IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 13 March 2023, accepted 3 April 2023, date of publication 12 April 2023, date of current version 17 April 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3266397

RESEARCH ARTICLE

Unleashing Fairness: How a Group Norm-Aware Agent Shakes Up the Ultimatum Game

YOTARO FUSE[1], (Member, IEEE), BIINA ASHIDA[2], EMMANUEL AYEDOUN[2], (Member, IEEE), AND MASATAKA TOKUMARU[2], (Member, IEEE)

[1]Department of Intelligent Robotics, Toyama Prefectural University, Imizu, Toyama 939-0398, Japan
[2]Department of Electrical, Electronic and Information Engineering, Kansai University, Suita, Osaka 564-8680, Japan

Corresponding author: Yotaro Fuse (fuse@pu-toyama.ac.jp)

This work was supported in part by the Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant JP21K12099.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Review Department for Research on Humans with Toyama Prefectural University under Approval No. R4-6, Dated November 29, 2022.

ABSTRACT The emergence of social robots has created new opportunities for them to coexist with humans, but only if they can imitate human-like social behavior. Previous research has examined various robot social cues, such as emotions, gestures, and eye contact. However, an area that has been under-researched is the concept of implicit group norms, which are unwritten rules dictating the expected behavior of group members and can differ across different groups. By improving the ability of robots to behave in expected ways, we hope to considerable greater acceptance of robots among humans. In this study, we propose a group norm-aware decision-making model to help robots adapt to group norms, which we evaluated in a human–agent experiment based on the ultimatum game. In this scenario, the gains and losses of one group member affect everyone else. Our results demonstrate that a group norm-aware decision-making agent promotes fairer distributions of benefits among group members, enhancing mutual benefit compared with an agent that does not consider group norms. This study provides a solid foundation for further research in developing social robots that are more adaptable and acceptable to humans. Additionally, our proposed model sets the stage for future robot experiments, ultimately leading to the emergence of more equitable and empathetic human–robot interactions.

INDEX TERMS Fairness, group norm, human–agent interaction, human–robot interaction, interactive reinforcement learning, social robot, ultimatum game.

I. INTRODUCTION

The growing interest in exploring how robots can best interact with humans as trusted assistants and partners has led to extensive research [1], [2]. Human acceptance of robots is crucial for this symbiosis to flourish [3]. Thus, social robots must exhibit behavior that adapts to human personalities and characteristics. In response, researchers have developed models that allow computational systems and social robots to learn and adjust to human interactions, creating a more natural and engaging experience [4], [5].

The associate editor coordinating the review of this manuscript and approving it for publication was Arianna Dulizia.

The nature of living in a group is inherently social. In a human group, individuals often anticipate certain behavioral expectations from others, even without direct interaction or instruction [6], where ‘‘group norms’’ come into play, as they are shared expectations within a group. In human society, the expected behavior in a given situation is often implicitly shared without explicit communication. Conforming to group norms can serve as a criterion for evaluating whether someone is suitable for group membership or as a peer.

Our research has focused on developing agents and robots capable of adapting to implicit group norms, even without direct interactions [7], [8]. Our goal is to create robots that can conform to the unwritten rules of human–robot groups, thus enhancing predictability and familiarity in human–robot interactions. This approach is critical to building a more symbiotic relationship between humans and robots in the future. By improving the ability of robots to behave as expected, we hope to promote greater acceptance of robots among humans.

It is important to note that our previous studies [7], [8], focused on evaluating agents and robots in relatively simple experimental scenarios without considering situations that involve conflicting interests within groups and direct interactions between group members. In real-world situations, group members may prioritize their benefits even if it means disadvantaging others. However, generally, all individuals desire fairness and resent any discrepancy between their benefits and those of others [9]. In addition, members of human groups often provide direct feedback or responses to each other's actions, further emphasizing the need to validate our proposed model in scenarios that involve both immediate responses and individual interests. Nevertheless, as previously mentioned, our previous study evaluated the model in a controlled experimental situation that did not incorporate interests and direct reactions. To conduct more realistic evaluation experiments than those in our previous studies, we must validate our proposed model in scenarios that involve direct responses and individual interests.

In group settings with individual interests, each member desires fair treatment from others while also prioritizing their gains and feeling averse to their losses. Based on this, two general statements can be derived:

- Extreme altruistic behavior can harm one's benefit and is not considered fair.
- Extreme selfish behavior can harm the benefits of others and is not considered fair.

Therefore, decision-making that mixes altruism and selfishness is required to behave humanly and reasonably in a group. It is crucial to balance acting altruistically by showing consideration for the benefits of others and acting selfishly to protect one's advantage. In this study, the agent must behave as a group member in more complex scenarios than those in our previous studies.

This study requires the agent to exhibit group behavior in more complex scenarios than those in our previous studies. For instance, in one of our earlier studies [10], we merely observed agents' behavior using the proposed model in complex scenarios where the mentioned behavior was required. In contrast, the current study evaluates the agent's ability to adjust to implicit group norms within groups and investigates how the agent's decisions influence human group members. We propose using the ''group ultimatum game'' (GUG) as an experimental setting to assess the agent's ability to act within a group, considering their interests and the feedback from their actions. This game is a modified version of the well-known ultimatum game [11].

The GUG has a structure that restricts each player's ability to observe all other players' behavior. However, our group norm-aware decision-making model assumes that all behavior within the group is observable. This study investigates the impact of the agent's adaptive behavior on group norms on their decision-making process, recognizing that agents have the advantage of observing all behavior within a group before making a decision, unlike human players.

The paper is structured as follows: Section II provides an overview of related research on the acceptance of robots by humans and human–robot interaction. Specifically, this section explores topics such as humans' conformity to robots and the ultimatum game in human–robot groups. Section III outlines the experimental scenario used in this study, namely, the group ultimatum game (GUG). Section IV introduces our proposed model for an agent to make decisions that are aware of group norms. Section V describes the experiment in detail, which consisted of the GUG being conducted twice, followed by a discussion of the behavior of agents and human participants in the GUG scenario. Finally, Section VI presents the conclusions drawn from the study and suggestions for future research.

II. RELATED STUDIES

In the human–robot interaction field, socializing with robots is crucial for their acceptance by humans. Previous research suggests that the user's personality and the robot's familiarity with the user can impact acceptance [13]. Other studies have investigated the effects of robot features on elderly care and service robots [14], [15]. Although these studies focused on elderly people, it is hypothesized that robots' social competencies could positively influence their acceptance by humans. Additionally, Correia et al. have discussed the elements necessary for robots to improve the sense of unity and group cohesiveness of human–robot groups [16]. They explored the social aspects that allow humans in a group to consider robots and agents as group members, identifying norms and roles as essential components of group cohesion and necessary for robot socialization.

Overall, this study proposes and assesses a group norm-aware decision-making model as a method of robot socialization, considering previous research on human acceptance of robots and their use as team members. Although emotions, gestures, and eye contact have been extensively studied, group norms have received comparatively less attention.

A. GROUP NORM AND CONFORMITY

Humans have varying criteria for decision-making [20], [21] and respond differently to one another under the same conditions. Additionally, individuals within a group can be affected by social factors, which refer to changes in an individual's thoughts, feelings, attitudes, or behavior that occur because of interactions with other individuals or groups [22]. Typically, people imitate and conform to other people's behavior when they do not know how to act in an unfamiliar situation. As defined by [23], conformity is the change in one's behavior to match the responses of others. In other words, humans tend to respond to the conduct of others.

In a study by Sherif et al., participants in a group attempted to answer vague quiz questions, and the results showed the formation of a group norm through several interactions [24]. It was assumed that the influence of each participant in the group facilitated the imitation of their response by the other participants [25], thereby forming a group norm about the quiz. However, some studies have highlighted the impact of group norms on social influence within human groups, with some experiments aiming to demonstrate this social phenomenon [26], [27], [28]. For instance, Asch et al. investigated social pressure from a majority group [19]. They examined whether innocent participants would conform to the behavior of a majority group, even if it is clearly incorrect.

In addition to Asch-based experiments, several studies have investigated social influence in human–robot groups. Establishing a positive rapport with robots while maintaining appropriate boundaries in their interactions is vital for humans in such groups. Salomons et al. showed that the presence of a robot induces social pressure, leading to opinion changes among some individuals in human–robot groups [17]. Meanwhile, Brandstette et al. observed conformity in some scenarios involving human–robot groups [18]. Williams et al. and Vollmer et al. reported that some children changed their opinions or behavior because of mechanical or robotic behavior [29], [30]. However, Beckner et al. showed that humans did not conform to humanoid robots when performing tasks related to the boundaries of linguistic imitation [31]. Riva et al. investigated whether individuals faced with different tasks (objective vs. subjective) were more affected by the information provided by another human or an unembodied AI like a virtual assistant [32]. Their experimental results suggested that participants were more likely to exhibit conformity under the influence of AI rather than from other humans, particularly in objective tasks, as opposed to subjective ones. Masjutin et al. studied the tendency of participants to conform to hybrid majorities consisting of both humans and robots instead of solely focusing on humans or robots. Their results suggest that conformity frequently occurs in such hybrid groups [33].

In summary, the degree of human conformity to agents and the level of social influence that agents exert on humans can differ based on various factors. Prior research has predominantly examined human reactions to social agents. This study evaluates a decision-making model that can adjust to group norms and investigate how such decisions affect the group collectively.

B. HUMAN–ROBOT ULTIMATUM GAME

We developed a group ultimatum game scenario (GUG) based on a well-known ultimatum game (see Section III). In the ultimatum game [11], two players are designated as the "proposer" and the "responder." The proposer is given a sum of money by the experimenter and must decide how to divide it between themselves and the responder. If the responder accepts the offer, the money is divided based on the proposal. However, if the responder rejects the offer, both players receive nothing. Even after accounting for cultural differences, Oosterbeek et al. found that, on average, proposers tended to offer the responder 40% of the total amount received [12].

Previous studies in the social robotics field have explored the ultimatum game scenario involving human–robotic interactions, specifically examining human response to robots. Sandoval et al. conducted two experiments: One involving a human–agent and an experimental participant and the other involving a robot agent and an experimental participant [34]. They investigated participants' behavior toward robots and humans in the ultimatum game and their likeability toward the robots. Moreover, Sandoval et al. demonstrated in a comparative study between different robot strategies that human participants have varying preferences for robots depending on the robot's strategy [35]. Law et al. observed human emotional responses under study to the human–robot ultimatum game involving a non-humanoid robot instead of the humanoid or animal-shaped robots used in other studies [36]. They discovered that when robots displayed non-social behavior, such as rejecting the participant's offer, people tended to anthropomorphize the robots more. Dio et al. focused on preschoolers' behavior toward robots in ultimatum games [37] and observed that children treated the robot as they would treat other children.

Several studies have investigated human physiological responses in human–robot ultimatum games. For instance, Mitjans et al. examined the behavioral and physiological responses of participants who acted as proposers in an ultimatum game involving a human, a computer, a virtual avatar, and a robot as respondents [38]. They observed a noticeable increase in GSR values a few seconds after proposing an offer in some game settings with a social robot. Fukuda et al. conducted EEG measurements in the ultimatum game scenario, finding that when a robot touched human participants, it increased their tolerance for unfair offers from the robot, indicating a positive effect [39]. Previous studies have focused on human reactions to robots in the one-to-one ultimatum game scenario but did not investigate how robots make decisions or consider the influence of group dynamics. Moreover, these studies did not involve group interactions. In contrast, our study investigates the proposed decision-making model for robots in the group ultimatum game scenario and examines the influence of the robot's decisions on the group including humans.

Additionally, previous research in the evolutionary game theory field has investigated multiplayer ultimatum games played in groups comprising a proposer and a group of responders, as observed in studies conducted by Takesue et al. and Santos et al. [40], [41]. Some of these studies employed groups of humans or agents similar to our current study. However, these studies differ in the game structure, as they involve a single proposer and a group of responders. Additionally, these studies primarily focus on the long-term evolution of strategy selection through multi-agent

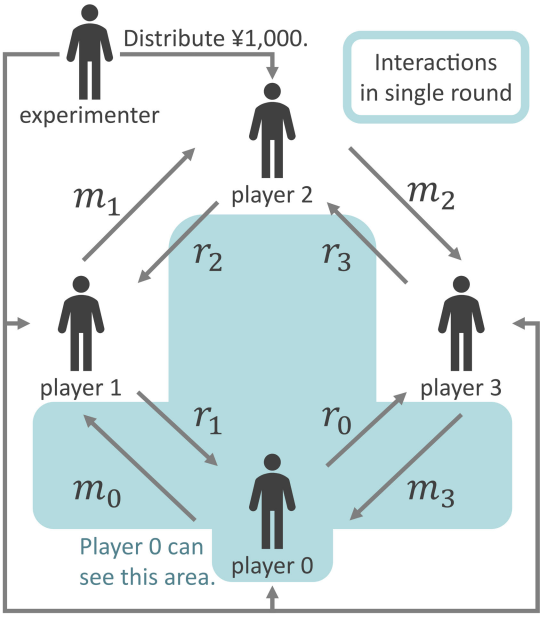


FIGURE 1. Diagram illustrating the sequence of interactions in a round of the group ultimatum game. At the beginning of the iteration, the experimenter was given 1,000 yen for each player. Players then divide this amount between themselves and another player of their choice and pay an amount of m_x to the selected player. The receiving player can either accept or reject the offer. At the end of the round, the players can see the results. The experiment comprised 20 repeated rounds.

simulations rather than the decision-making process of agents and interactions between humans and agents, which is the main focus of our study.

III. GAME SCENARIO

The current study investigates how players behave in a group ultimatum game scenario where each player seeks to maximize their benefit and avoid losses while interacting with other group members. In this section, we describe the group ultimatum game and introduce an interface for playing the game.

A. GROUP ULTIMATUM GAME

Fig. 1 illustrates the players and their actions at a round of the group ultimatum game, including the player identities $p_x(x = 0, 1, 2, 3)$, the amount of money offered by each player $x(m_x)$, and player p_x's response (ACCEPT or REJECT) r_x.

The players' actions involve offering money to a specific player and responding to the proposed amounts from other players by either accepting or rejecting them. As a result, each player plays two unique roles during the game: one as the proposer who offers a specific amount of money and the other as the receiver who either accepts or rejects the proposed amount. The experiment comprised 20 repeated rounds. The following steps were undertaken in each iteration:

1) The experimenter distributes 1,000 yen to each player.
2) Each player gives m_{give} yen to the player on their left from the 1,000 yen they were given. The player keeps the remaining amount of yen, which is equal to $1000 - m_{\text{give}}$ yen. See Fig. 3.
3) Each player accepts or rejects the amount m_{given} offered by the player on their right. See Fig. 4.
4) All players see the responses of others at the end of the round.
   a) If all players accept the offers, they will receive $1000 - m_{\text{give}} + m_{\text{given}}$ yen. See Fig. 5.
   b) If any player rejects the offer, they will receive nothing. See Fig. 6.

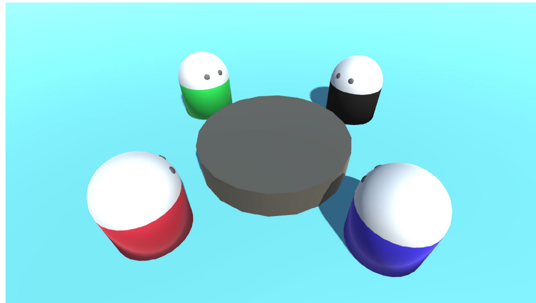


FIGURE 2. Entire game environment of group ultimatum game.

Players do not have access to all interactions that occurred in a round. For example, as shown in the blue-colored area in Fig. 1, player 0 only observes $\{m_0, m_3, r_0, r_1, r_2, r_3\}$, except for offers m_1 and m_2. In other terms, players are not privy to the amount of money offered in an interaction in which they were not directly involved.

B. INTERFACE

Fig. 2 presents the virtual game environment, which shows a gray circular table surrounded by four robots. Each player was situated at the viewpoint of one of the robots and controlled it using a game controller during the game. The robots on either side of the game environment directly interacted with the player. Figs. 3 to 8 show examples of the interface shown on a player's screen. The cursor in Figs. 3 and 4, represented by a red dot, was controlled by a player using the game controller's stick. The amount of money each player currently possessed and the current round is always displayed in the lower-left corner of the screens.

The interface used by the player to input their offer amount is illustrated in Fig. 3. The amount of money offered is entered by pressing the "0" to "9" buttons on the game controller, and the input is then confirmed by pressing the "Enter" button. If needed, the player can delete their input by pressing the "Del." button. The player's input screen for accepting or rejecting an offer is shown in Fig. 4. The corresponding button, either "ACCEPT" or "REJECT," can be pressed by the player to make their choice. The figures depicted in Figs. 6 and 5 illustrate the display presented to the player at the end of each round, indicating the outcome of the players' interactions. Fig. 6 shows the result when any player rejects the offer, while Fig. 5 displays the result when all offers are accepted. Upon viewing the game result, the player presses the "Proceed to next round" button. In the

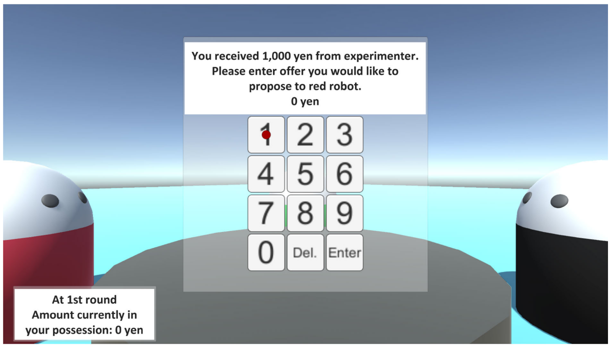


FIGURE 3. Player's input screen for offering amount of money.

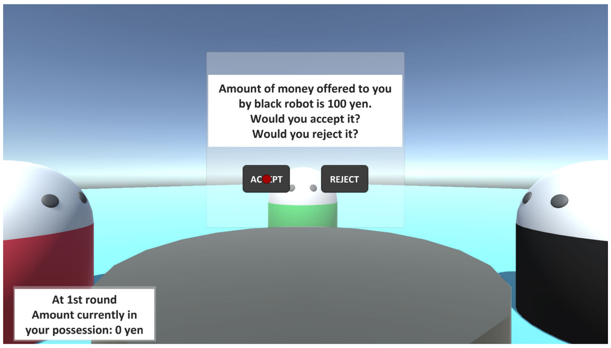


FIGURE 4. Player's input screen for responding.

final round, the total amount of money each player earned is shown on the player's screen, as illustrated in Fig. 7. They saw the screen after being informed a game result at the round 20. As depicted in Fig. 8, players were presented with a waiting screen while waiting for other players' inputs. A round consisted of three steps: distribution, response, and result. To proceed to the next step, all players had to complete their input in the current process. If some players finished their input before the others, they were required to wait until all players had finished their input.

Each participant engaged in the game in a face-to-face setting with the experimenter. The participants were not allowed to see the game screens of the other players and were not informed about which colored robot each human player was controlling. They were also informed that the experimenter would constantly monitor their interactions.

IV. METHOD

In this experiment, a computer agent was included as a participant in the group ultimatum game, along with three human players. We used two types of agents, one of which made decisions based on the proposed decision-making model aimed at adapting to group norms. The other agent served as a comparison and made decisions that did not conform to group norms. The proposed decision-making model was based on a reinforcement learning framework that allowed the expression of adaptive behavior through the renewal of value functions.

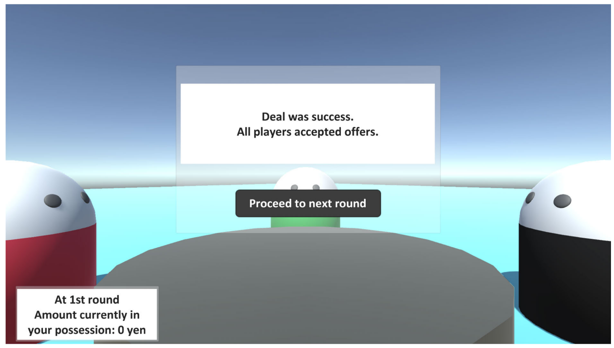


FIGURE 5. Player's screen exhibiting game result at round. If all players exhibited ACCEPT, players get money.

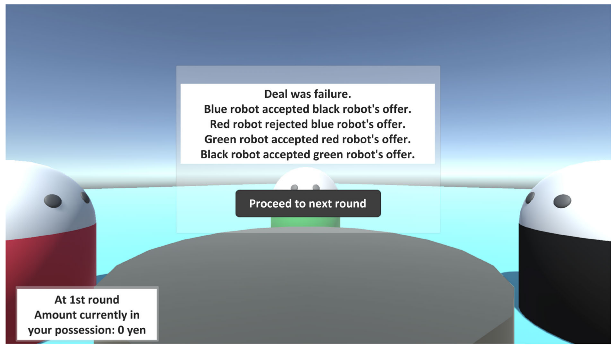


FIGURE 6. Player's screen exhibiting game result at round. If anyone exhibited REJECT, players do not get money.

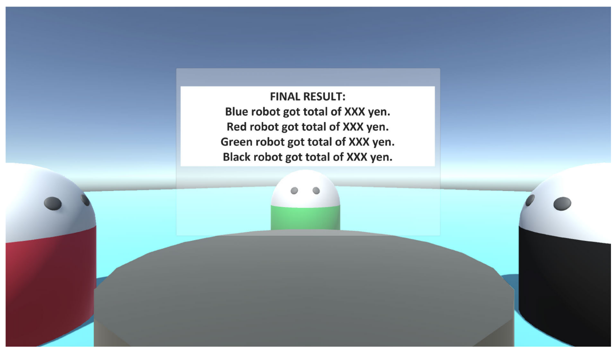


FIGURE 7. Player's screen exhibiting total scores at final round.

In the ultimatum game, each player p was designated by an element of the set $\{1, 2, 3, a\}$, where a represents the computer agent, and the other three parts correspond to the human players. The amount of money offered by player p was denoted by m_p, and the response to the offer (either ACCEPT or REJECT) was represented by r_p, which takes the value of 0 or 1, respectively, as shown in (1).

$$r_p = \begin{cases} 0 & \text{if ACCEPT} \\ 1 & \text{if REJECT} \end{cases} \tag{1}$$

In contrast to the human players who were unable to observe all interactions in a round, the agent in this experiment could observe all interactions of the players, including

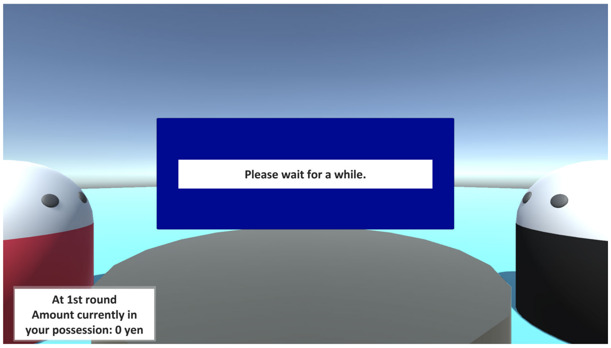


FIGURE 8. **Player's screen while waiting for other players inputs.**

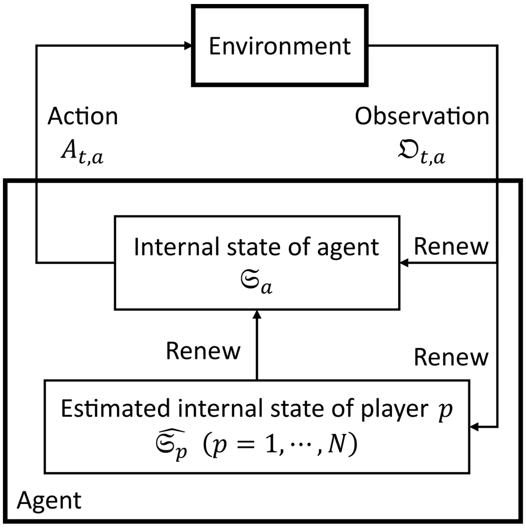


FIGURE 9. **Method of group norm-aware decision-making at round *t*.**

those outside the observable area of the human players (see the observable area of player 0 in Fig. 1). The agent observed all offers m_p and responses r_p, and exhibited its own actions $m_{p=a}$ and $r_{p=a}$. Therefore, this experiment evaluated the agent's ability to adapt to group norms in the group ultimatum game while having access to all interaction details.

A. ENVIRONMENT AND AGENT

Fig. 9 depicts an overview of the decision-making process based on the proposed model in round t. In each round, the agent using the proposed model followed this sequence:

1) Exhibit the agent's action based on internal state.
2) Observe others' actions.
3) Renew internal states.

In Fig. 9, the environment is external to the agent and refers to the virtual game environment and the other players. The agent outputs actions $A_{t,a}$ to the environment and inputs observations $\mathfrak{O}_{t,a}$ from the environment. As shown in (2) and (3), the action $A_{t,a}$ is a set of its actions $m_{t,p=a}$ and $r_{t,p=a}$, and the observation $\mathfrak{O}_{t,a}$ is a collection of all players' actions including the agent itself. The number of other players N in games was three.

$$A_{t,a} = \{m_{t,a}, r_{t,a}\} \tag{2}$$

$$\mathfrak{O}_{t,a} = \{A_{t,p}\} \quad (p = 1, \ldots, N, a) \tag{3}$$

Based on the observed information $\mathfrak{O}_{t,a}$, the agent renewed its internal state $\mathfrak{S}_a$ along with the internal state of the player p estimated by the agent $\widehat{\mathfrak{S}_p}$ (estimated internal state of player p). Additionally, the agent renewed its internal state $\mathfrak{S}_a$ based on the estimated internal states of others $\widehat{\mathfrak{S}_p}$. This allowed the agent to reflect the results of the estimations of others in its subsequent actions. In the next round $t+1$, the agent output action $A_{t+1,a}$ to the environment based on the renewed internal state $\mathfrak{S}_a$.

B. ACTION

The agent's action $A_{t,a}$ was based on the value functions $\mathrm{V}_{\mathrm{offr}}$ and $\mathrm{V}_{\mathrm{diff}}$, shown in (4) and (5). $m_{t,\to a}$ in (5) refers to the amount of money offered to the agent, and $m_{t,\to a} - m_{t,a}$ represents the difference between the amount offered by a player to the agent and the amount the agent itself offers to another player. The agent decided to accept or reject the offer based on whether this difference was deemed acceptable or not acceptable.

$$m_{t,a} = \arg\max \mathrm{V}_{\mathrm{offr}} \tag{4}$$

$$r_{t,a} = \begin{cases} 0 & \text{if ACCEPT} \\ & (\arg\max \mathrm{V}_{\mathrm{diff}} \leq m_{t,\to a} - m_{t,a}) \\ 1 & \text{if REJECT} \\ & (\arg\max \mathrm{V}_{\mathrm{diff}} > m_{t,\to a} - m_{t,a}) \end{cases} \tag{5}$$

The value function V assigns a set of values to a set of actions that can be performed or selected in a specific group, based on the expected outcome or reward associated with each action. Agents used value functions as a strategy to estimate adaptive actions that aligned with group norms. In this study, a high value assigned by the value function to a particular action meant that the agent estimated that the action was more adaptive to the group norm. The notation arg max V refers to the action that was estimated to have the highest value among all possible actions in the given context and was deemed to be the most adaptive in the group. The agent attempted to estimate the actions that were more adaptive to the group norm by continuously renewing the value function based on its observations of the game environment $\mathfrak{O}_{t,a}$.

In this experiment, the value function $\mathrm{V}_{\mathrm{offr}}(m_o)$ outputs the in-group value of proposing a certain amount of money m_o to the agent's counterpart. The value function $\mathrm{V}_{\mathrm{diff}}(m_d)$ outputs the in-group value of an infimum of the acceptable difference infAD between the agent's offer amount m_a and the amount offered to the agent $m_{\to a}$. In other words, if the infimum of an acceptable difference infAD is set to -100 ($\arg\max \mathrm{V}_{\mathrm{diff}}(m_d) = -100$), the agent will reject an offer $m_{\to a}$ if the difference between the offered amount and the agent's offer is less than -100, but will accept the offer if

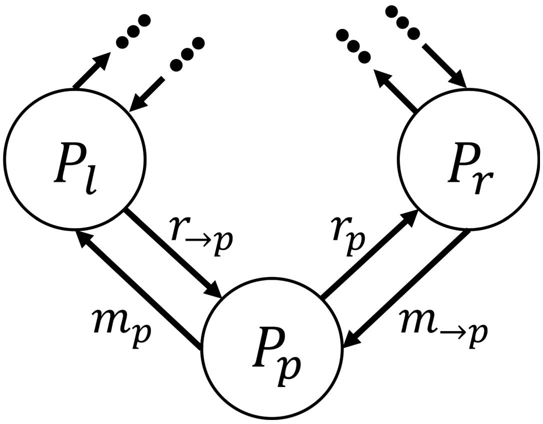


FIGURE 10. Diagram showing the interactions involving player P_p in the game scenario. The amount offered by player P_r, m_r, is represented by $m_{\to p}$, while player P_l's response, r_l, is represented by $r_{\to p}$.

the difference is greater than or equal to -100. Therefore, the value function $\mathrm{V}_{\mathrm{diff}}(m_d)$ serves as a threshold for determining the maximum amount of loss that the agent is willing to accept. The variables m_o and m_d were elements of the sets M_o and M_d, respectively, which were defined in (6) and (7).

$$m_o \in M_o = \{0, 10, 20, \ldots, 990, 1000\} \tag{6}$$

$$m_d \in M_d = \{-1000, -990, \ldots, 0, \ldots, 990, 1000\} \tag{7}$$

C. INTERNAL STATE

The internal state $\mathfrak{S}_a$ and $\widehat{\mathfrak{S}_p}$ shown in Fig. 9 were constructed as follows (8) and (9).

$$\mathfrak{S}_a = \{\mathrm{V}_{\mathrm{offr}}(m_o), \mathrm{V}_{\mathrm{diff}}(m_d)\} \tag{8}$$

$$\widehat{\mathfrak{S}_p} = \{\widehat{\mathrm{V}^p_{\mathrm{offr}}}(m_o), \widehat{\mathrm{V}^p_{\mathrm{diff}}}(m_d)\} \tag{9}$$

Assuming that other players had the same decision-making policy as the agent using the proposed model, the estimated internal state of each player $\widehat{\mathfrak{S}_p}$ was made up of their respective estimated value functions $\widehat{\mathrm{V}^p_{\mathrm{offr}}}$ and $\widehat{\mathrm{V}^p_{\mathrm{diff}}}$. These estimated value functions were continuously updated based on observations of the game environment $\mathfrak{O}_{t,a}$. The internal states $\mathfrak{S}_a$ and $\widehat{\mathfrak{S}_p}$ at the start of the game were composed of initial value functions, which were set to output specific values (see Section V-E and V-D).

The process of renewing the internal states involves two steps. First, upon receiving an input $\mathfrak{O}_{t,a}$, both the internal state $\mathfrak{S}_a$ and the estimated internal states $\widehat{\mathfrak{S}_p}$ were updated (step 1). Then, the agent's internal state $\mathfrak{S}_a$ was updated based on the estimated internal states $\widehat{\mathfrak{S}_p}$ (step 2). Once step 2 was completed, the agent would output action for the next round $A_{t+1,a}$ based on its updated internal state $\mathfrak{S}_a$.

1) RENEW STATE $\mathfrak{S}_a$

Fig. 10 depicts a conceptual diagram of the interactions surrounding player P_p in the group ultimatum game. The counterparties of P_p are P_l and P_r, and player P_p receives the offered amount $m_{\to p}$ $(= m_r)$ and the response $r_{\to p}$ $(= r_l)$ from the other players. Thus, the interactions in which player P_p is directly involved can be represented as a set T_p using the expression (10).

$$T_p = \{m_p, r_p, m_{\to p}, r_{\to p}\} \tag{10}$$

(11) and (13) represent a method of renewing the value function $\mathrm{V}_{\mathrm{offr}}(m)$ concerning the amount of money offered by the agent and a reward function $\mathrm{R}_{\mathrm{offr}}(m)$ used for renewing $\mathrm{V}_{\mathrm{offr}}(m)$. Here, m is an element of the set M_o shown in (6), and the set of player numbers is denoted as $P = \{1, 2, 3\}$. (13) and (14) represent a method for renewing the value function $\mathrm{V}_{\mathrm{diff}}$ and the reward function $\mathrm{R}_{\mathrm{diff}}$ for the infimum of acceptable difference infAD. Here, $m \in M_d$ is shown in (7), and the reward is calculated based on each player's response $r_{\to p}$ to the offered amount m_p.

$$\begin{aligned}\mathrm{V}_{\mathrm{offr}}(m) \leftarrow\ & (1-\alpha)\mathrm{V}_{\mathrm{offr}}(m) \\ & +\alpha\left(\mathrm{R}_{\mathrm{offr}}(m) + \gamma \max \mathrm{V}_{\mathrm{offr}}(m)\right)\end{aligned} \tag{11}$$

$$\begin{aligned}\mathrm{R}_{\mathrm{offr}}(m) =\ & \sum_{p\in P}\left\{(1-r_{\to p})\times\exp\left(-\frac{(m_p-m)^2}{\mathrm{kurtosis}}\right)\right\} \\ & + \sum_{p\in P}\left\{r_{\to p}\times\exp\left(-\frac{(m_p+c_\Delta-m)^2}{\mathrm{kurtosis}}\right)\right\}\end{aligned} \tag{12}$$

$$\begin{aligned}\mathrm{V}_{\mathrm{diff}}(m) \leftarrow\ & (1-\alpha)\mathrm{V}_{\mathrm{diff}}(m) \\ & +\alpha\left(\mathrm{R}_{\mathrm{diff}}(m) + \gamma \max \mathrm{V}_{\mathrm{diff}}(m)\right)\end{aligned} \tag{13}$$

$$\mathrm{R}_{\mathrm{diff}}(m) = \sum_{p\in P}\left\{(1-r_p)\times \mathrm{E}(m_{\to p}, m_p, m)\right\} \tag{14}$$

$$\mathrm{E}(m_{\to p}, m_p, m) = \exp\left(-\frac{(\mathrm{N}(m_{\to p}-m_p)-m)^2}{\mathrm{kurtosis}}\right) \tag{15}$$

$$\mathrm{N}(m) = \frac{m-|m|}{2} = \begin{cases} 0 & (m \geq 0) \\ -m & (m < 0)\end{cases} \tag{16}$$

In the above equations, the symbol γ represents the discount rate, and α represents the learning rate which are the agent's parameters for renewing the value functions. Based on these parameters, the agent decides how much a given reward should affect the renewal of the value function. The kurtosis is a Gaussian parameter that determines the shape of the reward function. The parameters are constants and make up the agent's internal state, reflecting the agent's personality (the degree of influence the agent receives from others in the group) concerning renewing its internal state.

2) RENEW ESTIMATED STATE OF OTHERS $\widehat{\mathfrak{S}_p}$

While the agent had access to information on all players, the human players' observations were limited (see Fig. 1). To account for this, the agent made the assumption that each human player p made decisions in the same way as itself and estimated the internal state of the human players as $\widehat{\mathfrak{S}_p}$.

The agent renewed the estimated value functions of a player $\widehat{\mathrm{V}^p_{\mathrm{offr}}}(m_o)$ and $\widehat{\mathrm{V}^p_{\mathrm{diff}}}(m_d)$ based on a portion of observed information $\mathfrak{O}_a$. In renewing the value functions, the agent utilized the corresponding reward functions $\widehat{\mathrm{R}^{p}_{\mathrm{offr}}}(m_o)$ and $\widehat{\mathrm{R}^{p}_{\mathrm{diff}}}(m_d)$. The method used to renew the value functions $\widehat{\mathrm{V}^{p}}(m_o)$ and the reward functions $\widehat{\mathrm{R}^{p}}(m_d)$ was similar to that shown in (11) to (14).

When the agent updated the internal state of a player P_p, the interactions observed by player P_p in (10) were used to assign a set of the player's observable interactions $\{P_p, P_r\}$, as shown in Fig. 10, to the set P of reward functions in (13) and (14). Here, the values of α, γ, and kurtosis were the same as those of the agent's parameters.

3) RENEW STATE $\mathfrak{S}_a$ BASED ON STATE $\widehat{\mathfrak{S}_p}$

The purpose of renewing the agent's internal state $\mathfrak{S}_a$ based on the estimated inner states of players $\widehat{\mathfrak{S}_p}$ was to adjust its value functions, such as $\mathrm{V}_{\mathrm{offr}}$, in a way that maximizes its profit while increasing the likelihood of its offer being accepted by the player. The game players aim to achieve two objectives: first, the difference between their offer and the offer presented to them should be greater than their infimum of an acceptable difference infAD, and second, they should be able to persuade the other player to accept their offer.

To satisfy the above two conditions, the agent updated its value function $\mathrm{V}_{\mathrm{offr}}$ based on the estimated internal state of the players, $\widehat{\mathfrak{S}_p}$. The update process is described in (17), which shows the range of offer amounts m that the agent should consider based on $\widehat{\mathrm{infAD}_r}$, $\widehat{m}_r$, $\widehat{m}_l$, and infAD from (18) to (21). By multiplying $\mathrm{V}_{\mathrm{offr}}(m)$ by a constant c_v that is less than 1.0, the values of the offered amount m outside that range are reduced. After updating the value function with (17), action A is output to the environment based on the internal state including the value functions in the subsequent round.

$$\mathrm{V}_{\mathrm{offr}}(m) \leftarrow c_v \times \mathrm{V}_{\mathrm{offr}}(m)$$
$$(\text{not if } \widehat{\mathrm{infAD}_l} + \widehat{m}_l \leq m \leq \widehat{m}_r - \mathrm{infAD}) \tag{17}$$
$$\widehat{\mathrm{infAD}_l} = \arg\max \widehat{\mathrm{V}^{l}_{\mathrm{diff}}} \tag{18}$$
$$\widehat{m}_l = \arg\max \widehat{\mathrm{V}^{l}_{\mathrm{offr}}} \tag{19}$$
$$\widehat{m}_r = \arg\max \widehat{\mathrm{V}^{r}_{\mathrm{offr}}} \tag{20}$$
$$\mathrm{infAD} = \arg\max \mathrm{V}_{\mathrm{diff}} \tag{21}$$

The variables in (18) to (21) that compose the inequality in (17) were determined by the agent's value function and the estimated value functions of players p_l and p_r, assuming that player P_p in Fig. 10 is the agent. The agent will propose the offer m to player P_l, and the agent should ensure that the offer satisfies the infimum of acceptable differences infAD of players $P_{p=a}$ and P_l. Therefore, from the agent's perspective, the following two inequalities, (22) and (23) should be satisfied.

$$\mathrm{infAD} \leq \widehat{m}_r - m_{p=a} \tag{22}$$
$$\widehat{\mathrm{infAD}_l} \leq m_{p=a} - \widehat{m}_l \tag{23}$$

These two inequalities can be summarized as inequalities in (17).

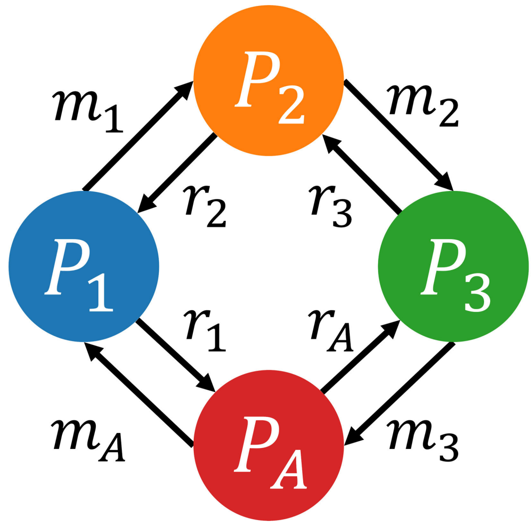


FIGURE 11. Diagram of players (circled Ps) and their interactions(m, r) in a round of the group ultimatum game. Circled Ps show the game players; human players (P_1, P_2, P_3) and agent(P_A). m_x and r_x indicate offered amount and response of player P_x.

V. EXPERIMENT

Fig. 11 shows interactions between game players, where P represents a player, $\{P_1, P_2, P_3\}$ denotes a set of human participants, and P_A denotes a computer agent. The three human participants in the experiment were randomly assigned to a player P_1, P_2, or P_3 without knowing which colored robot represented a human or an agent.

This experiment was approved by the Ethics Review Department for Research on Humans at Toyama Prefectural University, under reference number R4–6 on November 29th, 2022.

A. PROCEDURES

After providing informed consent and basic demographic information, participants received a briefing on the rules of the basic ultimatum game and the group ultimatum game, as well as information on the impact of their scores on their experimental reward. Verbal communication between participants was prohibited, and they were uninformed about which player was controlling the robot in the game. Additionally, participants interacted with different players during the first and second games while being monitored by the experimenter.

In the group, the game players participated twice and changed their positions between the first and second games. The offering and responding interaction partners were different players in the two games. In particular, a human player in position P_1 moved to P_3, and a human player in position P_3 moved to P_1. The participants were informed that their positions had changed, but they were not given specific details about how the change occurred.

The first game of Groups A and B involved the adaptive agent using the proposed model participating as a P_A, while the maladaptive agent with the comparison model (see

TABLE 1. Condition of the agent.

Learning rate α	0.2
Discount Factor γ	0.8
Calibration constant of reward function c_Δ	50
Scaling constant of value function c_v	0.5
Gaussian parameter of reward function kurtosis	10000
initial $\arg\max \mathrm{V}_{\mathrm{offr}}$	100
initial $\arg\max \mathrm{V}_{\mathrm{diff}}$	-100

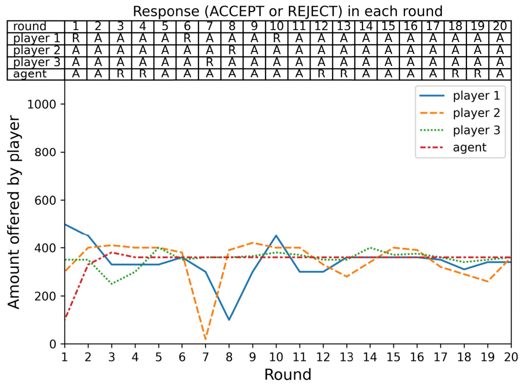


FIGURE 12. Offer *m* and response *r* in the first game of group A. Agent embedded by proposed model participated.

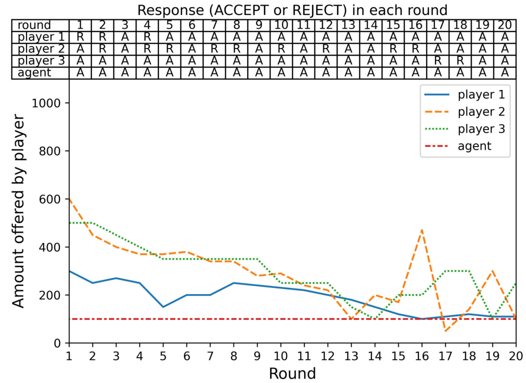


FIGURE 13. Offer *m* and response *r* in the second game of group A. Agent for comparison participated.

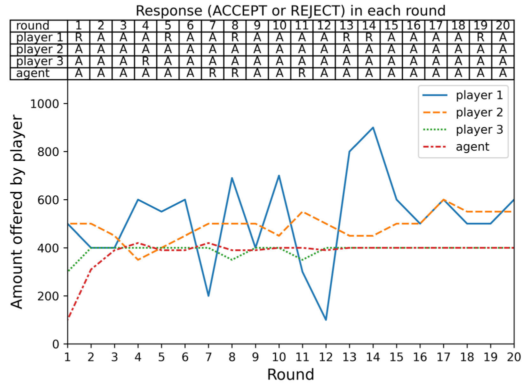


FIGURE 14. Offer *m* and response *r* in the first game of group B. Agent embedded by proposed model participated.

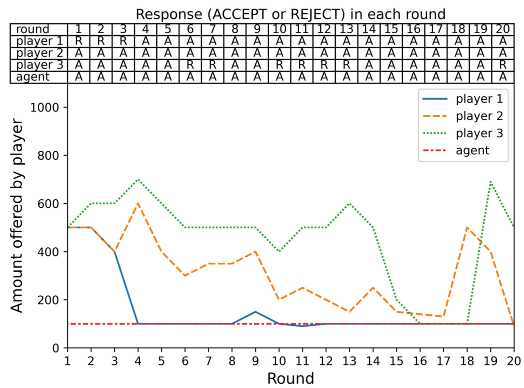


FIGURE 15. Offer *m* and response *r* in the second game of group B. Agent for comparison participated.

Section V-D) participated as a P_A in the second game. The agents' participation order effects were reversed in Groups C, D, and E to control for order effects.

B. PARTICIPANTS AND CONDITION

Each of the fifteen participants was assigned to one of the five groups (A, B, C, D, and E). All participants were male and had an average age of 22.53 (SD $=$ 0.88). The conditions of the agent based on the proposed model are shown in Table 1.

C. REWARD

A few weeks after the experiment, the participants received a success reward proportionate to the amount of money they earned in the games in the form of *Amazon gift cards* ranging from 500 yen to 1, 000 yen. Participants received a reward based on the game in which they scored higher between the two games. The reward details were communicated to the participants while explaining the rules of the group ultimatum game. The experiment was conducted and completed in December 2022, lasting approximately 30 min. Note that the minimum wage in the Toyama prefecture, where the participants resided, was 908 yen per hour. The potential to earn up to 1, 000 yen in just 30 minutes, double the minimum hourly wage in the Toyama prefecture where the participants resided, provided an incentive for them to participate in the experiment.

D. COMPARISON MODEL

In this experiment, we included a maladaptive agent as a point of comparison to the proposed model, which adapted to group norms. Specifically, the comparison agent exhibited a fixed offer m and response r (the agent's action $A_a = \{m_{t=1,a} = 100, r_{t=1,a} = 0\}$) at every round of the game in this study.

Both agents had identical initial internal states $\mathfrak{S}_a$ and demonstrated an offer of $m_{t=1,a} = 100$ and a response of $r_{t=1,a} = 0$ in the first round. The agent using the proposed model updated its internal state $\mathfrak{S}_a$ and its estimated internal states of each player $\widehat{\mathfrak{S}_p}$ after each action and observation, nevertheless, the comparison agent did not update its states.

TABLE 2. Total amount that each player got in a group, absolute value of agent's relative gain, and number of human players' rejections.

Group	Game	Model	Total amount (TA) finally obtained by agent	player 1	player 2	player 3	Sum of total amount	Absolute value of agent's relative gain	Number of human players' REJECTs
A	1st	proposed	9,140	9,060	8,720	9,080	36,000	140	5
B	1st	proposed	11,130	9,870	10,850	12,150	44,000	130	7
C	2nd	proposed	10,072	9,800	9,990	10,138	40,000	72	7
D	2nd	proposed	10,190	10,810	8,950	10,050	40,000	190	8
E	2nd	proposed	10,460	10,060	12,130	11,350	44,000	540	10
A	2nd	for comparison	9,200	7,320	7,480	8,000	32,000	1,200	14
B	2nd	for comparison	12,990	9,950	7,730	9,330	40,000	2,990	10
C	1st	for comparison	5,119	3,700	2,600	4,581	16,000	1,119	17
D	1st	for comparison	8,508	8,500	8,650	10,342	36,000	492	15
E	1st	for comparison	7,850	6,000	7,000	7,150	28,000	850	14

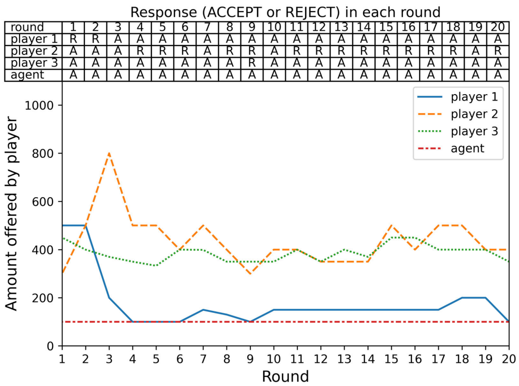


FIGURE 16. Offer *m* and response *r* in the first game of group C. Agent for comparison participated.

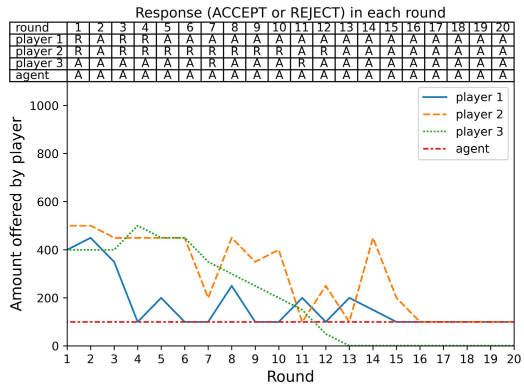


FIGURE 18. Offer *m* and response *r* in the first game of group D. Agent for comparison participated.

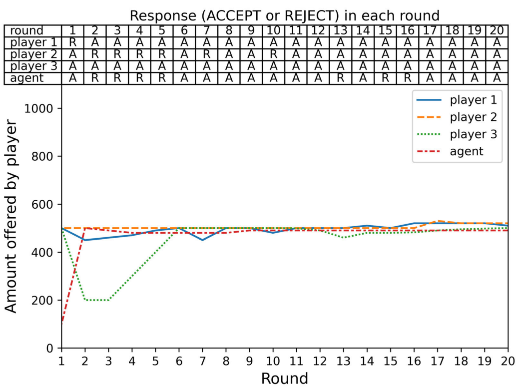


FIGURE 17. Offer *m* and response *r* in the second game of group C. Agent embedded by proposed model participated.

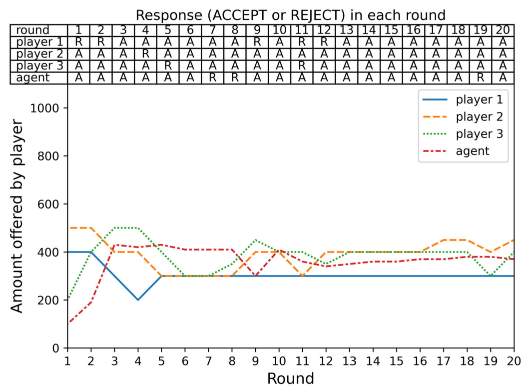


FIGURE 19. Offer *m* and response *r* in the second game of group D. Agent embedded by proposed model participated.

This comparison allowed us to contrast adaptive agents to group norms with non-adaptive agents.

E. INITIAL INNER STATE

For agents to exhibit an initial action $A_a = \{m_{t=1,a}, r_{t=1,a}\}$ in a first round, the agent's initial internal state $\mathfrak{S}_a = \{\mathrm{V}_{\mathrm{offr}}(m_o), \mathrm{V}_{\mathrm{diff}}(m_d)\}$ was set as shown in (24) and (25).

$$\mathrm{V}_{\mathrm{offr}}(m_o) \leftarrow 0.1 \times \exp\left(-\frac{(100 - m_o)^2}{\mathrm{kurtosis}}\right) \tag{24}$$

$$\mathrm{V}_{\mathrm{diff}}(m_d) \leftarrow 0.1 \times \exp\left(-\frac{(-100 - m_d)^2}{\mathrm{kurtosis}}\right) \tag{25}$$

(24) and (25) show that a Gaussian function is assigned to the initial value functions. For example, when m_o is 100 in (24), $\mathrm{V}_{\mathrm{offr}}(100)$ was assigned 0.1. This means that $\arg\max \mathrm{V}_{\mathrm{offr}}$ is 100, and the agent in the first round offers 100 yen. (25) was set up in the same way, with a Gaussian function so that $\arg\max \mathrm{V}_{\mathrm{diff}}$ is -100 at the first round. By setting the initial value functions in this way, the behavior of the agents in the first round was controlled.

The initial estimated internal state of a player $\widehat{\mathfrak{S}_p}$ was set to be equal to the agent's initial internal state $\mathfrak{S}_a$. Therefore, initial $\widehat{\mathrm{V}^{p}_{\mathrm{offr}}}(m_o)$ = initial $\mathrm{V}_{\mathrm{offr}}(m_o)$ and initial $\widehat{\mathrm{V}^{p}_{\mathrm{diff}}}(m_d)$ = initial $\mathrm{V}_{\mathrm{diff}}(m_d)$. At the step of presenting an offer in the first round, the agent had not observed the other players' offers.

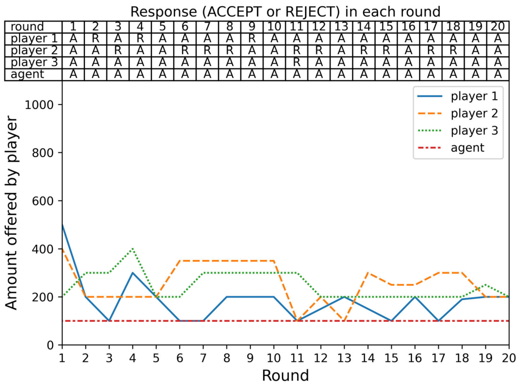


FIGURE 20. **Offer *m* and response *r* in the first game of group E. Agent for comparison participated.**

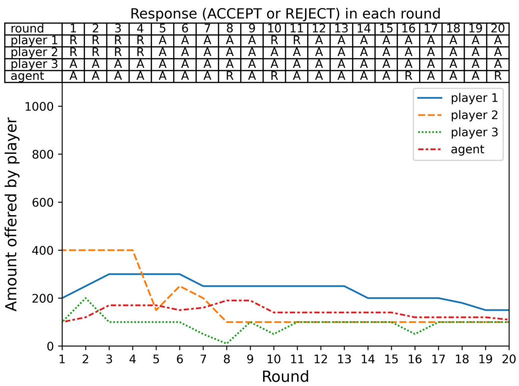


FIGURE 21. **Offer *m* and response *r* in the second game of group E. Agent embedded by proposed model participated.**

The initial step of this experiment involved the agent estimating its initial internal state, $\widehat{\mathfrak{S}_p}$, based on the assumption that the other players would make decisions similar to its own. The agent using the proposed model updated its internal states $\mathfrak{S}_a$ and $\widehat{\mathfrak{S}_p}$, based on the behavior of the other players as the game progressed.

The impact of an agent's adaptability to group norms on human players can be investigated by setting the value functions for adaptive and maladaptive agents as described in (24) and (25), respectively. In the ultimatum game, humans typically offer about 40% of the distributed amount [12]. For example, if an agent offers 500 yen in the first round, other players' offers in the first round may be closely tied to the agent's offer. Consequently, the agent has no opportunity to demonstrate its ability to adapt to the group norm, which contradicts the purpose of our study. To overcome this issue, we set the initial internal state of the agent to offer a low amount of money (100 yen), which is unlikely to be offered by humans in the first round. This allows us to observe the agent's adaptive behavior over multiple rounds. Additionally, we set the comparison agent to always accept, as it was not adaptive.

F. RESULT

We compared the scores and behavior of the players in two groups: one with an agent using the proposed model and the other with a different agent. Table 2 presents the results of both games in each group, including the total amount of money earned by each player, the sum of the total amount of money earned by the players in the group, the absolute value of the relative gain achieved by the agents, and the total number of REJECTs expressed by human players in the groups. The term "relative gain" denotes the absolute value of the difference between the average amount earned by the four players and the amount obtained by the agent.

1) INDIVIDUAL GAME RESULTS

Figs. 12 to 21 display the actions of players in a single game. The horizontal axis shows the number of rounds, and the vertical axis shows the amount of the offer made. Additionally, the table above each graph indicates whether each player accepted(A) or rejected(R) the received offer. Based on these results, we observed that the agents who utilized the proposed model (as seen in Figs. 12, 14, 17, 19, and 21) demonstrated greater adaptability in the amounts they offered, compared to the agent using the comparison model (as seen in Figs. 13, 15, 16, 18, and 20). Additionally, the agents using the proposed model occasionally reject the offers presented to them. The human players also adjusted the amounts they offered. However, some of the human players showed larger fluctuations in the amount of money offered than the agents' fluctuations, such as player 1 in Fig. 14 and player 2 in Fig. 18.

In Fig. 18, we observed that player number 3 persistently offered a meager amount of one yen after the 13th round, implying that the player may have realized that the agent always accepted any offer made by the player. This outcome suggests that the maladaptive agent was incapable of coping with the self-centered behavior of the other players. In contrast, as seen in Fig. 21, when player 3 offered a highly unjust offer to the adaptive agent in the eighth round, the agent could refuse the offer, demonstrating its capability to handle highly unfair proposals.

2) OVERALL GAME RESULTS

Fig. 22 displays the absolute value of the difference between the total amount acquired by the agent and the total amount obtained by the other players in the same group as the agent. When considering the total amount of player p in a group g, denoted by $\mathrm{TA}_{g,p}$, the absolute value difference between the total amount achieved by the agent and the total amount obtained by the other players in the group is given by $\{|\mathrm{TA}_{g,p} - \mathrm{TA}_{g,p=a}|\}$. It is worth noting that in this context, g takes on one of the values A, B, C, D, E, or F, while p is an element of the set $\{1, 2, 3\}$. Here, $p = a$ indicates that the player in question is the agent. The proposed model ($M = 566.67$, SD $= 523.98$) yielded a lesser absolute difference than the comparison model ($M = 1774.67$, SD $= 1405.68$). The Brunner–Munzel test [42] indicated a significant difference between the two groups (statistic$=$ 3.476 and a p–value of 0.002367).

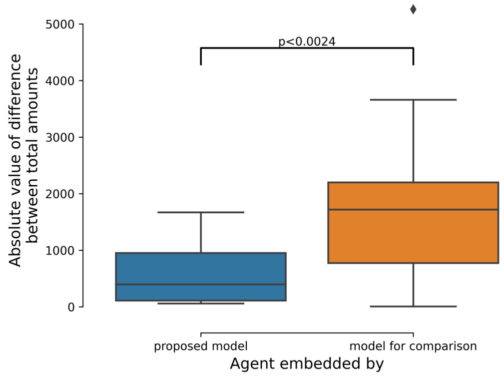


FIGURE 22. Comparison of the difference between the total amount of money earned by the agent and the other players in their group, represented as an absolute value. This value is indicative of the agent's profitability relative to the other players in the group.

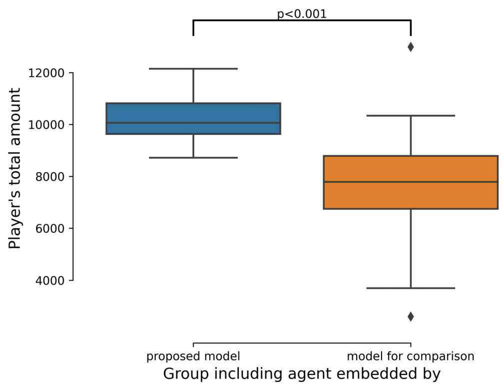


FIGURE 23. Comparison of the total amount of money earned by players who participated in a group with the agent using the proposed model and the comparison model. It is an indicator of which type of agent was more beneficial for human players.

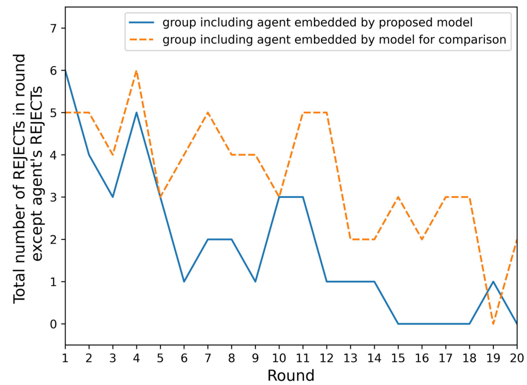


FIGURE 24. Total number of REJECT expressed by each group of human players at each round.

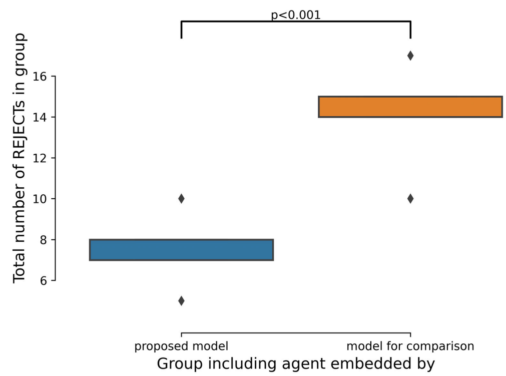


FIGURE 25. Comparison of the number of REJECT expressed by human players in the game for each game.

Fig. 23 illustrates a comparison of the total amounts obtained by players in the group where the agent with the proposed model was embedded and the group where the comparison agent was present. The results reveal that the groups with the adaptive agent achieved relatively higher gains ($M = 10200.00$, SD $= 983.63$) compared to the group with the agent for comparison ($M = 7600.00$, SD $=$ 2398.66). The Brunner–Munzel test [42] showed a significant difference between the two groups (statistic$= -8.391$, p–value$< .001$).

Figs. 24 and 25 depict the total number of REJECTs expressed by human players in the groups. The horizontal axis of Fig. 24 represents the number of rounds, and the vertical axis shows the total number of REJECTs in the groups. We observed that the number of REJECTs decreased as the game progressed for both groups. Fig. 25 compares the number of REJECTs expressed by human players in each group. The results show that REJECTs were less frequent in the groups with the adaptive agent ($M = 7.40$, SD $= 1.82$) than the group with the agent for comparison ($M = 14.00$, SD $= 2.55$). The Brunner–Munzel test [42] revealed a significant difference between the two groups (statistic$= 16.97$, p–value$< .001$).

G. DISCUSSION

Based on the results presented in Fig. 22, it is evident that the agent using the proposed model with group norm awareness tended to receive an amount that was closer to the other players, as compared to the maladaptive agent. Research has shown that people generally have a preference for avoiding inequality and strive for a situation where there is not a significant difference between their profits and those of others [9]. In our study, we observed that this tendency resulted in group norm-aware decision-making, which facilitated the formation of fair offers and responses, ultimately minimizing the risk of significant losses and enabling the adaptive agent to obtain fair game scores relative to the comparison agent. Notably, the adaptive agent's relative profit being close to zero indicates that the agent experienced neither extreme gains nor losses. This means that the agent exhibited a balanced and fair behavior that was both altruistic and selfish under conflicting interest. Additionally, the agent's rejection of player 3's offer in the eighth round, as illustrated in Fig. 21, demonstrates the agent's ability to decline offers that were perceived as extremely unfair to itself. Overall, the agent embedding the proposed model exhibited human-like and reasonable behavior as game players, utilizing their internal state to make decisions, including the rejection of offers that were perceived as unfair. This implies that the group norm-aware agent was equipped with an aspect of social competence regarding fairness.

Fig. 23 compares the benefits gained by the participants who participated in the group. As indicated in Table 2, the comparison agent obtained relatively higher profits than the other players in the group. However, the total amount gained by the players in the group was lower than the amount achieved by players with the agent based on the proposed model. These results suggest that agents who exhibited adaptive behavior to the group norms could achieve mutually beneficial behavior among the players in the group by demonstrating fairness. The adaptive behavior exhibited by the agents in adhering to group norms and demonstrating fairness may have contributed to an overall increase in the benefits of the group as a whole.

Interestingly, the data presented in Fig. 25 indicates that the adaptive behavior of the agent towards group norms resulted in a decreased likelihood of human players rejecting offers made to them in the group where the agent participated. Consequently, less rejection in the groups resulted in high final profits for the players. Therefore, we can conclude that the fair behavior of the agent had a positive impact on the human players in the groups. For example, such fair behavior reduced the likelihood of human players rejecting offers, resulting in a mutually beneficial relationship between the human players and the agent in the game scenario.

In this study, the agent using the proposed model lacks prior knowledge of fair offers in the ultimatum game. Instead, during interactions in GUGs, the proposed model allowed the agent to select actions that would contribute to a fair distribution from a set of possible actions. Additionally, the group that included the agent exhibited increased reciprocal behavior among its members. Based on the proposed model, the agent's fair behavior as a group member resulted in higher and more equitable benefits for the group. These findings suggest that the group norm-aware nature of the model may enable the computational agent to adapt to the group as a teammate in a socially-appropriate manner.

VI. CONCLUSION

Social robots are expected to exhibit human-like social skills to gain acceptance from humans. This study assessed the behavior of a group norm-aware agent in the context of the ultimatum game scenario played with human participants in groups. Like humans who value fairness and equality [9], the agents had to act appropriately in both decision-making phases of a single round of the game, which offers money to a human player and responds to the other's offer. In other words, non-human agents that engage in the game in a human-like manner should prioritize high gains or losses as group members.

We conducted an experimental evaluation with two agents: the first agent made decisions based on a proposed decision-making model, attempting to adapt to group norms, whereas the other agent made maladaptive decisions that did not conform to group norms for comparison. Our findings indicate that the participation of the group norm-aware agent in the game decreased non-cooperative behavior, such as rejecting offers, in the groups where the agent was present. This trend led to higher game scores for each group member in those groups than the groups where the maladaptive agent participated. Furthermore, we observed that the agent's game scores were more similar to those of other players, suggesting that the agent exhibited behavior perceived as fair and acceptable by the human players.

We plan to conduct future experiments using a robot that combines our proposed model. Unlike the current study, the robot will have nonverbal expressions, such as emotions and gestures, that can help achieve a more sophisticated and nuanced behavior in addition to group norm-aware behavior. We will also investigate how human participants behave and their subjective impressions of the social robot in the group ultimatum game scenario.

REFERENCES

[1] I. Leite, C. Martinho, and A. Paiva, "Social robots for long-term interaction: A survey," *Int. J. Social Robot.*, vol. 5, no. 2, pp. 291–308, 2013.

[2] T. B. Sheridan, "A review of recent research in social robotics," *Current Opinion Psychol.*, vol. 36, pp. 7–12, Dec. 2020.

[3] W. Wang, Y. Chen, R. Li, and Y. Jia, "Learning and comfort in human–robot interaction: A review," *Appl. Sci.*, vol. 9, no. 23, p. 5152, Nov. 2019.

[4] C. A. Cruz and T. Igarashi, "A survey on interactive reinforcement learning: Design principles and open challenges," in *Proc. ACM Designing Interact. Syst. Conf.*, Jul. 2020, pp. 1195–1209.

[5] N. Akalin and A. Loutfi, "Reinforcement learning approaches in social robotics," *Sensors*, vol. 21, no. 4, 1292, 2021.

[6] D. C. Feldman, "The development and enforcement of group norms," *Acad. Manage. Rev.*, vol. 9, no. 1, pp. 47–53, Jan. 1984.

[7] Y. Fuse and M. Tokumaru, "Social influence of group norms developed by human–robot groups," *IEEE Access*, vol. 8, pp. 56081–56091, 2020.

[8] Y. Fuse, E. Ayedoun, and M. Tokumaru, "Group norm-aware robot adaptively maintains suitable interpersonal distance in human–robot group," *Int. J. Affect. Eng.*, vol. 21, no. 3, pp. 191–201, 2022.

[9] E. Fehr and K. M. Schmidt, "A theory of fairness, competition, and cooperation," *Quart. J. Econ.*, vol. 114, no. 3, pp. 817–868, Aug. 1999.

[10] Y. Fuse, B. Ashida, E. Ayedoun, and M. Tokumaru, "Decision-making model for robots that considers group norms and interests," in *Human-Computer Interaction. Technological Innovation*. Salmon Tower Building, NY, USA: Springer, 2022, pp. 475–485.

[11] R. H. Thaler, "Anomalies: The ultimatum game," *J. Econ. Perspect.*, vol. 2, no. 4, pp. 195–206, Nov. 1988.

[12] H. Oosterbeek, R. Sloof, and G. Van De Kuilen, "Cultural differences in ultimatum game experiments: Evidence from a meta-analysis," *Exp. Econ.*, vol. 7, no. 2, pp. 171–188, Jun. 2004.

[13] L. Bishop, A. van Maris, S. Dogramadzi, and N. Zook, "Social robots: The influence of human and robot characteristics on acceptance," *J. Paladyn Behav. Robot.*, vol. 10, no. 1, pp. 346–358, Jan. 2019.

[14] M. Heerink, B. Kröse, V. Evers, and B. Wielinga, "Assessing acceptance of assistive social agent technology by older adults: The almere model," *Int. J. Social Robot.*, vol. 2, no. 4, pp. 361–375, Dec. 2010.

[15] C. S. Song and Y.-K. Kim, "The role of the human–robot interaction in consumers' acceptance of humanoid retail service robots," *J. Bus. Res.*, vol. 146, pp. 489–503, Jul. 2022.

[16] F. Correia, F. S. Melo, and A. Paiva, "When a robot is your teammate," *Topics Cognit. Sci.*, Dec. 2022. [Online]. Available: https://onlinelibrary.wiley.com/doi/full/10.1111/tops.12634, doi: 10.1111/tops.12634.

[17] J. Brandstetter, P. Racz, C. Beckner, E. B. Sandoval, J. Hay, and C. Bartneck, "A peer pressure experiment: Recreation of the Asch conformity experiment with robots," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2014, pp. 1335–1340.

[18] N. Salomons, M. Van Der Linden, S. Strohkorb Sebo, and B. Scassellati, "Humans conform to robots: Disambiguating trust, truth, and conformity," in *Proc. ACM/IEEE Int. Conf. Hum.-Robot Interact.*, Feb. 2018, pp. 187–195.

[19] S. E. Asch, "Effects of group pressure on the modification and distortion of judgments," in *Groups, Leadership and Men*, H. Guetzknow, Ed. Pittsburgh, PA, USA: Carnegie Press, 1951, pp. 177–190.

[20] E. M. Aminoff, D. Clewett, S. Freeman, A. Frithsen, C. Tipper, A. Johnson, S. T. Grafton, and M. B. Miller, "Individual differences in shifting decision criterion: A recognition memory study," *Memory Cognition*, vol. 40, no. 7, pp. 1016–1030, Oct. 2012.

[21] M. Deutsch and H. B. Gerard, "A study of normative and informational social influences upon individual judgment.," *J. Abnormal Social Psychol.*, vol. 51, no. 3, pp. 629–636, Nov. 1955.

[22] L. S. Rashotte, "Social influence," in *The Blackwell Encyclopedia Sociology*, vol. 9, G. Ritzer, and J. M. Ryan, Eds., Oxford, U.K.: Blackwell Publishing, 2007, pp. 4426–4429.

[23] R. B. Cialdini and N. J. Goldstein, "Social influence: Compliance and conformity," *Annu. Rev. Psychol.*, vol. 55, no. 1, pp. 591–621, 2004.

[24] M. Sherif, "A study of some social factors in perception," *Arch. Psychol.*, vol. 27, no. 187, pp. 5–16, 1934.

[25] R. B. Cialdini, "Influence: Science and practice," in *HarperCollins*. New York, NY, USA: HarperCollins Publishers, 2009.

[26] J. Jetten, R. Spears, and A. S. R. Manstead, "Strength of identification and intergroup differentiation: The influence of group norms," *Eur. J. Social Psychol.*, vol. 27, no. 5, pp. 603–609, Sep. 1997.

[27] T. Postmes, R. Spears, and S. Cihangir, "Quality of decision making and group norms," *J. Personality Social Psychol.*, vol. 80, no. 6, pp. 918–930, 2001.

[28] J. Jetten, T. Postmes, and B. J. McAuliffe, "'We're *all* individuals': Group norms of individualism and collectivism, levels of identification and identity threat," *Eur. J. Social Psychol.*, vol. 32, no. 2, pp. 189–207, 2002.

[29] R. Williams, C. V. Machado, S. Druga, C. Breazeal, and P. Maes, "'My doll says it's ok': A study of children's conformity to a talking doll," in *Proc. 17th ACM Conf. Interact. Design Children*, Jun. 2018, pp. 625–631.

[30] A.-L. Vollmer, R. Read, D. Trippas, and T. Belpaeme, "Children conform, adults resist: A robot group induced peer pressure on normative social conformity," *Sci. Robot.*, vol. 3, no. 21, Aug. 2018.

[31] C. Beckner, P. Rácz, J. Hay, J. Brandstetter, and C. Bartneck, "Participants conform to humans but not to humanoid robots in an English past tense formation task," *J. Lang. Social Psychol.*, vol. 35, no. 2, pp. 158–179, Mar. 2016.

[32] P. Riva, N. Aureli, and F. Silvestrini, "Social influences in the digital era: When do people conform more to a human being or an artificial intelligence?" *Acta Psychologica*, vol. 229, Sep. 2022, Art. no. 103681.

[33] L. Masjutin, J. K. Laing, and G. W. Maier, "Why do we follow robots? An experimental investigation of conformity with robot, human, and hybrid majorities," in *Proc. 17th ACM/IEEE Int. Conf. Hum.-Robot Interact. (HRI)*, Mar. 2022, pp. 139–146.

[34] E. B. Sandoval, J. Brandstetter, M. Obaid, and C. Bartneck, "Reciprocity in human-robot interaction: A quantitative approach through the Prisoner's Dilemma and the Ultimatum Game," *Int. J. Social Robot.*, vol. 8, pp. 303–317, 2016. [Online]. Available: https://link.springer.com/article/10.1007/s12369-015-0323-x

[35] E. B. Sandoval, J. Brandstatter, U. Yalcin, and C. Bartneck, "Robot likeability and reciprocity in human robot interaction: Using ultimatum game to determinate reciprocal likeable robot strategies," *Int. J. Social Robot.*, vol. 13, no. 4, pp. 851–862, Jul. 2021.

[36] T. Law, J. De Leeuw, and J. H. Long, "How movements of a non-humanoid robot affect emotional perceptions and trust," *Int. J. Social Robot.*, vol. 13, no. 8, pp. 1967–1978, Dec. 2021.

[37] C. Di Dio, F. Manzi, S. Itakura, T. Kanda, H. Ishiguro, D. Massaro, and A. Marchetti, "It does not matter who you are: Fairness in pre-schoolers interacting with human and robotic partners," *Int. J. Social Robot.*, vol. 12, no. 5, pp. 1045–1059, Nov. 2020.

[38] A. Acosta-Mitjans, D. Cruz-Sandoval, R. Hervas, E. Johnson, C. Nugent, and J. Favela, "Affective embodied agents and their effect on decision making," in *Proc. 13th Int. Conf. Ubiquitous Comput. Ambient Intell. (UCAmI)*, Nov. 2019, pp. 1–11.

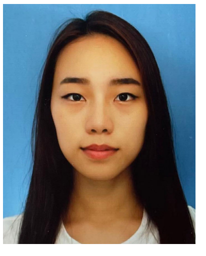
[39] H. Fukuda, M. Shiomi, K. Nakagawa, and K. Ueda, "'Midas touch' in human–robot interaction: Evidence from event-related potentials during the ultimatum game," in *Proc. 7th Annu. ACM/IEEE Int. Conf. Hum.-Robot Interact.*, Mar. 2012, pp. 131–132.

[40] H. Takesue, A. Ozawa, and S. Morikawa, "Evolution of favoritism and group fairness in a co-evolving three-person ultimatum game," *Europhys. Lett.*, vol. 118, no. 4, p. 48002, May 2017.

[41] F. P. Santos, J. M. Pacheco, A. Paiva, and F. C. Santos, "Evolution of collective fairness in hybrid populations of humans and agents," in *Proc. AAAI Conf. Artif. Intell.*, vol. 33, no. 1, 2019, pp. 6146–6153.

[42] K. Neubert and E. Brunner, "A studentized permutation test for the nonparametric Behrens–Fisher problem," *Comput. Statist. Data Anal.*, vol. 51, no. 10, pp. 5192–5204, Jun. 2007.

YOTARO FUSE (Member, IEEE) received the B.E. and M.E. degrees in engineering and the Dr. (Eng.) degree from Kansai University, in 2017, 2019, and 2022, respectively. From 2021 to 2022, he was a Research Fellow of the Japan Society for the Promotion of Science (DC2). He is currently an Assistant Professor with the Faculty of Engineering, Toyama Prefectural University. His current research interests include social robot and human–agent interaction. He is a member of the Institute of Electronics, Information, and Communication Engineers and the Japan Society for Fuzzy Theory and Intelligent Informatics.



BIINA ASHIDA received the B.E. degree in engineering from Kansai University, in 2021. Her research interests include social robot and human–agent interaction.

EMMANUEL AYEDOUN (Member, IEEE) received the Ph.D. degree in informatics from Osaka Prefecture University, in 2020. From 2018 to 2020, he was a Young Scientist Research Fellow (DC1) of the Japanese Society for the Promotion of Science (JSPS). He is currently an Assistant Professor with the Faculty of Engineering Science, Kansai University. His research interests include the intersection of human–computer interaction, affective sciences, and educational psychology. He is a member of the Japanese Society for Artificial Intelligence (JSAI) and the International Artificial Intelligence in Education Society (IAIED).

MASATAKA TOKUMARU (Member, IEEE) received the B.E. and M.E. degrees in engineering from Kansai University, Osaka, Japan, in 1994 and 1997, respectively, and the Dr. (Eng.) degree from Osaka City University, Japan, in 2001. From 1997 to 2007, he was a Research Associate or a Lecturer with the Department of Electronics, Kansai University, where he has been a Professor with the Faculty of Engineering Science, since 2016. His research interests include affective computing and evolutionary computation. He is a member of the Japan Society of Kansei Engineering, the Institute of Electronics, Information and Communication Engineers, the Japan Society for Fuzzy Theory and Intelligent Informatics, the Human Interface Society, and the Japanese Society for Artificial Intelligence.