Generating social capital: civil society and institutions in comparative perspective*. Palgrave Macmillan New York.

Izuma, K., Saito, D. N., & Sadato, N. (2008). Processing of social and monetary rewards in the human striatum. *Neuron*, *58*(2), 284-294.

Kaneuchi, S., Miura, A., & Karasawa, M. (2021). Do selfish behavior and altruistic behavior generalize?: An experimental investigation using the dictator game. *The Japanese Journal of Psychology*, *92*(1), 40－45.

Kikuchi, M., Watanabe, Y., & Yamagishi, T. (1997). Judgment accuracy of other's trustworthiness and general trust: an experimental study. *The Japanese Journal of Experimental Social Psychology*, *37*, 23-36.

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2018). Package "Emmeans". R Package Version 4.0-3.

Lieberman, M. D. (2013). *Social: Why our brains are wired to connect*. OUP Oxford.

Milinski, M., Semmann, D., & Krambeck, H. J. (2002). Reputation helps solve the 'tragedy of the commons'. *Nature*, *415*(6870), 424-426.

Miwa, K., & Terai, H. (2006). Analysis of human-human and human-computer agent interactions from the viewpoint of design of and attribution to a partner. *Proceedings of the 28th Annual Meeting of the Cognitive Science Society (CogSci2006),* 597-602.

Miwa, K., & Terai, H. (2012). Impact of two types of partner, perceived or actual, in human–human and human–agent interaction. *Computers in Human Behavior*, *28*, 1286-1297.

Nowak, M. A. (2012). Why we help. *Scientific american*, *307*(1), 34-39.

Nowak, M. A., & Sigmund, K. (1992). Tit for tat in heterogeneous populations. *Nature*, *355*(6357), 250-253.

Putnam, R. D., Leonardi, R., & Nanetti, R. Y. (1992). *Making democracy work: Civic traditions in modern Italy*. Princeton university press.

Putnam, R. D. (2000). Bowling alone: America's declining social capital: originally published in journal of democracy 6 (1), 1995. *Culture and Politics: A Reader*, 223-234.

Rotter, J. B. (1967). A new scale for the measurement of interpersonal trust. *Journal of Personality, 35,* 651-665.
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*, 1755-1758.
Uslaner, E. M. (1999). Democracy and social capital. In M. E. Warren (Ed.), *Democracy and trust.* Cambridge: Cambridge University Press.
Wright, T. L. (1973). Situational and personality parameters of interpersonal trust in a modified prisoner's dilemma game.
Yamagishi, T. (2011). Trust: The evolutionary game of mind and society. Springer Science & Business Media.
Yamagishi, T., & Yamagishi, M. (1994). Trust and commitment in the United States and Japan. *Motivation and Emotion*, *18*, 129-166.

## Appendix

Algorithm schematics of each program for acceptance or rejection by participants

### Random

[r1 - r7]
accept OR reject -> r1 - r7: 14.29%

### Adaptive

[r1 - r3]
accept -> %SAME%
reject -> r5 - r7: 33.33%
[r4]
accept OR reject -> r1 - r3, r5 - r7: 16.66%
[r5 - r7]
accept -> %SAME%
reject -> r1 - r3: 33.33%

### Egocentric

[r1]
accept -> r1: 10%, r2: 20%, r3: 70%
reject -> r5: 10%, r6: 20%, r7: 70%
[r2]
accept -> r2: 30%, r3: 70%
reject -> r5: 10%, r6: 20%, r7: 70%
[r3]
accept -> r3: 100%
reject -> r5: 10%, r6: 20%, r7: 70%
[r4]
accept -> r1: 10%, r2: 20%, r3: 70%
reject -> r5: 10%, r6: 20%, r7: 70%
[r5]
accept -> r5: 10%, r6: 20%, r7: 70%
reject -> r1: 10%, r2: 20%, r3: 70%
[r6]
accept -> r6: 30%, r7: 70%
reject -> r1: 10%, r2: 20%, r3: 70%
[r7]
accept -> r7: 100%
reject -> r1: 10%, r2: 20%, r3: 70%

### Exocentric

[r1]
accept -> r1: 100%
reject -> r5: 70%, r6: 20%, r7: 10%
[r2]
accept -> r1: 70%, r2: 30%
reject -> r5: 70%, r6: 20%, r7: 10%
[r3]
accept -> r1: 70%, r2: 20%, r3: 10%
reject -> r5: 70%, r6: 20%, r7: 10%
[r4]
accept -> r1: 70%, r2: 20%, r3: 10%
reject -> r5: 70%, r6: 20%, r7: 10%
[r5]
accept -> r5: 100%
reject -> r1: 70%, r2: 20%, r3: 10%
[r6]
accept -> r5: 70%, r6: 30%
reject -> r1: 70%, r2: 20%, r3: 10%
[r7]
accept -> r5: 70%, r6: 20%, r7: 10%
reject -> r1: 70%, r2: 20%, r3: 10%