Computers in Human Behavior 28 (2012) 1286–1297

Contents lists available at SciVerse ScienceDirect

Computers in Human Behavior

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

# Impact of two types of partner, perceived or actual, in human–human and human–agent interaction

Kazuhisa Miwa *, Hitoshi Terai

*Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan*

ARTICLE INFO

*Article history:*
Available online 9 March 2012

*Keywords:*
Social dilemma
Human–human interaction
Human-agent interaction
Prisoner's dilemma game
Media Equation

ABSTRACT

Participants engaged in the Prisoner's dilemma game with a partner through a computer terminal. We define two types of partner: a perceived partner and an actual partner, and manipulated the two factors independently. A perceived partner means a partner with whom participants imagined themselves to be interacting; instruction given by an experimenter controls the image of the perceived partner. An actual partner can change its behavior. In one scenario participants actually interacted with a human partner, in another scenario their partner was either a mostly cooperating computer agent or a mostly defecting computer agent. Three experiments were performed. The result suggested that the participants' selection behavior was largely influenced by the instruction given about the partner by the experimenter and not influenced by the partner's actual behavior. The analysis of the participants' impressions of the partner showed that the effect of instruction about the partner disappeared. Individual likeability for a partner was very influenced by the partner's behavior; as the participants incurred more defect actions from the partner, individual likeability for the partner decreased. On the other hand, social likeability for a partner was not so influenced by the partner's behavior, but rather related to the participants' own behavior. The participants who made more defect actions rated their partner's social likeability lower.



## 1. Introduction

In this paper, we investigate commonalities and differences in human social responses to another human or a computer agent. In particular, we deal with responses in situations in which there are dilemmas. A social dilemma is defined as an individual conflict that emerges when: (1) individuals can choose between cooperative and non-cooperative alternatives, (2) positive benefits are gained by persons taking non-cooperative actions, and (3) if all members of the group select non-cooperative actions, then each individual gains less benefit than when all members select cooperative actions (Dawes, 1980). Such dilemma situations have been actively investigated in many academic fields such as psychology, sociology, and economics as a representative situation in which crucial features of human social responses are observed. Social dilemmas relate to common problems that underlie many social phenomena such as destruction of environment, ambient pollution, development of desertification, overhunting, protected trade, entrance examinations, illegal parking, and dust pollution (Hardin, 1968).

There is a long history of studies on social dilemmas. One representative series of laboratory studies analyzed decision making using a dilemma task (Messick & Brewer, 1983). The most popular task is the Prisoner's dilemma game. There are two types of the dilemma tasks. One is the "one-shot dilemma game" and the other is the "iterated dilemma game". The latter was used in this study. In the one-shot game, the rational choice is the defect (non-cooperative) action. In the iterated game, the rational action is determined by multiple complicating factors because a single action influences the partner's successive actions. Altruistic egoism is identified as one principle for rational behavior in the iterative dilemma game (Pruitt & Kimmel, 1977; Taylor, 1987) where participants try to induce cooperative behavior from another by performing cooperative actions. They establish mutual cooperation while behaving altruistically, and try to maximize everyone's benefits. Pruitt and Kimmel (1977) proposed two criteria for inducing mutual cooperation in the Prisoner's dilemma game: (1) group members notice the importance of long-term benefits by establishing mutual cooperation, stopping egoistic pursuits for short-term benefits, and (2) mutual trust appears among the group members, confirming that the other members do not fail to cooperate.

The Tit for Tat (TFT) strategy is known to be an effective strategy for inducing mutual cooperation where participants do whatever the other player did in the previous trail. They cooperate when the other cooperated and defect when the other did not. In Axelrod (1980a), a computer tournament was held in which a total of 14 strategies were tested. Selection rules were submitted by experts in game theory from a variety of disciplines: psychology, political science, economics, sociology, and mathematics. The winner was TFT. A more

* Corresponding author.
*E-mail address:* miwa@is.nagoya-u.ac.jp (K. Miwa).

0747-5632/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.chb.2012.02.012

thoroughgoing tournament in which 62 strategies were tested; again showed the winner was TFT (Axelrod, 1980b). In representative situations, it has been demonstrated that when TFT strategy is followed, mutual trust between all members emerges in the group.

The findings above regarding the human nature of social responses to others have been accumulated mainly in social psychological studies. In the current study, we investigate, when people face a dilemma situation, what commonalities and differences in human social responses emerge in human–human interaction: i.e., when the partner is a human, or in human–agent interaction: i.e., when the partner is a computer agent.

### 1.1. Two types of partners

To investigate human social responses in a dilemma situation, cognition about others is a crucial issue. Cognition about others such as responses to others and impression formulation about others is determined based on two different styles of processing: top-down and bottom-up processing. The bottom-up processing means the evidence based processing performed based on others' actual figures and actions. On the other hand, many social psychological studies have indicated that human cognition about others is greatly influenced by the top-down processing, driven by socialized knowledge and knowledge about others. Robust evidence supporting such top-down processing: e.g., having a schema about others (Cohen, 1981; Fiske & Taylor, 1991) or a stereotype (Dion, 1972) activates specific information processing and compensates for insufficient information (Cantor & Mischel, 1977).

In understanding the relationship between responses to another human or a computer agent, these two types of processing are embodied by two types of partner: perceived and actual, which are defined as critical concepts in the current study. Fig. 1 illustrates the experimental setting in the current study.

#### 1.1.1. Perceived partner

A perceived partner influences the top-down processing in cognition about others, and is the partner with whom participants imagine they are interacting. This partner is represented in the participants' mind. Participants can believe that they are interacting with either to a human or to a computer agent, independently of whether they are actually interacting with a human or an agent. Instruction given from an experimenter controls the image of the perceived partner.

#### 1.1.2. Actual partner

An actual partner involves the bottom-up processing. An actual partner usually changes its behavior and the information that participants receive. An important point is that the two types of partners are basically independent factors; i.e., a perceived partner in the mind is basically independent of how an actual partner behaves. For example, we can imagine a case in which we believe ourselves to be interacting with a human (perceived partner), but his/her behavior is stylized because we actually are interacting with a computer agent (actual partner).

### 1.2. Effect of perceived partner

In preceding studies, the participants' image of the partner with which they interacted, a human or a computer agent, was manipulated by instructions from the experimenter, telling them that their partner was a human or a computer agent. In most cases, the actual partner was a computer agent whose behavior and responses were systematically controlled even when the participants had been instructed that their partner was human.

Those studies indicated that participants' responses differ depending on the instruction about whom the partner is. In an earlier attempt (Yamamoto, Matsui, Hiraki, Umeda, & Anzai, 1994), the participants played Shiritori, a popular word game in Japan, with a partner through a computer terminal. Shiritori is a game played by saying a word that starts with the last syllable of the word given by the previous player. The participants were actually playing Shiritori with a computer program installed in a computer. Instruction was manipulated; in one situation the participants were informed that the partner was a computer program and in the other situation they were informed the partner was a human at another campus connected by the Internet. The result showed that the degree of enjoyment of the game rated by the participants depended on the instruction given. When the participants were instructed that the partner was a human, they engaged in the game longer and felt happier while playing.

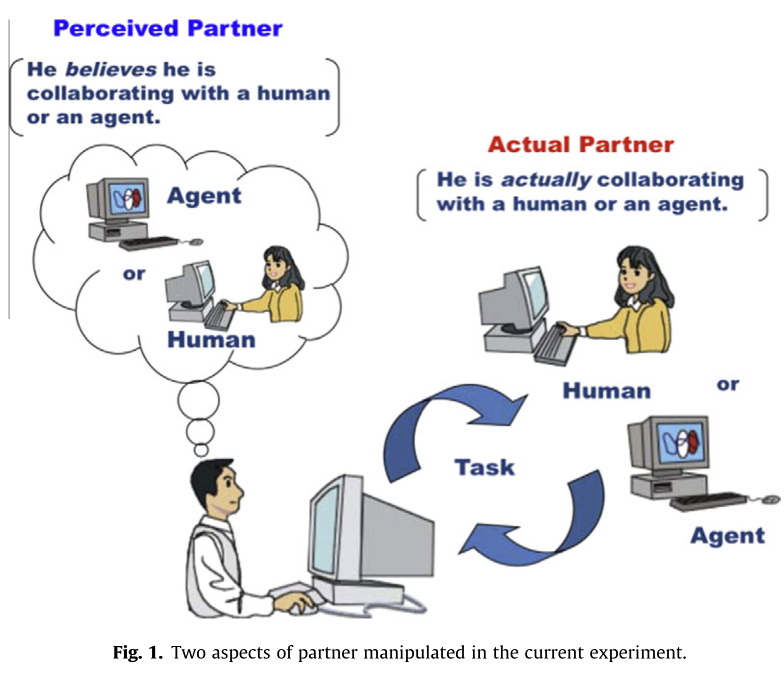


**Fig. 1.** Two aspects of partner manipulated in the current experiment.

In Sundar and Nass (2000), the participants were tutored through a computer terminal. In Experiment 1, the participants were told that their tutor was the computer in front of them or the imaginary programmer who produced the tutoring program. Messages were completely identical for both the conditions. A questionnaire accessing the participants' impressions of their tutor showed that judgments of friendliness, playfulness, effectiveness, and style similarity, differed depending on the instruction given about the tutor. In Experiment 2, the participants were instructed as to whether the partner was a computer or a person in another room connected through the Internet. In the latter case, the actual partner was also a computer program. The result was consistent with that in Experiment 1. The evaluation, from three viewpoints: friendliness, playfulness, effectiveness, was changed by the instruction given.

These results show that participants' mental representation of a partner, a human or computer agent, has a relatively large impact on the pattern of interaction between the participants and their partner.

### 1.3. Effect of actual partner

In Burgoon et al. (2000), participants engaged in the Desert Survival Problem. In the experiment, they were told that their partner was a human, even though in all cases but one the partner was a computer agent that generated monotonous responses. As a result, the effect of the partner's behavior revealed significant results in the participants' evaluation of the expertise of the partner.

Several studies did not manipulate the actual partner, but controlled its appearances. In Gong (2008), the participants were informed that they were interacting with an anthropomorphic computer agent even when a real human face was used for the interface. In this situation, the appearance of the partner was consistently manipulated. The result of this study provided support for the assumed linear relationship between the degree of anthropomorphism of computer representations and people's social responses. When facial representations on computers progressed from low-anthropomorphism to medium-anthropomorphism to high-anthropomorphism, and to real human images, people gave them more positive social judgments, greater attribution of homophily, higher competency and trustworthiness ratings, and were more influenced by them in choices of dilemma decision-making.

In the above studies, appearance of a partner was manipulated while the representation of the partner was fixed. However, in most studies, the manipulation of an actual partner did not only change the behavior nor the appearance of the partner but also influenced the representation of the partner. For example, in Mol, Krahmer, Maes, and Swerts (2009), the participants watched an animation and were asked to explain the story to a partner. Gestures emerging through the explanation were recorded. One of the main results was that the participants produced a lower number of gestures when talking to the artificial agent than when talking to a human partner. In the "interaction with a human" condition, the participants thought that they were relating to a human, while in the "interaction with an agent" conditions, they believed they were interacting with a computer agent.

This type of confound design was also found in other studies where media for communication or the appearances of a partner, i.e., the anthropomorphic degree of a partner, were controlled. In such studies, the factors focused on (i.e., media types or the anthropomorphic degree) and the representation of the partner (i.e., participants believe themselves to be interacting with a human or a computer agent) were not independently manipulated. In Kiesler, Waters, and Sproull (1996), the participants engaged in the Prisoner's dilemma game. They analyzed to what degree the participants' decisions were changed. As a result, especially with the cooperative behavior, the degree of keeping to their decision was higher when the partner was a human than when the partner was a computer agent. Parise, Kiesler, Sproull, and Waters (1999) in which the appearance of the partner was manipulated showed that the participants made and kept promises to cooperate with the human-like agent as much as they did with a human confederate. General evaluations of likability of the agent did not lead the participants to cooperate with it.

### 1.4. Independent manipulation of partner

It seems that both types, i.e., the representation of a partner constructed as a perceived partner and an actual partner, may be essentially important to determine interaction between participants and their partner. For example, in Aharoni and Fridlund (2007), the participants took a mock job interview through a computer terminal. In the experiment, the participants were instructed on whether the partner was a human or a computer agent, thus controlling the way the participant represented the partner. Additionally, two experimental conditions were set up; one case in which the participants were informed they had been accepted for the job as a result of the interview and the other case in which they were informed they had been rejected, thus controlling the actual behavior of the partner. The participants' facial expressions were recorded during the interview by a video camera. The effect of the instruction emerged in several kinds of expressions such as smiling and silence filling. The participants smiled more when they were instructed that the partner was a human. On the other hand, impressions of the partner depended on the partner's final decision, acceptance or rejection, more than the effect of instruction. This study implies that both factors, representation and actual behavior, are crucial for determining interaction with a partner. Much evidence needs to be accumulated by systematic experiments in which the two factors are independently manipulated. This motivation is the starting point of the current study.

Dehn (2000) pointed out that in many human computer interaction (HCI) studies multiple factors simultaneously varied between the control and experimental conditions. Therefore, it was sometimes difficult to determine which factor causes the result. Gong (2008) more specifically indicated five factors that are sometimes confounded in HCI studies. He indicated that there is an essential difference between "being human-like" versus "actually being human". This suggestion is interpreted in the context above, that being human-like means that the representation of a partner is not a human, but behavior of the partner is human-like. On the other hand, actually being human means that the representation of a partner is a human.

## 2. Research questions and hypotheses

We aim to understand commonalities and differences of human social responses in human–human and human–agent interaction; and also, by independently controlling two types of partners, perceived and actual, to investigate how the schema-based top-down and evidence-based bottom-up processing in cognition about others relates to the nature of interaction. Research questions in the current study and related hypotheses are as follows.

### 2.1. Research question RQ1

RQ1: How is the participants' strategy selection behavior influenced by the two aspects of their partner: perceived or actual?

**Hypothesis H1.** Participants will be influenced to a greater degree by the representation of the partner (perceived partner) than by the partner's actual behavior.

This hypothesis comes from the preceding studies. As we mentioned, many studies have indicated that experimental instruction about a partner largely impacts on participants' behavior. Therefore, in our case, the participants' decision to cooperate or defect, may differ largely, depending on whether they believe they are interacting with a human or a computer agent.

**Hypothesis H2.** Participants will be influenced to a greater degree by behavior of the partner, i.e., the partner's selected pattern of cooperate or defect, than by the representation of the partner (perceived partner).

This hypothesis comes from our naïve prediction: the more the partner defects the more the participant will choose to defect. "An eye for an eye, a tooth for a tooth" may be one of the most basic principles of human behavior.

**Hypothesis H3.** The influence of one aspect of a partner varies depending on the other aspect of a partner.

### 2.2. Research question RQ2

RQ2: How are participants' impressions about their partner influenced by the two aspects of a partner: perceived or actual?

**Hypothesis H4.** Participants will be influenced to a greater degree by the representation of a partner (perceived partner) than by actual behavior.

**Hypothesis H5.** Participants will be influenced to a greater degree by behavior of the partner, i.e., the partner's selected pattern of cooperate or defect, than by the representation of the partner (perceived partner).

Two hypotheses above refer to how participants' impressions of a partner are determined by the two aspects of a partner. Additionally, the impressions may also relate to the participants' own behavior. Therefore, the following query is added to RQ2: how do the participants' impressions about a partner relate to their own actions to a partner?

**Hypothesis H6.** Participants impressions of a partner relate to participants' own selective actions rather than relate to the partner's behavior.

## 3. Experimental paradigm in current study

### 3.1. Task

In the current study, we use the Prisoner's dilemma game in which each participant faces two alternative actions: (1) cooperation: i.e., doing a socially responsible thing and (2) defection: i.e., acting based on self-interest regardless how this might harm the partner. Each participant is better off defecting regardless of the partner's choice, but the sum of the participants' payoffs is maximized if both participants choose to cooperate; so a dilemma emerges. Table 1 shows the payoff matrix used in the current study. For example, when both participants offer cooperation, both receive 120 Yen; however, when one participant offers cooperation while the other offers defection, the former receives nothing while the latter receive 180 Yen.

**Table 1**
Payoff matrix in Prisoner's dilemma game.

| | B: Cooperate | B: Defect |
|---|---|---|
| A: Cooperate | A: 120 Yen/B: 120 Yen | A: 0 Yen/B: 180 Yen |
| A: Defect | A: 180 Yen/B: 0 Yen | A: 60 Yen/B: 60 Yen |

### 3.2. Experimental manipulation

#### 3.2.1. Actual partner's behavior

Participants repeat 11 trials in the dilemma game with a controlled partner. In the game, behavior of the participant is defined as: a selective pattern of cooperate and defect actions.

As an independent variable, the first experimental factor is related to the actual partner's behavior. This factor was controlled by manipulating the partner with which participants actually competed. Three cases were set up: (1) competition with a human partner (w/ Human), and (2) competition with a computer agent. The latter case was subdivided into two sub cases: (2a) competition with an agent who uses the cooperation strategy in decision-making (w/ C-agent), and (2b) competition with an agent who uses the defection strategy (w/ D-agent). The C-agent offered ten cooperate actions and only one defect action in the 6th trial among 11 trials. The D-agent offered six cooperate actions and five defect actions in the 2nd, 4th, 5th, 7th, and 10th trials. The first factor was manipulated as follows. When competing with a human partner, each terminal for communication was connected to the Internet, and each participant solves the task with a partner who simultaneously engages in this task using another Internet connected computer. When competing with a computer agent, each participant solves the task with the agent installed on the computer they are using.

Participants were not informed what strategy their partner actually uses. Available information about the partner was only the partner's behavior. In collaboration with a computer agent, the rate of defect actions was accurately manipulated. In the w/ Human condition, there were no such controls. Each participant played the game with another participant without any instruction about the strategy.

#### 3.2.2. Perceived partner

The second factor is related to the representation of the partner which was manipulated by instructing participants to compete with a program installed on a computer (w/ agent) or with a human partner sitting in front of another computer terminal (w/ human).

The Wizard of Oz method in the field of the development of computer agents, in which for collecting corpus data of natural conversations participants are guided to believe their partner is a computer agent but the actual partner is a human, corresponds to one condition in the current experiment (Dahlbak, Jonsson, & Ahrenberg, 1993).

## 4. Structure of experiments

We conducted a total of three experiments. The standard experimental procedure is as follows. The participants made 11 decisions one by one. After seeing their partner's decision in the preceding trial, they were then required to make their next decision. Fig. 2 shows an example screen shot of the computer terminal used in the experiment.

In the human subject condition (w/ human instruction), participants introduced themselves face-to-face, and then moved to their respective computer terminals. In the computer agent condition (w/ agent instruction), participants sat in front of an assigned computer terminal and immediately engaged in the task.

Table 1 shows the payoff matrix used throughout the three experiments.

Experiment 1 is a fundamental experiment. To increase the reality of the competitive situation in the dilemma game, the participants were instructed that they would be actually paid based on

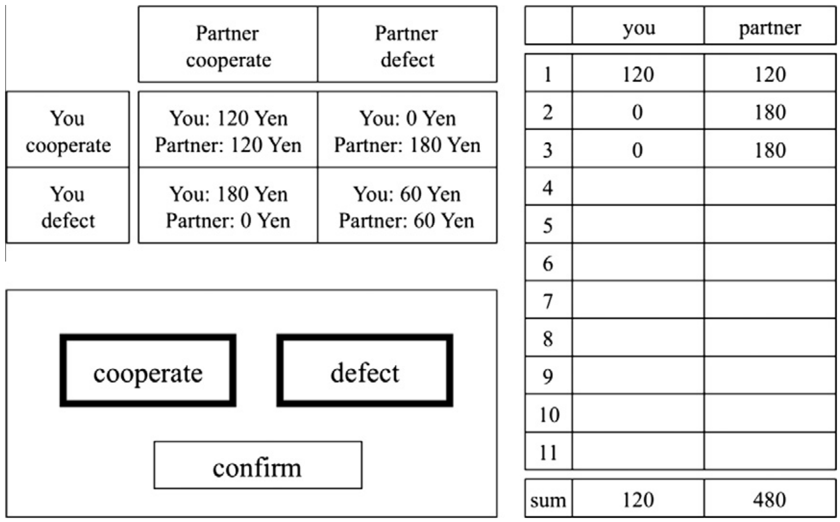


**Fig. 2.** Example screen shot of computer terminal for experiment.

the payoff matrix in Table 1. After the experiment, the maximum expense: i.e., 1980 Yen (=180 Yen/trial × 11 trials) was actually paid to every participant. In Experiments 2 and 3, the participants were instructed that they are playing a game with the objective to obtain as many points as possible.

In Experiments 1 and 2, when the participants were instructed that their partner was a human, they knew their partner's identity because they introduced themselves face-to-face. In Experiment 3, the participants' self-introduction to a partner in the initial stage of the experiment was excluded from the experimental procedure.

After all 11 decisions were made, a questionnaire, developed by Hayashi (1978), was completed to determine personality impressions about the partner. The impressions were analyzed from two viewpoints: *social desirability*, and *individual likeability*.

The social desirability was evaluated by five items: responsible/undependable, tidy/careless, profound/shallow, dominant/subservient, and prudent/imprudent; and the individual likeability was evaluated by two items: lovable/hateful and friendly/unfriendly. The participants rated each item using a 1–7 scale, with maximum likeability (or desirability) as 7 and minimum likeability (or desirability) as 1. Average scores were used for evaluation.

In the following, statistical analysis of each experiment is presented with each experimental result; and discussion about each analysis is made as a while after the results of all three experiments.

## 5. Experiment 1

### 5.1. Participants

One hundred forty-five undergraduates participated, and were randomly assigned to one of six conditions, in which the number of participants in each condition was as equal as possible. The numbers of participants in (1) w/ human (instruction) and w/ Human (actual partner), (2) w/ human and w/ C-agent, (3) w/ human and w/ D-agent, (4) w/ agent and w/Human, (5) w/ agent and w/ C-agent, and (6) w/ agent and w/ D-agent were 24, 26, 25, 26, 23, and 21, respectively.

### 5.2. Result

In the following statistical analysis, proportions and percentages were converted by the arcsine transformation, and .05 was used as the level of statistical significance.

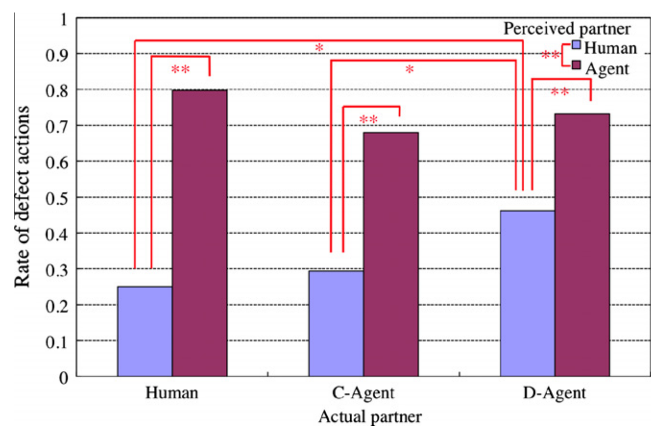


**Fig. 3.** Rate of defect actions in participants' decision-making.

#### 5.2.1. Selection behavior

First, to consider RQ1, we analyzed the participants' selection behavior. Fig. 3 summarizes the participants' selection behavior where the vertical axis indicates the rate of the participants' defect actions, and the horizontal axis indicates each experimental condition. A two (instruction) × three (actual partner) ANOVA revealed that the main effect of instruction reached significance ($F(1,139) = 87.53$, $p < .01$), but not the main effect of the actual partner ($F(2,139) = 1.88$, n.s.). The interaction also reached significance ($F(2,139) = 4.05$, $p < .05$). The simple main effect of instruction at every level for the actual partner factor, w/ Human, w/ C-agent, and w/ D-agent, revealed significance ($F(1,139) = 56.96$, $p < .01$; $F(1,139) = 26.00$, $p < .01$; $F(1,139) = 12.66$, $p < .01$). The simple main effect of the actual partner at w/ human instruction revealed significance ($F(2,139) = 4.61$, $p < .05$), where a Ryan' multiple comparison analysis showed that the rate of defection was higher at w/ D-agent than at w/ Human and w/ C-agent ($MSe = 310.19$, $p < .05$). Otherwise the simple main effect of the actual partner at w/ agent instruction did not reach significance ($F(2,139) = 1.31$, n.s.).

The analysis shows that people generally selected more defect actions when instructed that their partner was a computer agent. When instructed that their partner was a human, the rate of defect actions decreased; and they selected more defect actions when competing with a partner who selected more defect actions than when competing with a partner who selected fewer defect actions.

#### *5.2.2. Impressions about the partner caused by the partner's behavior*

Fig. 4 shows the participants' impressions about their partner where the vertical axis indicates the average score of the participants' rating, and the horizontal axis indicates each experimental condition. For individual likeability about a partner, a two (instruction) × three (actual partner) ANOVA revealed that the main effect of instruction and the actual partner reached significance ($F(1,132) = 17.07$, $p < .01$; $F(2,132) = 13.53$, $p < .01$). The interaction also reached significance ($F(2,132) = 11.54$, $p < .01$). The simple main effect of instruction at w/ Human as the actual partner revealed significance ($F(1,132) = 39.77$, $p < .01$), but not at w/ C-agent ($F(1,132) = 0.13$, n.s.) and at w/ D-agent ($F(1,132) = 0.24$, n.s.). The simple main effects of the actual partner at both levels of instruction, w/ human and w/ agent, revealed significance ($F(2,132) = 10.48$, $p < .01$; $F(2,132) = 14.59$, $p < .01$), where a Ryan's multiple comparison analysis showed that the rate of defection was higher at w/ Human and w/ C-agent than at w/ D-agent ($MSe = 183.18$, $p < .05$) when instructed that the partner was human, and higher at w/ C-agent than at w/ Human and w/ D-agent ($MSe = 183.18$, $p < .05$) when instructed that the partner was an agent respectively. For social desirability about a partner, a two (instruction) × three (actual partner) ANOVA revealed that both the main effects of instruction and the actual partner reached significance ($F(1,132) = 23.50$, $p < .01$; $F(2,132) = 3.27$, $p < .05$). The interaction also reached significance ($F(2,132) = 6.07$, $p < .01$). The simple main effects of instruction at w/ Human and w/ C-agent revealed significance ($F(1,132) = 27.86$, $p < .01$; $F(1,132) = 7.67$, $p < .01$), but not at w/ D-agent ($F(1,132) = 0.12$, n.s.). The simple main effect of the actual partner at w/ human instruction was also significant ($F(2,132) = 9.05$, $p < .01$), where a Ryan's multiple comparison analysis showed that the rate of social desirability was higher at w/ Human than at w/ C-agent and w/ D-agent ($MSe = 79.76$, $p < .05$), but the difference between at w/ C-agent and at w/ D-agent was not found ($MSe = 79.76$, $p < .05$).

#### *5.2.3. Impressions of the partner caused by participants' own behavior*

To investigate RQ2 about how participants' impressions about their partner relate to their own actions. The participants were divided to two groups. We calculated the average number of the participants' defect actions across the experimental conditions. Those who offered more defect actions than the average were categorized as a high defect group; those who offered less were categorized as a low defect group. Fig. 5 shows the participants' impressions of a partner where the vertical axis indicates the average score of the participants' rating, and the horizontal axis indicates each of the low and high defect groups. A two (instruction) × two (participants' action) ANOVA revealed that the main effect of instruction for individual likeability of a partner reached significance ($F(1,134) = 12.74$, $p < .01$), but the main effect of the participants' action did not ($F(1,134) = 0.00$, n.s.). The interaction also reached significance ($F(1,134) = 5.75$, $p < .05$). There was a simple main effect of instruction at the low defect group ($F(1,134) = 17.80$,

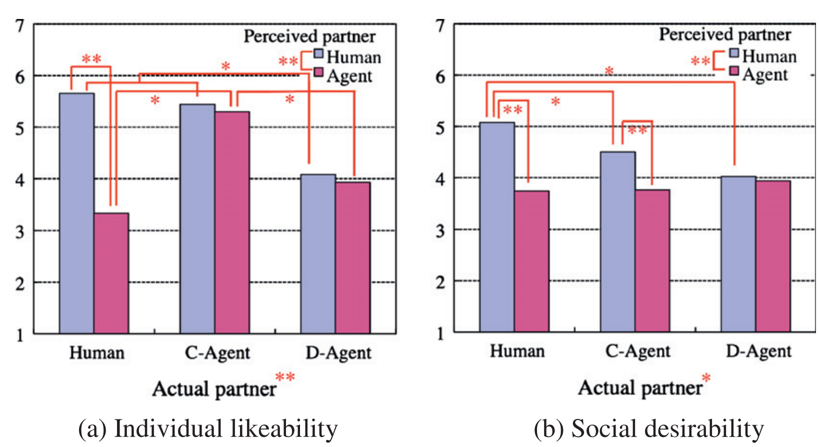


**Fig. 4.** Average score of participants' impressions about partner as function of experimental setting.

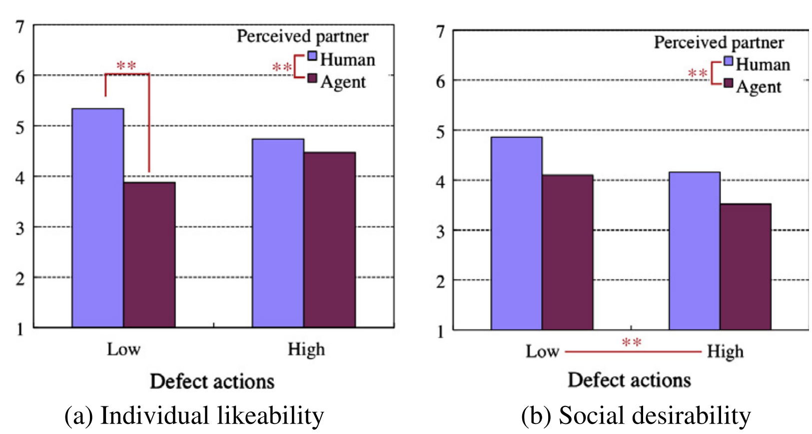


**Fig. 5.** Average score of participants' impressions of partner as function of participants' selection behavior, i.e., low and high defect actions.

$p < .01$), but the simple main effect at the high defect group was not found ($F(1,134) = 0.68$, n.s.). None of simple main effects of the participants' action at any level of the instruction factor were found ($F(1,134) = 3.04$, n.s., $F(1,134) = 2.72$, n.s.). For social desirability about a partner, a two (instruction) × two (participants' action) ANOVA revealed that both the main effects of instruction and the participants' action reached significance ($F(1,134) = 22.54$, $p < .01$; $F(1,134) = 18.95$, $p < .01$). There was no significant interaction ($F(1,134) = 0.23$, n.s.).

## 6. Experiment 2

Four experimental conditions were set up, excluding the two conditions in which the actual partner was a human (w/ Human condition) from the six conditions performed in Experiment 1.

### 6.1. Participants

Sixty-six undergraduates participated in the experiment. They were randomly assigned to one of the four conditions, in which the number of participants in each condition was as equal as possible. As a result, the numbers of participants in (1) w/ human (instruction) and w/ C-agent (actual partner), (2) w/ human and w/ D-agent, (3) w/ agent and w/ C-agent, and (4) w/ agent and w/ D-agent were 16, 16, 18, and 16, respectively.

### 6.2. Result

#### 6.2.1. Selection behavior

Fig. 6 shows the participants' selection behavior where the vertical axis indicates the rate of the participants' defect actions, and the horizontal axis indicates each experimental condition. A two (instruction) × two (actual partner) ANOVA revealed that the main effect of instruction reached significance ($F(1,62) = 18.04$, $p < .01$), but neither the main effect of the actual partner nor the interaction was found ($F(1,62) = 0.34$, n.s.; $F(1,62) = 0.49$, n.s.).

Consistent to the result of Experiment 1, the participants selected more defect actions when instructed that their partner was a computer agent.

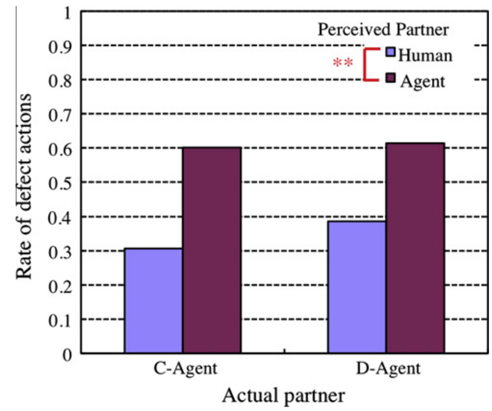


**Fig. 6.** Rate of defect actions in participants' decision-making.

#### 6.2.2. Impressions about the partner caused by the partner's behavior

Fig. 7 shows the participants' impressions about their partner where the vertical axis indicates the average score of the participants' rating, and the horizontal axis indicates each experimental condition. A two (instruction) × two (actual partner) ANOVA revealed that the main effect of an actual partner on individual likeability for a partner reached significance ($F(1,62) = 22.48$, $p < .01$), but the main effect of instruction did not ($F(1,62) = 1.88$, n.s.). The interaction also reached significance ($F(1,62) = 4.74$, $p < .05$). The simple main effect of instruction at w/ D-agent revealed significance ($F(1,62) = 6.30$, $p < .05$), but not at w/ C-agent ($F(1,62) = 0.32$, n.s.). The simple main effect of the actual partner at w/ agent instruction revealed significance ($F(1,62) = 23.93$, $p < .01$), but not at w/ human instruction ($F(1,62) = 3.29$, n.s.). For social desirability for a partner, a two (instruction) × two (actual partner) ANOVA revealed that the main effect of instruction reached significance ($F(1,62) = 9.28$, $p < .01$), but the main effect of the actual partner did not ($F(1,62) = 0.14$, n.s.). The interaction was not significant ($F(1,62) = 1.19$, n.s.).

#### 6.2.3. Impressions of partner caused by the participants' own behavior

Fig. 8 shows the participants' impressions about their partner where the vertical axis indicates the average score of the participants' rating, and the horizontal axis indicates each of the low and high defect groups. A two (instruction) × two (participants' actions) ANOVA revealed that neither the main effect of instruction nor of the participants' action for individual likeability about a partner reached significance ($F(1,62) = 0.86$, n.s.; $F(1,62) = 0.92$, n.s.). There was no significant interaction ($F(1,62) = 1.04$, n.s.). For social desirability about a partner, a two (instruction) × two (participants' action) ANOVA revealed that the main effect of instruction reached significance ($F(1,62) = 10.34$, $p < .01$). The main effect of the participants' action did not reach significance ($F(1,62) = 2.35$, n.s.). The interaction also reached significance ($F(1,62) = 4.31$, $p < .05$). The simple main effect of instruction at the high defect group was significant ($F(1,62) = 14.00$, $p < .01$), but the simple main effect at the low defect group was not found

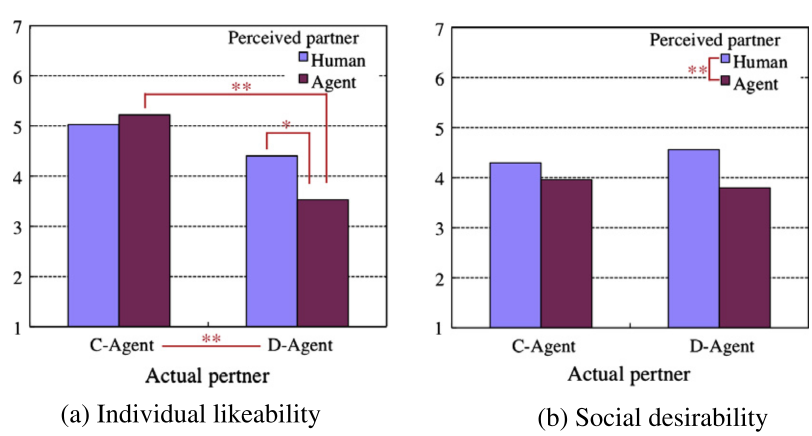


**Fig. 7.** Average score of participants' impressions about partner as a function of experimental setting.

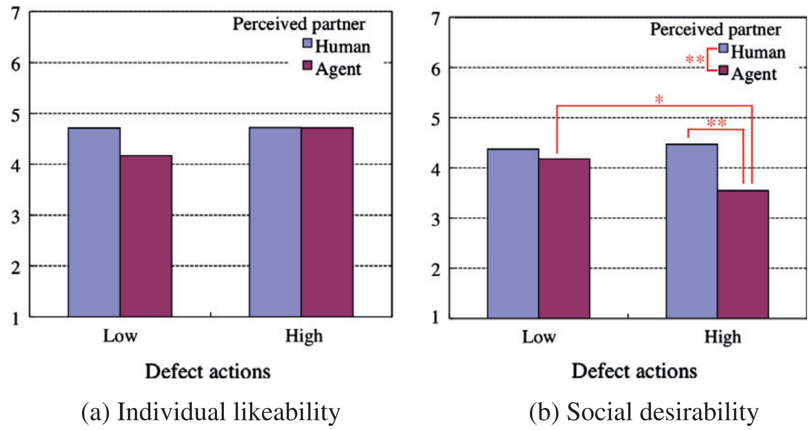


**Fig. 8.** Average score for participants' impressions of partner as a function of participants' selection behavior, i.e., low and high defect actions.

($F(1,62) = 0.65$, n.s.). The simple main effect of the participants' action at w/ agent instruction reached significance ($F(1,62) = 6.52$, $p < .05$), but not at w/ human instruction ($F(1,62) = 0.15$, n.s.).

## 7. Experiment 3

Participants' selection behavior may be crucially influenced depending on whether or not the participants know a partner's personal identity. For example, Lee and Nass (2002) investigated how the decisions of each participant were influenced by decisions of other members of a group. Result showed that the participants followed the opinions of the other members when they were instructed that the other group members were humans than when instructed they were computer agents. Additionally, in this study, such a group conformity effect on participants' decisions was investigated, comparing two situations: the public compliance condition in which each personal opinion was exposed to the other members and the private conformity condition in which each opinion was not exposed. In the former case, the conformity effect differed depending on the instruction, but did not differ in the latter case. The result implies that in our experiments, the instruction effect may disappear when the participant giving the responses is not identifiable, i.e., the identity of who selects the decision is not exposed. Throughout Experiments 1 and 2, the instruction effect was consistently confirmed, but may not be replicated in situations in which personal identity is not revealed. The participants in Experiment 3 did not know who their partner was even though they believed themselves to be interacting with an anonymous person.

### 7.1. Participants

Eighty-three undergraduates participated in the experiment. They were randomly assigned to one of four conditions, in which the number of participants in each condition was as equal as possible. As a result, the numbers of participants in (1) w/ human (instruction) and w/ C-agent (actual partner), (2) w/ human and w/ D-agent, (3) w/ agent and w/ C-agent, and (4) w/ agent and w/ D-agent were 24, 20, 19, and 20, respectively.

### 7.2. Result

#### 7.2.1. Selection behavior

Fig. 9 shows the participants' selection behavior where the vertical axis indicates the rate of the participants' defect actions, and the horizontal axis indicates each experimental condition. A two (instruction) × two (actual partner) ANOVA revealed that the main

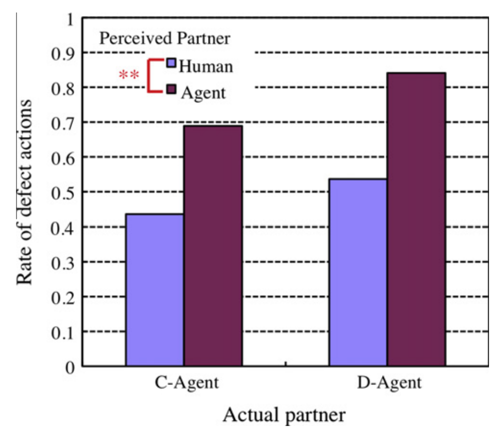


**Fig. 9.** Rate of defect actions in participants' decision-making.

effect of instruction reached significance ($F(1,79) = 19.02$, $p < .01$), but neither the main effect of the actual partner nor the interaction reached significance ($F(1,79) = 3.33$, n.s., $F(1,79) = 0.24$, n.s.). The participants selected more defect actions when informed that their partner was a computer agent.

Consistent to the results of Experiments 1 and 2, the participants selected more defect actions when instructed that their partner was a computer agent.

#### 7.2.2. Impressions of the partner caused by partner's behavior

Fig. 10 shows the participants' impressions of their partner where the vertical axis indicates the average score of the participants' rating, and the horizontal axis indicates each experimental condition. A two (instruction) × two (actual partner) ANOVA revealed that the main effect of the actual partner on individual likeability for a partner reached significance ($F(1,79) = 10.22$, $p < .01$). Neither the main effect of instruction nor the interaction reached significance ($F(1,79) = 0.00$, n.s.; $F(1,79) = 0.02$, n.s.). For social desirability for a partner, a two (instruction) × two (actual partner) ANOVA revealed that the main effect of the actual partner also reached significance ($F(1,79) = 4.26$, $p < .05$). Neither the main effect of instruction nor the interaction reached significance ($F(1,79) = 1.27$, n.s.; $F(1,79) = 0.03$, n.s.).

#### 7.2.3. Impressions of a partner caused by the participants' own behavior

Fig. 11 shows the participants' impressions of a partner where the vertical axis indicates the average score of the participants'

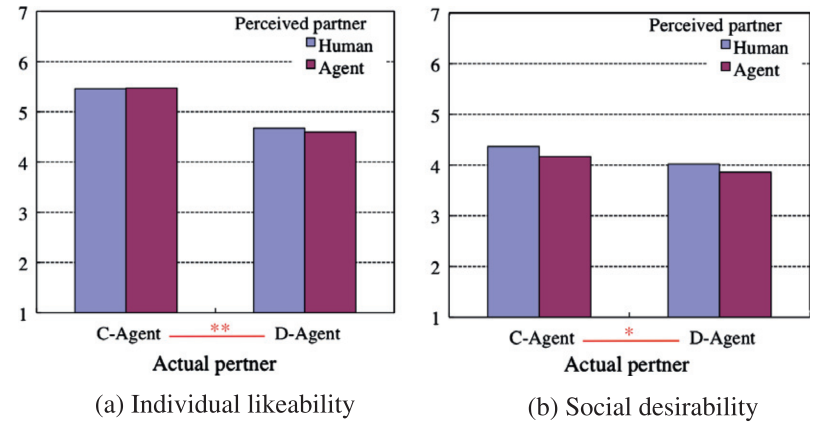


**Fig. 10.** Average score for participants' impressions of partner as a function of experimental setting.

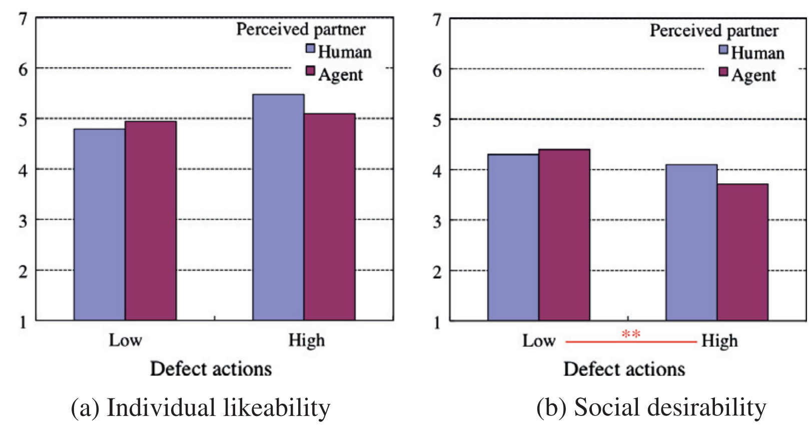


**Fig. 11.** Average score for participants' impressions of partner as a function of participants' selection behavior, i.e., low and high defect actions.

rating, and the horizontal axis indicates each of the low and high defect groups. A two (instruction) × two (participants' action) ANOVA revealed that neither the main effect of instruction nor the participants' action for individual likeability for a partner reached significance ($F(1,79) = 0.12$, n.s.; $F(1,79) = 2.05$, n.s.). There was no significant interaction ($F(1,79) = 0.95$, n.s.). For social desirability for a partner, a two (instruction) × two (participants' action) ANOVA revealed that the main effect of the participants' action reached significance ($F(1,79) = 8.46$, $p < .01$). But neither the main effect of instruction nor the interaction reached significance ($F(1,79) = 0.89$, n.s.; $F(1,79) = 2.42$, n.s.).

## 8. Summary of experiments

One series of human computer interaction studies that have had a large impact on the related academic society are the Media Equation studies (Reeves & Nass, 1996). The research paradigm of the Media Equation studies is as follows (Nass & Moon, 2000). First, researchers focused on experimental situations from past social psychological studies in which the nature of human–human interaction had been investigated. Then they replicate almost the same situations but have participants interact with a computer agent instead. They observe whether the participants' responses to the agent are identical to the responses to a human by analyzing the participants' behavior and evaluating the participants' impressions of their partner. Those studies have generally supported the idea that human beings often relate to computers just as they do to other human beings in a variety of domains such as politeness (Nass, Moon, & Carney, 1999), reciprocity (Fogg & Nass, 1997), personality (Moon & Nass, 1996), in-group/out-group differences (Nass, Fogg, & Moon, 1996), and ethnicity (Nass, Isbister, & Lee, 2001). In the experiments performed based on the Media Equation research paradigm, no control condition as a baseline for comparison was set up. In our experiment, we set up a control condition in which human–human interaction was observed; and directly compared participants' responses to a computer agent with those in the control condition. The Media Equation studies had stressed the equality of interaction with a computer agent to that with a human; however, the results in the direct comparison experiments showed that there were relatively big differences between the responses to a human and to a computer agent.

Table 2 shows the summary of the results of three experiments, indicating the impacts of two aspects of a partner on participants' selection behavior and impressions about the partner. In the table, "O" means a main effect of a focused factor, or simple main effects of a focused factor at every level of the other factor when the interaction reached significance, "o" means a simple main effect of a focused factor at a single level of the other factor when the interaction reached significance, and "X" means no effect. Note that in Experiment 1 we set up three conditions as the factor of the actual partner; but two among the three conditions, the w/ C-agent and w/ D-agent conditions, are only considered for analysis because we focus on the influence of the partner's selection behavior. In the two conditions above, the partner's behavior was systematically manipulated; however, in the w/ Human condition, the behavior of the partner was not controlled.

The overall result shows that for RQ1, referring to participants' selection behavior, the effect of instruction (the representation of a

**Table 2**
Summary of results of three experiments: O shows main effect, o shows simple main effect at a single level, and X shows that no effect was found.

| | Instruction × partner's behavior ANOVA | | Instruction × participant's behavior ANOVA | |
|---|---|---|---|---|
| | Instruction (perceived) | Partner's behavior (actual) | Instruction (perceived) | Participant's behavior |
| *Selective behavior* | | | | |
| Experiment 1 | O | o | – | – |
| Experiment 2 | O | X | – | – |
| Experiment 3 | O | X | – | – |
| *Impressions: Individual desirability* | | | | |
| Experiment 1 | X | O | o | X |
| Experiment 2 | o | o | X | X |
| Experiment 3 | X | O | X | X |
| *Impressions: Social desirability* | | | | |
| Experiment 1 | o | X | O | O |
| Experiment 2 | O | X | o | o |
| Experiment 3 | X | O | X | O |

partner) is dominant rather than the effect of the actual partner's behavior. Therefore, Hypothesis H1 that participants are influenced to a greater degree by the representation of the partner (perceived partner) is confirmed but H2 that participants will be influenced to a greater degree by behavior of the partner (actual behavior) is not confirmed. Only in Experiment 1, the simple main effect of the actual partner factor when instructed that the partner was human reached significance but it did not when instructed that the partner was an agent. However, this instruction effect is generally very strong and therefore is seldom influenced by, and almost independent of, the actual partner's behavior. Therefore, Hypothesis H3 that the influence of one aspect of a partner varies depending on the other aspect of a partner is not confirmed.

In RQ2, the effect of instruction on participants' impressions about a partner is minimal. In six of the twelve cases, there was no effect of instruction on participants' impressions about a partner. The main effect of instruction was only confirmed in two of the twelve cases. This means that in RQ2, Hypothesis H4 is not confirmed. However, in terms of individual likeability, there were consistent effects of the factor of an actual partner's behavior, supporting Hypothesis H5 in RQ2. The more the partner defected against the participants the less the participant felt individual likeability for the partner. On the other hand, for social desirability, there were consistent effects of the participants' own behavior. Participants who selected more defect actions against a partner rated the partner low on social desirability. This means that from the viewpoint of the social desirability Hypothesis H6 is confirmed.

## 9. Discussion and conclusions

### 9.1. Social responses in dilemma situation

The three experiments consistently showed that when instructed that the partner was a computer agent, the rate of defect actions increased. As mentioned in the introduction, in the iterative dilemma game, the altruistic egoism behavior is often observed. The increase of defect actions when instructed to interact with a computer agent implies that the altruistic egoism behavior appears only when a partner is a human; and such behavior is inhibited when a partner is a computer agent. Additionally, this change of selection strategy is brought about by the representation of the partner, rather than the actual behavior of the partner. This change may be due to two reasons. One is that the altruistic behavior is inhibited because the participants may feel distrust for the non-human partner. The participants may fear that the non-human partner might repeat defect responses consistently even if they offer cooperative actions to the partner. The other is the participants' underestimation of the non-human partner's ability to produce adaptive responses. The participants may think that the computer agent may not be equipped with intelligent functions to respond adaptively to the participants' selection. They may hope that a computer agent as a partner may not respond to them with defect actions even if the participants themselves repeat defect actions; they may expect not to fall into mutual distrust.

The TFT strategy is effective for maximizing the benefits under such an uncertain situation. Therefore, we analyzed the rate of the participants' selections made according to the TFT strategy (see Fig. 12). In Experiment 1, a two (instruction) × three (actual partner) ANOVA revealed that the main effect of instruction and an actual partner reached significance ($F(1,139) = 24.99$, $p < .01$; $F(2,139) = 7.04$, $p < .01$). The interaction also reached significance ($F(2,139) = 5.91$, $p < .01$). The simple main effects of instruction at w/ Human and w/C-Agent actual partner ($F(1,139) = 7.26$, $p < .01$; $F(1,139) = 29.23$, $p < .01$) were significant, but the simple main effect at w/ D-Agent actual partner was not found ($F(1,139) = 0.31$, n.s.). The simple main effects of the actual partner at both levels of instruction, w/ Human and w/ Agent, revealed significance ($F(2,139) = 5.11$ $p < .01$; $F(2,139) = 7.84$, $p < .01$), where a Ryan's multiple comparison analysis showed that the rate of usage of TFT was higher at w/Human than at /w D-Agent when instructed the partner was human ($MSe = 160.29$, $p < .05$), and higher at w/ Human and w/ D-Agent than at w/ C-Agent when instructed the partner was an agent ($MSe = 160.29$, $p < .05$). In Experiment 2, the same ANOVA revealed that the main effect of instruction revealed significance ($F(1,62) = 4.24$, $p < .05$), but the main effect of an actual partner did not ($F(1,62) = 1.95$, n.s.). The interaction also reached significance ($F(1,62) = 6.97$, $p < .05$). The simple main effect of instruction at w/ C-Agent revealed significance ($F(1,62) = 11.05$, $p < .01$), but not at w/ D-Agent ($F(1,62) = 0.17$, n.s.). The simple main effect of an actual partner at w/ Agent instruction revealed significance ($F(1,62) = 8.15$, $p < .01$), but not at w/ Human instruction ($F(1,62) = 0.78$, n.s.). In Experiment 3, the same ANOVA revealed that the main effect of instruction revealed significance ($F(1,79) = 6.60$, $p < .05$), but the main effect of an actual partner did not ($F(1,79) = 0.22$, n.s.). The interaction also reached significance ($F(1,79) = 6.48$, $p < .05$). The simple main effect of instruction at w/ C-Agent revealed significance ($F(1,79) = 13.07$, $p < .01$), but not at w/ D-Agent ($F(1,79) = 0.00$, n.s.). The simple main effect of an actual partner at w/ Agent instruction revealed significance ($F(1,79) = 4.53$, $p < .05$), but not at w/ Human instruction ($F(1,79) = 2.17$, n.s.).

The three experiments consistently showed that when instructed that the partner was a human, the change of actual partner's behavior across the C-agent and D-agent conditions did not influence the rate of the participants' TFT usage. On the other hand, when instructed to interact with a computer agent, the rate of TFT usage decreased in the C-agent condition compared to in the

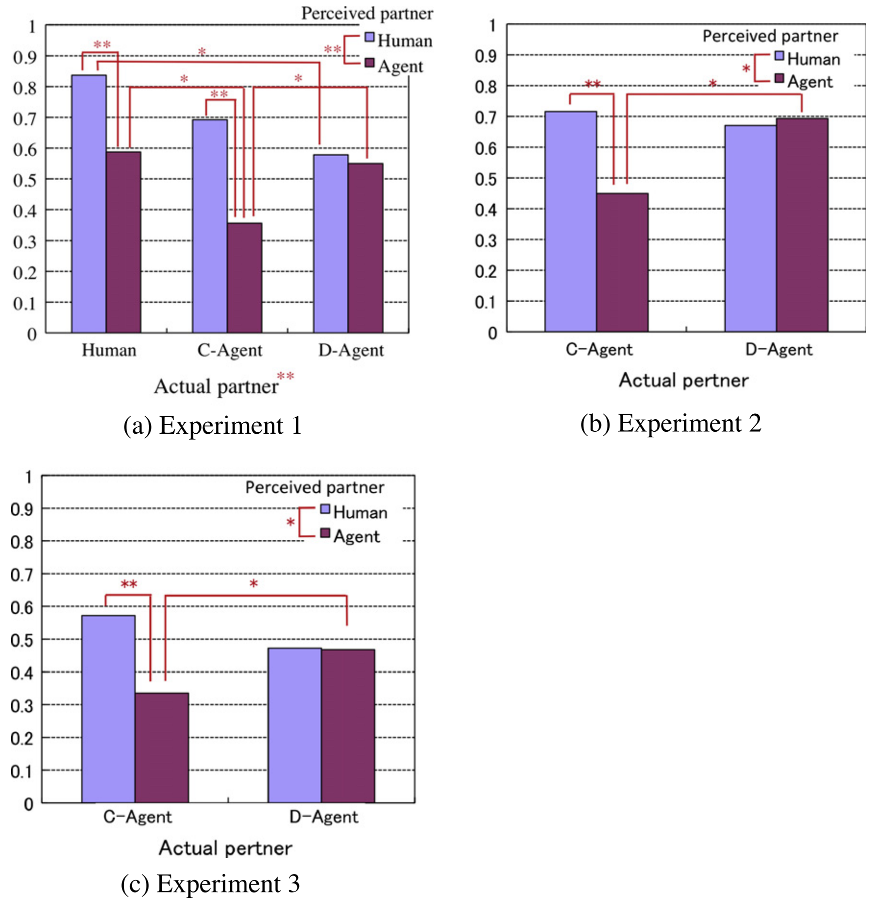


**Fig. 12.** Rate of participants' usage of TFT strategy for selection.

D-agent condition. This happened because the participants repeatedly responded with defect actions when the partner continuously made cooperative actions in the C-agent condition. The above analysis implies that the altruistic egoism strategy emerges only in human–human interaction; but in human–agent interaction participants sometimes repeat defect actions after noticing that the partner does not respond to their defect actions with defect actions.

Basically, when the participants believed they were interacting with a human partner, they tended to select more cooperate actions even in a one-shot dilemma game. For example, Kiyonari et al. studied a one-shot Prisoner's dilemma game (Kiyonari, Tanida, & Yamagishi, 2000) where the optimal (points-maximizing) strategy was simply defection. The situation was very realistic because participants were given real money, and the amount of which was based on the one time selection. Moreover, each of the pair of participants did not face the other. The result showed that even in such a situation, people generally prefer cooperative actions to defect actions. This finding suggests that people prefer cooperation to defection; it is a very strong and fundamental nature of human behavior when working with other humans. On the other hand, in our experiment, the simple instruction that a partner is a computer agent caused our participants largely shift their behavior to offer defect actions.

### 9.2. Partner impressions

Participants' impressions of likeability were higher when they related to a partner who selected fewer defect actions. Their impressions of social desirability were influenced more by the participants' own behavior than by the partner's behavior. The participants who selected many defect actions against a partner rated social desirability of the partner lower. Kelley and Stahelski (1970) pointed out that cognition about others substantially differs between cooperative persons who prefer cooperative actions and non-cooperative persons who prefer defect actions. They confirmed that the latter persons tend to think that the partner is also a non-cooperative person. Similarly, in our experiments, the non-cooperative persons who made many defect actions rated the partner's social desirability lower.

This characteristic may also be interpreted as a kind of cognitive dissonance behavior. The participants may justify their defect actions by reducing a partner's social desirability (Festinger, 1957), canceling their inner cognitive dissonance. We should note that this tendency was confirmed not to depend on an experimenter's instruction. Participants' own behavior consistently affected their impressions of social desirability about the partner in both w/ human and w/ agent instruction conditions, meaning that the participants tried to justify their defect actions not only when interacting with a human but also when interacting with a computer agent.

### 9.3. Nature of experimental tasks

In our experiments, the Prisoner's dilemma game was used as an experimental task. Are the findings observed in the current experiments also confirmed when other tasks are used? The experiments in Miwa, Baba, and Terai (2005) and Miwa and Terai (2006) were performed based on the same research paradigm as in the current study, but using a rule discovery task, Wason's 2-4-6 task, which is widely used in laboratory studies on human scientific discovery (Wason, 1960). In the study, hypothesis-testing behavior was investigated. It is well known that human hypothesis-testing has a positive test bias, a strong tendency that positive instances consistent with a formed hypothesis are employed for testing the hypothesis (Klayman & Ha, 1987).

The first experiment showed that the participants' hypothesis testing behavior is largely influenced by an actual partner's behavior but not by instruction. The rate of using positive instances in hypothesis-testing was consistent regardless of whether they believe their partner is a human or a computer agent. The second experiment dealt with reciprocity behavior, one of the representative principles widely observed in human behavior: people give a lot of information to a partner when they receive a lot of information from the partner. The statistical analysis showed that on contribution to a partner, the main effect of instruction did not appear but the interaction of instruction (human or computer agent) and the amount of received information reached significance. When instructed that their partner was a human, the participants gave a lot of information about a hypothesis to a partner from whom they received a lot of information. On the contrary, when instructed that their interaction was with a computer agent, this tendency disappeared. In the current experiment with the Prisoner's dilemma game, a strong main effect of instruction emerged in participants' selection behavior.

It seems that the degree of sociality behind the participants' behavior increases in the order of hypothesis-testing, reciprocity behavior in problem solving tasks such as Wason's 2-4-6 task, and selection behavior in the dilemma game in which complex factors of social interaction influence the participants' behavior. The above results indicate that the differences in responses to humans and computer agents become salient in situations where sociality is largely required.

## References


Aharoni, E., & Fridlund, A. J. (2007). Social reactions toward people vs. computers: How mere labels shape interactions. *Computers in Human Behavior, 23*, 2175–2189.

Axelrod, R. (1980a). Effective choice in the Prisoner's dilemma. *Journal of Conflict Resolution, 24*, 3–25.

Axelrod, R. (1980b). More effective choice in the Prisoner's dilemma. *Journal of Conflict Resolution, 24*, 379–403.

Burgoon, J. K., Bonito, J. A., Bengtsson, B., Cederberg, C., Lundeberg, M., & Allspach, L. (2000). Interactivity in human–computer interaction: A study of credibility, understanding, and influence. *Computers in Human Behavior, 16*, 553–574.

Cantor, N., & Mischel, W. (1977). Traits as prototypes: Effects on recognition memory. *Journal of Personality and Social Psychology, 35*, 38–48.

Cohen, C. E. (1981). Person categories and social perception: Testing some boundaries of the processing effects of prior knowledge. *Journal of Personality and Social Psychology, 40*, 441–452.

Dahlbak, N., Jonsson, A., & Ahrenberg, L. (1993). Wizard of oz studies: Why and how. *Knowledge-Based Systems, 6*, 258–266.

Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology, 31*, 169–193.

Dehn, D. M. (2000). The impact of animated interface agents: A review of empirical research. *International Journal of Human–Computer Studies, 52*, 1–22.

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

Festinger, L. (1957). *A theory of cognitive dissonance*. Standard University Press.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York: McGraw-Hill.

Fogg, B., & Nass, C. (1997). *Do users reciprocate to computers? Proceedings of the CHI conference (Atlanta, GA)*. New York: Association of Computing Machinery.

Gong, L. (2008). How social is social responses to computers? The function of the degree of anthropomorphism in computer representations. *Computers in Human Behavior, 24*, 1494–1509.

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

Hayashi, F. (1978). *The fundamental dimensions of interpersonal cognitive structure*. Bulletin of School of Education, Nagoya University (Vol. 25, pp. 233–247) [in Japanese].

Kelley, H. H., & Stahelski, A. J. (1970). Social interaction basis of cooperators' and competitors' beliefs about others. *Journal of Personality and Social Psychology, 21*, 190–197.

Kiesler, S., Waters, K., & Sproull, L. (1996). A Prisonor's dilemma experiment on cooperation with people and human-like computers. *Journal of Personality and Social Psychology, 70*, 47–65.

Kiyonari, T., Tanida, S., & Yamagishi, T. (2000). Social exchange and reciprocity: Confusion or a heuristic? *Evolution and Human Behavior, 21*, 411–427.

Klayman, J., & Ha, Y.-W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review, 94*, 211–228.

Lee, E.-J., & Nass, C. (2002). Experimental tests of normative group influence and representation effects in computer-mediated communication: When interacting via computers differs from interacting with computers. *Human Communication Research, 28*, 349–381.

Messick, D. M., & Brewer, M. B. (1983). *Review of personality and social psychology*(Vol. 4, pp. 11–44).

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. In *Proceedings of 28th annual meeting of the cognitive science society* (pp. 597–602).

Miwa, K., Baba, Y., & Terai, H. (2005). An experimental study on collaborative scientific activities with an actual/imaginary partner. In *Proceedings of the 6th international conference of computer supported cooperative, learning* (pp. 433–439).

Mol, L., Krahmer, E., Maes, A., & Swerts, M. (2009). The communicative import of gestures. *Gesture, 9*, 97–126.

Moon, Y., & Nass, C. (1996). How real are computer personalities? Psychological responses to personality types in human–computer interaction. *Communication Research, 23*, 651–674.

Nass, C., Fogg, B., & Moon, Y. (1996). Can computers be teammates? *International Journal of Human–Computer Studies, 45*, 669–678.

Nass, C., Isbister, K., & Lee, E. (2001). Truth is beauty: Researching embodied conversational agents. In J. Castells (Ed.), *Embodied conversational agents*. Cambridge, MA: MIT Press.

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56*, 81–103.

Nass, C., Moon, Y., & Carney, P. (1999). Are respondents polite to computers? Social desirability and direct responses to computers. *Journal of Applied Social Psychology, 29*, 1093–1110.

Parise, S., Kiesler, S., Sproull, L., & Waters, K. (1999). Cooperating with life-like interface agents. *Computers in Human Behavior, 15*, 123–142.

Pruitt, D. G., & Kimmel, M. J. (1977). Twenty years of experimental gaming: Critique, synthesis, and suggestions for the future. *Annual Review of Psychology, 28*, 363–392.

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. CSLI Publications.

Sundar, S., & Nass, C. (2000). Source orientation in human computer interaction: Programmer, networker, or independent social actor? *Communication Research, 27*, 683–703.

Taylor, M. (1987). *The possibility of cooperation*. Cambridge, Mass.: Cambridge University Press.

Wason, P. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology, 12*, 129–140.

Yamamoto, Y., Matsui, T., Hiraki, K., Umeda, S., & Anzai, Y. (1994). Interaction with a computer system: A study of factors for pleasant interaction. *Cognitive Studies, 1*, 107–120 [in Japanese].