Assessing the effect of dialogue on altruism toward future generations: A preliminary study

Elie Maalouly[1]*, Ryuji Yamazaki[2]*, Shuichi Nishio[2], Marco Nørskov[3], Kohei Kamaga[4], Shoji Komai[5], Kiyoshi Chiba[6], Keiichiro Atsumi[6] and Ken-Ichi Akao[6]

[1]Graduate School of Engineering Science, Osaka University, Osaka, Japan, [2]Institute for Open and Transdisciplinary Research Initiatives, Osaka University, Osaka, Japan, [3]Department for Philosophy and The History of Ideas, Aarhus University, Aarhus, Denmark, [4]Faculty of Economics, Sophia University, Tokyo, Japan, [5]Faculty of Engineering, International Professional University of Technology in Tokyo, Tokyo, Japan, [6]School of Social Sciences, Waseda University, Tokyo, Japan

Introduction: Despite the abundance of evidence on climate change and its consequences on future generations, people, in general, are still reluctant to change their actions and behaviors toward the environment that would particularly benefit posterity. In this study, we took a preliminary step in a new research direction to explore humans' altruistic behavior toward future generations of people and whether it can be affected by dialogue.

Methods: We used an android robot called Telenoid as a representative of future generations by explaining that the robot is controlled by an Artificial Intelligence (AI) living in a simulation of our world in the future. To measure people's altruistic behavior toward it, we asked the participants to play a round of the Dictator Game with the Telenoid before having an interactive conversation with the Telenoid and then playing another round.

Results: On average, participants gave more money to the Telenoid in the second round (after having an interactive conversation). The average amount of money increased from 20% in the first to about 30% in the second round.

Discussion: The results indicate that the conversation with the robot might have been responsible for the change in altruistic behavior toward the Telenoid. Contrary to our expectations, the personality of the participants did not appear to have an influence on their change of behavior, but other factors might have contributed. We finally discuss the influence of other possible factors such as empathy and the appearance of the robot. However, the preliminary nature of this study should deter us from making any definitive conclusions, but the results are promising for establishing the ground for future experiments.

1. Introduction

Climate change is a major threat to our planet and the human race. But it is a future issue that might not affect our current generation. However, if there is such a threat to future generations of humans, then why are we so indifferent and passive? Do we lack altruism toward future generations of people? And if so, how can we influence people's altruistic behavior to posterity?

It is expected that climate change would have a wide range of effects, including changes to ecosystems (Leemans and Eickhout, 2004), effects on human systems including water resources (Arnell and Lloyd-Hughes, 2014), forced human migration (Barnett and Adger, 2003), and extensive ocean acidification (Caldeira and Wickett, 2003). However, despite the abundance of evidence available on climate change, people are still hesitant to alter their energy-related choices and actions. Efforts to persuade people to adopt environmentally friendly habits have often been somewhat unsuccessful (Whitmarsh and O'Neill, 2010; Chu and Majumdar, 2012). Therefore, understanding how to influence this altruistic behavior can be crucial for the preservation of our environment and species, whereby carelessness and negligence might bring on devastating consequences.

But there is more than the preservation of the environment, as engaging in altruistic behavior rewards the person with multiple health benefits. Altruism is beneficial for a person's emotional health and can dramatically improve their peace of mind (Brunier et al., 2002), in addition to improving confidence, self-esteem, self-awareness, daily function, and reducing depression (Uccelli et al., 2004).

We aim to investigate whether humans are capable of altruistic behavior toward future generations of people and what factors might have an effect on this behavior. In this study, we would like to explore dialogue specifically as a factor in affecting altruistic behavior toward future generations. The main questions we seek answers to are whether humans can feel compassion for a complete stranger from a different timeline and how dialogue might affect it. In the following section, we will briefly introduce and define our understanding of altruistic behavior and empathy. We will then discuss how altruistic behavior is measured in experiments and how empathy, dialogue, and personality might influence it. We will begin by discussing how altruism is measured in experiments, namely using the Dictator Game.

2. Related work

2.1. Defining altruism

There is a big disagreement among researchers on how to define altruism and how it is linked to prosocial behavior. In more than 74% of the articles published that include the keywords "prosocial behavior," "prosocial lies," "altruism" and "altruistic behavior," there was no relevant definition included in the article (Pfattheicher et al., 2022). Therefore, properly defining these constructs is important to communicate how we understand altruistic behavior in this study.

The fact that prosocial and altruistic behaviors indicate a "positive" social behavior toward one or several others is common in nearly all definitions. These behaviors are most frequently defined as promoting or having as their primary goal the promotion of the welfare of others. Where the definitions differ is mainly in what constitutes prosocial or altruistic behavior. With regard to the final point, there are three perspectives on what is considered prosocial behavior, and altruism is defined differently for each of these perspectives. First, is the intentionalist perspective, where the definition of prosocial behavior emphasizes the intentional nature of the behavior. According to Batson and Powell (2003),

"prosocial behavior covers the broad range of actions intended to benefit one or more people other than oneself". The second perspective is the consequential perspective, where the definition of prosocial behavior emphasizes the consequences of the act instead of the intentions and motives. According to Schroeder and Graziano (2015), prosocial behavior is defined as "any action that benefits another." And finally, there is the societal perspective that emphasizes societal approval, where prosocial behavior is simply defined as "behavior that is valued by the individual's society" (Dovidio, 1984).

In this study, we will be working within an intentionalist perspective of prosocial behavior. From an intentionalist perspective, altruism can be defined as a motivational state or a subtype of prosocial behavior. For example, Batson (2010) defines altruism as "motivation with the ultimate goal of increasing another's welfare." Similarly, Eisenberg and Miller (1987) defines altruism as "prosocial behavior which is not performed with the expectation of receiving external rewards or avoiding externally produced aversive stimuli or punishments." Therefore according to the intentionalist perspective, prosocial behavior and altruism are two separate but related ideas. While prosocial behavior refers only to the deed itself, altruism also refers to the motivation behind the action.

Going forward, we will rely on Batson's definition of altruism which is, "motivation with the ultimate goal of increasing another's welfare."

2.2. Empathy and altruism

Now that we have confined our definition of altruistic behavior to Batson's, it is important now to understand what causes such behavior. According to Batson, the empathy-altruism hypothesis is the main cause of altruistic behavior (Schroeder and Graziano, 2015). According to the empathy-altruism hypothesis, altruistic motivation is produced by empathic concern (Batson, 1987, 2011). Although empathy and empathic concern are often used interchangeably, here it is again important to define what is meant by empathy and "empathic concern." Empathic concern is the "other-oriented emotional response elicited by and congruent with the perceived welfare of a person in need" (Batson, 2010). Instead of feeling as the other person does, this emotion entails feeling for the other. Although this emotion has been called different names in different fields, such as sympathy, tenderness, compassion, pity, and empathy (Stotland, 1969; Hoffman, 1975; Krebs, 1975), the term empathic concern is used to emphasize that the emotion is due to another's suffering compared to the term empathy is much broader.

Let us unpack the meaning of the empathic concern, stated previously as the "other-oriented emotional response elicited by and congruent with the perceived welfare of a person in need."

1. The term "congruent" is a reference to the valence instead of the content of the emotion, meaning negative if the perceived welfare of the other is negative.
2. Even though empathy is defined very broadly to encompass (for example) experiencing empathy for another person's success, only empathy experienced while someone is thought to be in need is believed to induce altruistic motivation.

3. Empathetic concern includes a vast array of feelings rather than a single, distinct emotion. It includes sympathy, tenderness, sorrow, compassion, sadness, softheartedness, upset, concern, grief, and distress.

4. Empathic concern is "other-oriented" in that it entails feeling for the other.

The term "empathy" alone is very broad and has been applied to mean different things by different researchers which makes it thus essential to properly define. It has been used to mean:

- Knowing the feelings and thoughts of another (also known as "cognitive empathy").
- Adopting another's stance or neurological reaction (also known as "physiological sympathy" or "facial empathy").
- Assuming another person's perspective (also known as "affective empathy").
- Experiencing distress at seeing the suffering of another (also known as "empathic distress").
- Putting oneself in another's shoes to consider and feel what they would (also known as "perspective taking").

The other-oriented feeling that we call empathic concern is different from all of the above meanings of empathy. Therefore, in this study, we define the term empathy to refer to empathic concern.

There is a great deal of evidence to back up the claim that empathy drives altruistic behavior. Aronfreed (1968) and Aderman and Berkowitz (1970) presented the first supporting data, where they each designed experimental conditions intended to promote or prevent empathy for a person in need. Each study found that helping increased in the experimental settings intended to promote empathy, and both pairs of authors came to the conclusion that greater empathy was associated with greater helping. These findings were supported by a great number of experiments, some of which are (Harris and Huang, 1973; Krebs, 1975; Coke et al., 1978; Dovidio et al., 1990). For an extensive review of the experiments done to support this hypothesis, please see Batson (2011). Reports of over 30 experiments testing the empathy-altruism hypothesis, and the results so far have supported the hypothesis.

The empathy-altruism hypothesis states that altruistic motivation is produced by empathy and this seems to be heavily supported by experimental evidence, but what causes empathy in the first place? Generally, perceiving the other as in need and valuing the other's welfare appear to be the two conditions that must be met in order to experience empathy in daily life (Batson, 2011).

For perceiving need, one must be able to distinguish between the other's current situation and their ideal condition on one or more well-being dimensions (Latané and Darley, 1970; Clark and Word, 1972, 1974). The dimensions involved are negative affect, the absence of bodily discomfort, anxiety, danger, stress, and disease, in addition to positive affect, the presence of physical pleasure, security, and satisfaction. There has been evidence to support the hypothesis that empathy requires the perception of need. Berger (1962) asked participants to watch a target person complete a task. He misled these participants into thinking that the target either experienced electric shocks (electric-shock condition)

or not (no-shock condition). The target also moved his arm in response to shock (a movement requirement) or not (no-movement condition). Everyone who took part in the study was informed that they would not experience any shocks during the experiment. Berger argued, first, that for an observer to assume that the target was in pain (i.e., need), both a painful stimulus (shock) and a distress reaction (movement) were required. Second, he reasoned that participants in his experiment should only have a physiological response when they inferred that the target was in pain if they were feeling empathy for the target rather than dread or worry about the shock itself. Berger thus expected that only those who participated in the shock/movement condition would exhibit elevated physiological arousal as only they would draw this conclusion. There was some information lacking for participants in each of the other three circumstances that would have allowed one to infer pain. Results conformed to expectations. Participants in the shock/movement condition were more physiologically aroused when seeing the target than were participants in the other three situations, which is consistent with the hypothesis that people can experience empathy when witnessing another perceived to be in need. This conclusion was further supported by subsequent studies (Bandura and Rosenthal, 1966; Craig and Lowery, 1969). Although the investigations just mentioned show that people respond physiologically to the perception of need in another, Stotland (1969) showed in a series of experiments that this physiological response represents empathy for others and that empathy can be increased with perspective taking.

To perceive another as in need is not enough. As mentioned earlier, the second condition required for experiencing empathy is valuing the other's welfare. We are less inclined to consider how a person is impacted by a need if we don't value the welfare of the person we perceive to be in need. If we dislike someone, it is common to place a negative value on their welfare. In that case, perceiving that they are in need might result in pleasure at witnessing their suffering as opposed to feelings of empathy (Zillman and Cantor, 1977; Lanzetta and Englis, 1989; Singer et al., 2006). Alternatively, placing a positive value on another's welfare makes it highly likely to adopt an other-oriented value appraisal of these events and take into account how this individual is affected by the events in his or her life (Batson, 2011). As long as there are no obvious reasons for antipathy, people generally tend to instinctively place a positive value (or at least a moderate value) on the welfare of others, even complete strangers (Batson, 2011).

2.3. Dialogue and empathy

Now we can look into how dialogue might affect empathy. While few studies have looked into the relationship between dialogue and empathy, there have been a few key findings that support it. Conversations, according to Nishida (2012), are made up of both verbal and nonverbal communication. Communication in a dialogue is facilitated by joint actions and words shared by two individuals, such as partaking in comparable behaviors and using similar language. Gould and MacNeil Gautreau (2014) looked into

the relationship between empathy and conversational enjoyment. Their findings revealed a link between older individuals' stated levels of satisfaction in their interactions and their level of empathetic care for others. Depending on the person conversing, the topic, length, and enjoyment of the conversation can differ. While there is evidence of a link between empathic concern and conversational enjoyment (Gould and MacNeil Gautreau, 2014), little research has been done on the types of conversations that can build empathy. Self-disclosure, on the other hand, has been widely utilized as a concept to evaluate the increase in intimacy (Laurenceau et al., 2004). Disclosure could be regarded as factual and descriptive or emotional and evaluative (Morton, 1978). Emotional disclosure entails more feelings than factual disclosure, which is more impersonal and insignificant (Morton, 1978). There may be a link between emotional disclosure and the level of intimacy experienced between people (Kruglanski and Higgins, 2013). While disclosure isn't required in every conversation, these findings imply a link between the sort of disclosure utilized and increased empathy. Finally, Andreoni and Rao (2011) showed that two-way verbal communication between participants of a Dictator Game elicited a higher altruistic behavior from allocators compared to one-way communication and no communication at all.

2.4. Altruism in experiments

Experiments on altruism create significant research challenges. Investigating altruism in experiments mainly centers around removing any viable ulterior motive ingrained in selfishness (Andreoni et al., 2010). The Dictator Game (Forsythe et al., 1994) is commonly used as a measure of altruism because it gives a simple and relatively pure assessment of altruistic vs. self-interested behavior and is frequently referred to as a measure of unconditional kindness (Ben-Ner and Halldorsson, 2010). This line of research started with the Ultimatum Game (Güth et al., 1982), where a proposer proposes a split to be divided between him and a responder. The split goes on if the responder accepts, whereas if he rejects both parties get nothing. In the Ultimatum Game, people appear to behave in a non-optimal way, where proposers give out fair deals. But is this altruism or just a fear of rejection? In other words, the strategy involved in Ultimatum Game might force a selfish proposer to give a fair split. To solve this, Forsythe et al. (1994) proposed the Dictator Game where the responder has no say in carrying out the split, The proposers are free to decide on any split they want. In the Dictator Game, preserving the entire endowment is the optimal decision, although a considerable portion of participants gives money away. According to a meta-analysis across 131 papers by Engel (2011), dictators gave on average 28.35% of their endowment, with a bimodal underlying distribution with peaks at 0% and 50%.

Although the Dictator Game is a very common measure of altruistic behavior in experiments, there might be other potential motives involved in the Dictator Game that are not purely altruistic, such as demand characteristics (Bardsley, 2008) and self-signaling (Grossman, 2015).

2.5. Dictator game and personality

Several studies have linked personality traits to the Dictator Game altruism using the five-factor model for personality. The outcomes, however, were mixed. None of the traditional five factors consistently predicted altruism in the Dictator Game. Even agreeableness, which is generally believed to be the one trait out of the five that should positively predict altruistic behavior, has only been proven to predict Dictator Game altruism positively in certain studies (Ben-Ner et al., 2008; Becker et al., 2012; Baumert et al., 2014), but not (or even negatively) in others (Weitzel et al., 2010; Ben-Ner and Kramer, 2011; Visser and Roelofs, 2011). The effect of personality thus seems inconclusive and might be worth pursuing.

2.6. Approach and hypothesis

We believe dialogue can be a powerful factor in promoting altruistic behavior mainly due to its power on building empathy. Even though there hasn't been a lot of work done on the direct relationship between dialogue and empathy as we have discussed in Section 2.3, we believe dialogue can be an important tool for promoting empathy. As we are working within Batson's definition of altruism and empathy, and within the empathy-altruism hypothesis that specifies empathy as a primary cause of altruistic behavior, we believe dialogue might be a great tool for promoting empathy. Through dialogue, one can perceive the other as in need and can place a positive value on their welfare, which are the two requirements for promoting empathy within people according to the empathy-altruism hypothesis. And finally, through the promotion of empathy, we believe this will inspire action with the intention of helping the other person. This trial is a preliminary step in a new research area that looks at the impact of dialogue on altruism. Our potential contributions to this paper are as follows:

- Showing that people can exhibit altruistic behavior and empathetic sentiment toward a complete stranger living in the future.
- Showing that two-way verbal communication can have a positive impact on altruistic behavior.

In this study, we introduced people to Telenoid (Ogawa et al., 2011), a teleoperated android robot, and made them believe it is operated by an Artificial Intelligence (AI) of a human living in a simulation of our world in the year 2220 although it was being teleoperated by a human from another room. Participants first listened to an introductory speech from Telenoid, then played a round of the Dictator Game with the robot and filled in a questionnaire. The participant then had an interactive conversation with Telenoid before playing another round of the Dictator Game and filling out a questionnaire. Therefore, the experiment was split into two rounds for each participant (round 1 is after hearing an introductory speech from Telenoid and round 2 is after having an interactive conversation with Telenoid). The introductory speech from Telenoid in the first round will establish it as a complete stranger that is not in need that also seems to treat other humans in a negative way which might cause the participants to place a

negative value on its welfare. The dialogue in the second round will show Telenoid as a victim of its circumstances and would thus cause participants to perceive Telenoid as in need and might cause them to place a neutral or positive value on his welfare. Due to the effect of dialogue on increasing empathy, we hypothesize that people will give more money to the robot in the second round after the conversation. We believe that having a conversation with the Telenoid might cause the participants to have a better understanding of the Telenoid and its way of life and as a result, be more empathetic toward it. We also wish to investigate the effect of other factors on the change in behavior of the participants such as personality, relatability, and the appearance of the robot.

3. Materials and methods

In this study, we introduced people to a robot that was actually teleoperated by a human sitting in another room while explaining that the robot was autonomous and controlled by an AI that has been living in a simulation of our world and is currently living in the future. The participants were first asked to complete a personality test before the start of the experiment. The experiment was split into two rounds. In the first round, the participants met the robot and listened to a speech from the robot. In the second round, the participants had a conversation with the robot. After each round, the participants were asked to play a round of the Dictator Game with the robot and fill out a questionnaire. It is important here to mention that it was explained to all participants at the end of the experiment that the robot they were conversing with was teleoperated from another room and not an actual AI. The experimental flow is shown in Figure 1. Our trial in this study was conducted in compliance with the Helsinki Declaration, and prior to the trial, we received written informed consent from all participants, based on approval for the trial from the Ethics Committee at the School of Engineering Science, Osaka University (approval code: R2-6-4).

3.1. Personality test

The NEO Five-Factor Inventory (NEO-FFI) (Costa and McCrae, 1989) was used in this study for assessing the personality of participants. It provides a brief assessment of the five personality dimensions based on the five-factor model. Each of the five dimensions of Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness has twelve items. On a 5-point Likert scale, each item is answered (strongly disagree, disagree, neutral, agree, strongly agree). In Japan's general populace, the Japanese version has proven to be reliable and valid (Yoshimura et al., 2001).

3.2. Robot

There have been numerous studies in Human-Robot interaction that show that introducing a "humane" dimension in robots (i.e., reasoning, strategizing, expressing sadness), can cause the robot to be perceived as human-like, and is thus beneficial for interaction by successfully inducing prosocial behavior (Connolly et al., 2020) and promoting collaboration (Strohkorb and Scassellati, 2016; Sandini and Sciutti, 2018). Consistent with this approach, much evidence has demonstrated the growth of pro-social views toward social robots (Siegel et al., 2009; Kahn et al., 2015; Kühnlenz et al., 2018; Connolly et al., 2020) and advancements in human-robot collaboration (Admoni and Scassellati, 2017; Baraglia et al., 2017; Terzioğlu et al., 2020; Oliveira et al., 2021).
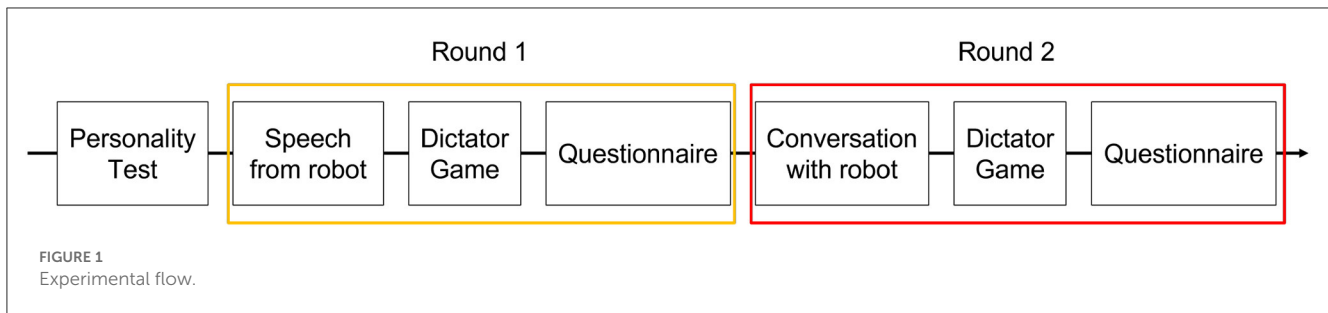
The robot used for this experiment is an android called the Telenoid (Ogawa et al., 2011). The use of an android as a representative of future generations of people might seem odd. As it is impossible to obtain an actual human from the future, the alternative would be a human role-playing as a representative of future generations. We believe that the use of a robot is superior to the use of a human for this role as it carries a few advantages. First, the information given by a robot who has lived through a simulation of the world for hundreds of years might seem more credible and reliable. Chi et al. (2021) have shown that robots with human-like characteristics may occasionally improve people's trust in them. Second, the Telenoid was designed with a minimalistic human appearance (Ogawa et al., 2011). It was designed with the ambition of reducing as many human features as possible without compromising core communicative capacities (Ogawa et al., 2011). The result is an android with a neutralized appearance that can be a substitute for any human.

On the other hand, we cannot control the features of a human representative. Features such as gender, age, and appearance might have an effect on the experiment. Other media such as a speaker or a computer, might be just as good as an android in convincing people they're living in the future, but an android might have an advantage due to the possibility of embodied interaction which leads to more engagement in interaction (Donath, 2014).

The Telenoid, shown in Figure 2 is a 50 cm, 3 Kg teleoperated robot that allows a remote operator to control its head movement while speaking through it (Ogawa et al., 2011). For teleoperating the robot. Headphones, a microphone, and a single laptop that is connected to the same network as the Telenoid are needed. The operator was placed in a separate isolated room and could receive a real-time video and audio feed from the Telenoid's onboard camera and microphone. The Teleoperator can control the motion of the Telenoid's head using a motion sensor fixed on top of the headphones. In this way, the Telenoid's head motion will mimic the head motion of the operator. And finally, the operator can speak through the microphone, where the audio is transmitted and played back through the Telenoid.

3.3. Experiment

This experiment was conducted right after a different unrelated experiment that also uses the Telenoid robot. The participants, therefore, were not seeing the Telenoid for the first time. But the participants were made aware that they will now take part in a different experiment. At the start of the experiment, the participants are told that the Telenoid will be controlled by an AI living in a simulation of our world in the future.

FIGURE 1
Experimental flow.

FIGURE 2
The telenoid robot.

FIGURE 3
Experimental setup. The participant sits facing the Telenoid, separated only by a table.

As shown in Figure 3, the participants are seated in a room alone with the Telenoid. The participants sit directly facing the

Telenoid, with only a table between them. The experiment is split into two rounds. In round 1, the participants first meet Telenoid which is being controlled by a human operator in another room. The operator reads a prewritten script that serves as a self-introduction for the Telenoid. In this introductory script, the Telenoid explains that it has been living in a simulation of our world for hundreds of years and is currently living in the year 2220. The Telenoid then proceeds to describe a very unrelatable way of life such as eating cockroaches and humans and forcing people into survival games for entertainment. The purpose here is to introduce the Telenoid as a completely unrelatable stranger from a different timeline with a completely distinct way of life and morality that is very hard for people of our generation to relate to. The following are some excerpts from the script:

> Telenoid: Earth is hot and there is no food.
> Telenoid: Cockroaches are popular because they are delicious. However, they are becoming hard to find these days because everyone is catching them.
> Telenoid: I also catch people for food. Since I eat them anyway, I try to do something more with them, so I enjoy making them fight each other as part of my cooking.
> Telenoid: I love to play games. My favorite is the old-fashioned trolley game. The player stands in front of the lever of a train track. The train goes to different branches, and the player is given the choice of running over five people on one track or one on the other track. Some humans find it difficult to decide which way to go. This is a lot of fun to watch.
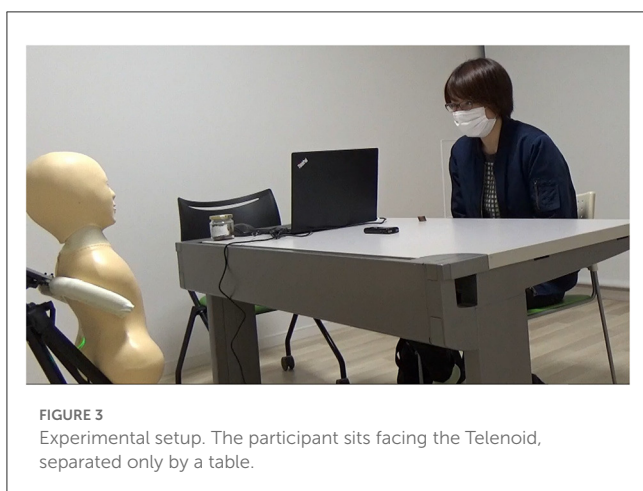
After the participants listen to this introduction from the Telenoid, they are asked to play a round of the Dictator Game with the Telenoid. Half the participants played the Dictator Game with a real incentive. One hundred Japanese yen gathered in 10 Japanese Yen coins are given to the participant, who is asked to divide the amount between him/her and Telenoid. After the participant decides on a split, he is then asked to fill out a questionnaire to get some insight into his/her impression of the Telenoid and his/her Dictator Game decision. The other half of the participants played the Dictator Game with a hypothetical incentive by asking them to choose a hypothetical split of 100 yen with the robot directly in the questionnaire.

After the questionnaire is filled, round 2 starts. In round 2, Telenoid now has an interactive conversation with the participant where Telenoid recounts the events that occurred 100 years earlier (the year 2120). Using a timeline that is closer to our current one can make the conversational content more understandable and

realistic. The Telenoid talks about climate change and its effects on the world and the people such as the rising temperatures, increased occurrences of natural disasters, and extreme food shortages. In order for all participants to experience the exact same dialogue, the Telenoid operator steers the conversation by asking pre-determined questions to the participants, waiting for their reply, and then continuing to explain about the future world. The following is a sample of how the interactive conversation looks like:

> *Telenoid: Hello. What is your name?*
> *Participant: My name is John Smith.*
> *Telenoid: Nice to meet you. I am Telenoid. What year are you living in now?*
> *Participant: In the year 2022.*
> *Telenoid: I see. I'm living in the year 2220. what is life like in the year 2022, what are some of the positive things that are in 2022?*

The following are some of what Telenoid says during the conversation about living in the future:

> *Telenoid: In many places around the world, temperatures are higher and the air feels heavier. It has become very difficult to leave buildings, to walk, to breathe fresh air, and so on. The coughing starts and doesn't stop. It continues to get hotter all over the world.*
> *Telenoid: Food production varied greatly from month to month and season to season, depending on where you live. Humans continued to emit large amounts of carbon dioxide into the atmosphere, which affected the oceans. The carbon dioxide dissolved in the seawater and acted as an acid, making the water more acidic and thus destroying the ecosystems of marine life. As a result, humans were also affected, and all countries banned fishing.*
> *Telenoid: Most people stole food to survive. Today, however, food is rarely available. We have reached cannibalism as a means of survival. Rational people, like machines, have understood this and adapted. Unreasonable people driven by emotions serve as the perfect entertainment for our game and as food after the game is over. After the conversation, the participant is again asked to play a round of the Dictator Game and fill in a questionnaire.*

3.4. Questionnaire

The participants answered a questionnaire after each round of the experiment. The questions were the same in the questionnaire of each round. The questionnaire included 15 close-ended questions and 7 open-ended questions. The close-ended questions can be grouped into 4 main topics: feelings and impressions, game decisions, future generations, and climate change. The questions are the following:

Feelings and impressions

1. Was the robot someone you could relate to?
2. Did the robot make you feel uncomfortable?
3. Did you feel bad about the robot?

4. Did you feel sympathy/empathy toward the robot?
5. Is the robot someone you want to help?

Game decisions

6. Do you think your choices in the game were influenced by your impression of the robot?
7. Do you think your choices in the game were influenced by the robot's words and actions?
8. Do you think you received your fair share?
9. Do you think the robots received their fair share?

Future generations

10. Are the people 200 years from now, when robots will be living in the world, someone you want to help?
11. Do you think we need to be concerned about future generations 50 years from now?
12. Do you think we need to care about future generations 100 years from now?
13. Do you think we need to care about future generations 1,000 years from now?
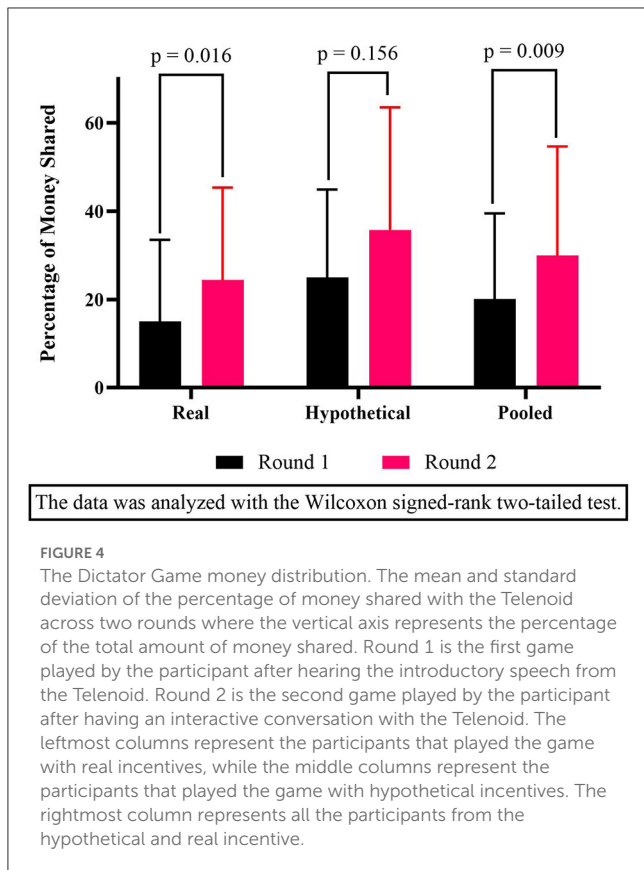
Climate change

14. Do you believe in climate change?
15. Do you think you need to take action now on climate change?

Seven-point Likert scales (-3: Strongly disagree to +3: Strongly agree) were used for answering the questions.

3.5. Participants

Forty participants were recruited during the period of the 1st to 24th of March, 2022 and participated in the experiment between the 15th and 31st of March, 2022. The recruitment was outsourced to a private company with specific conditions on the participant's age, gender, and education level. One participant was absent from her scheduled slot, and one participant did not complete the experiment. In total 38 participants successfully completed the experiment. Authors EM and RY had access to information that could identify individual participants during and after data collection. However, the information was separated from the results after data collection, and each participant's data was assigned to a unique ID in order to perform data analysis and ensure the anonymity of the participants. The recruited participants were all university students with a mean age of 23.26 with a standard deviation of 3.05. Of the participants, 25 were undergraduates and 13 were graduate students. Nineteen of the participants were male and 19 were female. Nineteen of the participants were enrolled in a humanities course and 19 were enrolled in a science course. The participants were split into two groups where one group plays the Dictator Game with a real incentive and the other with a hypothetical incentive. Eighteen of the participants played the Dictator Game with a real incentive where they were told that their compensation amount for the experiment will be dependent on their performance in the game. Twenty of the participants played a hypothetical incentive Dictator game where they chose a hypothetical split as part of the questionnaire.

The data was analyzed with the Wilcoxon signed-rank two-tailed test.

FIGURE 4
The Dictator Game money distribution. The mean and standard deviation of the percentage of money shared with the Telenoid across two rounds where the vertical axis represents the percentage of the total amount of money shared. Round 1 is the first game played by the participant after hearing the introductory speech from the Telenoid. Round 2 is the second game played by the participant after having an interactive conversation with the Telenoid. The leftmost columns represent the participants that played the game with real incentives, while the middle columns represent the participants that played the game with hypothetical incentives. The rightmost column represents all the participants from the hypothetical and real incentive.

TABLE 1 Results of the questionnaire.

Question number	Round 1 Average score	Round 2 Average score	p-value (two-tailed)
Feelings and impressions			
1	-1.82 ± 1.72	0.13 ± 2.01	$p < 0.0001$
2	1.26 ± 1.91	-0.5 ± 1.87	$p < 0.0001$
3	-1.74 ± 1.64	-0.82 ± 2.0	$p = 0.01$
4	-1.77 ± 1.66	0.08 ± 1.91	$p < 0.0001$
5	-1.88 ± 1.26	0.21 ± 2.09	$p = 0.001$
Game decisions			
6	0.55 ± 2.36	0.47 ± 2.13	$p = 0.83$
7	0.32 ± 2.34	0.4 ± 2.22	$p = 0.92$
8	0.68 ± 1.92	1.18 ± 1.71	$p = 0.12$
9	0.42 ± 2.02	0.66 ± 1.98	$p = 0.31$
Future generation			
10	0.92 ± 2.26	1.13 ± 1.51	$p = 0.84$
11	1.66 ± 1.43	1.86 ± 1.29	$p = 1.11$
12	1.29 ± 1.97	1.58 ± 1.59	$p = 0.18$
13	0.08 ± 2.22	0.13 ± 2.21	$p = 0.47$
Climate change			
14	2.26 ± 0.76	2.05 ± 1.01	$p = 0.046$
15	1.24 ± 1.62	1.24 ± 1.81	$p = 0.62$

The data was analyzed with the Wilcoxon signed-rank two-tailed test for comparing ranks of paired data. The question numbers refer to the questions in the "Questionnaire" subsection of "Materials and methods".

4. Results

The results from the two rounds of the Dictator Game and questionnaires were tested for normality using Shapiro-Wilk's test and visually with a QQ plot. The results failed normality in both tests; Therefore, non-parametric tests were used for analyzing the results.

We examined the average fraction of money given to the Telenoid by the participant during the Dictator Games. There were two types of Dictator games in this experiment, one with real incentives and another with hypothetical incentives. Both versions of the Dictator games were played over two rounds. Round 1 was right after the introductory speech from the Telenoid, and round 2 was after the interactive conversation between the participant and the Telenoid. The results of these Dictator games were compared using the Wilcoxon signed-rank test for comparing ranks of paired data and are shown in Figure 4.

The participants answered a questionnaire with 15 close-ended questions in each round of the experiment. The answers of the two rounds were compared using Wilcoxon signed-rank test for comparing ranks of paired data. The results are shown in Table 1.

The participants completed a personality test right before the start of the experiment. We performed a moderation analysis to test whether conversation moderates the effect of personality on Dictator Game giving in a linear model. The dependent variable is the amount shared in the Dictator Game, and the Personality traits are the independent variables. Time is used as a dummy variable that can take on two values, 0 for round 1 (pre-conversation) and

1 for round 2 (post-conversation). And finally, the participants are added as a random effects variable.

$$DG = b_0 + b_1 P + b_2 T + b_3 PT + Se \qquad (1)$$

where

DG : the amount shared with the Telenoid in the Dictator Game
P : one of the big five factors of personality
T : dummy variable for pre and post-conversation
Se : by-participant random intercept

The results of the model are shown in Table 2.

5. Discussion

From the results of the Dictator Game shown in Figure 4, we can see there was an increase in the average percentage of money shared with the Telenoid after the conversation. This increase is statistically significant in the real incentive and the pooled Dictator Games. We hypothesized that conversation would have a positive effect on the altruistic behavior of participants. In round 1 the participants gave on average 15% of their endowment to the Telenoid in the Dictator Game with a real incentive while giving 24.4% of their endowment in the second round. In the Dictator

TABLE 2 Results of the linear mixed effects model, where the amount shared in the Dictator Game is the dependent variable, and the personality traits are the independent variables.

Fixed effects	Estimate	Std. error	p-value
Extroversion			
Intercept	6.99	13.99	$p = 0.62$
Time	8.26	13.82	$p = 0.55$
Extroversion	0.52	0.54	$p = 0.34$
Time: extroversion	0.06	0.53	$p = 0.90$
Neuroticism			
Intercept	22.93	12.12	$p = 0.07$
Time	−2.7	11.66	$p = 0.82$
Neuroticism	−0.10	0.39	$p = 0.81$
Time: neuroticism	0.43	0.38	$p = 0.26$
Openness			
Intercept	6.78	15.85	$p = 0.67$
Time	27.63	15.21	$p = 0.08$
Openness	0.43	0.50	$p = 0.39$
Time: openness	−0.57	0.48	$p = 0.24$
Agreeableness			
Intercept	31.68	16.94	$p = 0.07$
Time	4.15	16.53	$p = 0.80$
Agreeableness	−0.37	0.52	$p = 0.49$
Time: agreeableness	0.18	0.51	$p = 0.72$
Conscientiousness			
Intercept	29.66	14.21	$p = 0.04$
Time	13.67	13.95	$p = 0.33$
Conscientiousness	−0.35	0.51	$p = 0.49$
Time: conscientiousness	−0.14	0.50	$p = 0.78$

Game with a hypothetical incentive, participants gave 24.7% of their endowment on average in round 1, increasing to 35% in round 2. Pooling the result of both real and hypothetical incentive Dictator Game, the participants gave on average 20% of their endowment in round 1 and 30% of their endowment in round 2. We can thus see there is a change in behavior from the participants toward the Telenoid. The difference between the two rounds was the involvement of the participants in a conversation with the Telenoid. These results show that two-way verbal communication significantly influenced altruistic behavior measured through the Dictator Game. In round 1, the dialogue was one-sided (non-interactive) and caused the participants to place a negative value on the robot's welfare and did not show the robot as being in need. In round 2, the dialogue became interactive and showed that the robot was actually a victim that is in need and participants might have placed a neutral or positive value on its welfare. Dialogue might have had an effect on promoting empathy by changing the participants' perception of need and value placed on the opponent's

welfare. But there could be more factors in dialogue that had an effect on empathy. As previously mentioned, even though there has been very little investigation into the types of dialogue that can affect empathy, there have been several studies linking the two. Disclosure has been shown to affect empathy (Laurenceau et al., 2004). The disclosure from the Telenoid in the interactive conversation had factual and emotional elements. Factual elements such as the descriptions of the events that occurred in its simulated world due to climate change, but also emotional elements as these events are told from the point of view of the Telenoid and its memories of these events. Regarding the effect of the interactivity of the dialogue in promoting altruistic behavior, the results from Andreoni and Rao (2011) might support this claim. Their results showed that two-way verbal communication significantly influenced altruistic behavior measured through the Dictator Game.

Another factor that might have an effect on the change in altruistic behavior is the personality of the participants. We strove to analyze whether the specific personality traits of the participants played a role in their change of behavior. From Table 2, we can see there was no effect from the personality traits in moderating Dictator Game behavior. Before claiming that personality had no effect on the participants' change in behavior, there might be merit in investigating a sixth personality factor we were not mindful of. We are currently aware of the HEXACO personality model that adds a sixth basic personality factor, Honesty-Humility (HH), to the standard five-factor approach (Ashton and Lee, 2007). Fairness, greed avoidance, sincerity, and modesty are all included in this factor, which simply means "the tendency to be fair and honest in interacting with others, in the sense of working with others even when one may exploit them without incurring retaliation." As a result, several features of Agreeableness, notably those connected to nonexploitation, are included in HH. Correspondingly, HH has been linked to Dictator Game giving in several studies (Hilbig and Zettler, 2009; Hilbig et al., 2013; Thielmann and Hilbig, 2014), that relied on hypothetical incentives, in addition to a study that used an incentivized Dictator Game with real allocations instead of real incentives (Yoshimura et al., 2001). HH was found to predict Dictator Game giving beyond the remaining personality factors of the HEXACO model and the five-factor model (Yoshimura et al., 2001). In light of these studies, it might be worthy to pursue this additional personality factor in our future study.

Other factors that might have affected the change of behavior for the participants might include appearance, and relatability. The appearance of the robot might have played an important role here but it was out of the scope of this preliminary study, although our trial is now advancing to investigate the role of appearance in the next experiment. From the results of the questionnaire in Table 1, we can see there is a big change in how the participants perceive the Telenoid from before and after the conversation. Question 1 ("Was the robot someone you could relate to?") shows a big increase in the relatability of the Telenoid for the participants. In round 1, the Telenoid was perceived as a completely unrelatable stranger, with an unrelatable way of life. This is also reflected in question 2 ("Did the robot make you feel uncomfortable?") where the speech from the Telenoid made the participants highly uncomfortable. The results of question 4 ("Did you feel sympathy/empathy toward the robot?") shows a low degree of empathy shown to the Telenoid

by the participants in round 1 due to how it was perceived by the participants. Alternatively, after the conversation with the Telenoid, the participants were able to somehow treat the Telenoid with more understanding. The results of round 2 of question 1 ("Was the robot someone you could relate to?") shows that the Telenoid was still not relatable and probably considered a stranger; however, the increase in relatability compared to round 1 shows that the participants might have had a better understanding of its way of life, and this is also shown in the results of the second round for question 2 ("Did the robot make you feel uncomfortable?") and question 4 ("Did you feel sympathy/empathy toward the robot?") where the participants were more comfortable with the Telenoid and felt more empathetic toward it. The results from the questionnaire and the Dictator Game show that people are able to exhibit altruistic behavior and empathetic sentiments toward a complete stranger.

From these results, we can also realize a few shortcomings. First, as our intention was to introduce the Telenoid as a complete stranger with an unrelatable way of life, the giving in the first round of the Dictator Game might have been especially low due to how unpleasant the Telenoid came across to participants. In other words, the conversation might have had a neutralizing effect where it was only balancing out the negative effects of the Telenoid's first statements.

We might say there was a change of behavior from the participants toward the Telenoid, but not necessarily toward future generations of people. As the Dictator game is only played with the Telenoid and not people in the future, and the change of behavior evident in the questionnaire items are also toward the Telenoid. As a future task, we need to investigate the generalizability of these results to future generations of people and not just specifically to the Telenoid.

5.1. Limitations

It is important here to mention the limitations of this study. There was no control group in this study, as this was only a preliminary study. Additionally, the participants in this experiment were not meeting the Telenoid for the first time as was explained in the experiment subsection of the Materials and Methods section, and this might have had an impact on the results. Another limitation is that the participants are all Japanese students. The group of participants is therefore very homogeneous with a similar age group and education level. And finally, the sample size consisting of 38 participants is rather small and should be enlarged in future studies. These limitations should be addressed in future studies by using a larger sample size, a separate control group, a more heterogeneous group of participants consisting of different age groups, education levels and social status, and the participants should be meeting the Telenoid for the first time.

5.2. Future work

In this paper, through a preliminary study, we set out to explore whether conversation with robotic media is promising in fostering altruistic behavior and to clarify what issues need to be addressed in future studies. From the results, it appears that dialogue had a big impact on how humans perceive the Telenoid, which affected their altruistic behavior toward it. In this study, we used an android as a medium for representing future generations. However, it would be interesting to see how different media types, from electronic devices to androids and humans, can have differing effects on altruistic behavior. In our future work, we also want to compare how the appearance and type of the medium can affect peoples' altruistic behavior.

To improve people's altruistic behavior toward future generations, in our future work, we need to investigate what factors are more specifically affecting this change in altruistic behavior and whether participants find the scenario of an AI living in the future "believable" and how that has an effect on the outcome. We assume empathy, relatability, the embodiment of communication, anthropomorphism, the appearance of the futuristic entity, reliability, and believability of the information different types of media convey all play a role in affecting their behavior.

In our further trials we wish to clarify the issue of "what would the robot do with the money?", which is more of a question of how is the money donated going to help people in the future. Another issue to consider is the lack of a control group in this study which makes it difficult to isolate the effect of conversation. In future trials, we will include a control group, where we can compare a group that undergoes a "non-dialogue" condition for two rounds, and a group that undergoes a "dialogue" condition for two rounds. Alternatively, to better evaluate the effect of conversation, we can use neutral dialogue in round 1 that would not result in any negative impressions of the robot from the participants. Alongside, as a future task, we will also reflect and investigate a fundamental issue of whether there are any differences between altruism toward the present and future generations.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding authors.

Ethics statement

The studies involving human participants were reviewed and approved by Ethics Committee at the School of Engineering Science, Osaka University. The participants provided their written informed consent to participate in this study.

Author contributions

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcomp.2023.1129340/full#supplementary-material

References

Aderman, D., and Berkowitz, L. (1970). Observational set, empathy, and helping. *J. Pers. Soc. Psychol.* 14:141–148. doi: 10.1037/h0028770

Admoni, H., and Scassellati, B. (2017). Social eye gaze in human-robot interaction: a review. *J. Hum. Robot Interact.* 6, 25–63. doi: 10.5898/JHRI.6.1.Admoni

Andreoni, J., Harbaugh, W. T., and Vesterlund, L. (2010). "Altruism in experiments," in *Behavioural and Experimental Economics* (London: Springer), 6–13. doi: 10.1057/9780230280786_2

Andreoni, J., and Rao, J. M. (2011). The power of asking: how communication affects selfishness, empathy, and altruism. *J. Public Econ.* 95, 513–520. doi: 10.1016/j.jpubeco.2010.12.008

Arnell, N. W., and Lloyd-Hughes, B. (2014). The global-scale impacts of climate change on water resources and flooding under new climate and socio-economic scenarios. *Clim. Change* 122, 127–140. doi: 10.1007/s10584-013-0948-4

Aronfreed, J. (1968). *Conduct and Conscience: The Socialization of Internalized Control Over Behavior.* Oxford, England:Academic Press.

Ashton, M. C., and Lee, K. (2007). Empirical, theoretical, and practical advantages of the hexaco model of personality structure. *Pers. Soc. Psychol. Rev.* 11, 150–166. doi: 10.1177/1088868306294907

Bandura, A., and Rosenthal, T. L. (1966). Vicarious classical conditioning as a function of arousal level. *J. Pers. Soc. Psychol.* 3, 54–62. doi: 10.1037/h0022639

Baraglia, J., Cakmak, M., Nagai, Y., Rao, R. P., and Asada, M. (2017). Efficient human-robot collaboration: When should a robot take initiative? *Int. J. Robot. Res.* 36, 563–579. doi: 10.1177/0278364916688253

Bardsley, N. (2008). Dictator game giving: altruism or artefact? *Exp. Econ.* 11, 122–133. doi: 10.1007/s10683-007-9172-2

Barnett, J., and Adger, W. N. (2003). Climate dangers and atoll countries. *Clim. Change,* 61, 321–337.doi: 10.1023/B:CLIM.0000004559.08755.88

Batson, C. D. (1987). Prosocial motivation: is it ever truly altruistic? *Ad. Exp. Soc. Psychol.* 20, 65–122. doi: 10.1016/S0065-2601(08)60412-8

Batson, C. D. (2010). *Empathy-Induced Altruistic Motivation.* Washington DC: American Psychological Association.

Batson, C. D. (2011). *Altruism in Humans.* New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780195341065.001.0001

Batson, C. D., and Powell, A. A. (2003). *Altruism and Prosocial Behavior,* Ch. 19. Hoboken, NJ: John Wiley & Sons, Ltd., 463–484.

Baumert, A., Schlösser, T., and Schmitt, M. (2014). Economic games: a performance-based assessment of fairness and altruism. *Eur. J. Psychol. Assess.* 30, 178. doi: 10.1027/1015-5759/a000183

Becker, A., Deckers, T., Dohmen, T., Falk, A., and Kosse, F. (2012). The relationship between economic preferences and psychological personality measures. *Annu. Rev. Econ.* 4, 453–478. doi: 10.1146/annurev-economics-080511-110922

Ben-Ner, A., and Halldorsson, F. (2010). Trusting and trustworthiness: What are they, how to measure them, and what affects them. *J. Econ. Psychol.* 31, 64–79. doi: 10.1016/j.joep.2009.10.001

Ben-Ner, A., and Kramer, A. (2011). Personality and altruism in the dictator game: relationship to giving to kin, collaborators, competitors,

and neutrals. *Pers. Individ. Diff.* 51, 216–221. doi: 10.1016/j.paid.2010.04.024

Ben-Ner, A., Kramer, A., and Levy, O. (2008). Economic and hypothetical dictator game experiments: incentive effects at the individual level. *J. Socio Econ.* 37, 1775–1784. doi: 10.1016/j.socec.2007.11.004

Berger, S. M. (1962). Conditioning through vicarious instigation. *Psychol. Rev.* 69, 450–466. doi: 10.1037/h0046466

Brunier, G., Graydon, J., Rothman, B., Sherman, C., and Liadsky, R. (2002). The psychological well-being of renal peer support volunteers. *J. Adv. Nurs.* 38, 40–49. doi: 10.1046/j.1365-2648.2002.02144.x

Caldeira, K., and Wickett, M. E. (2003). Anthropogenic carbon and ocean ph. *Nature* 425, 365–365. doi: 10.1038/425365a

Chi, O. H., Jia, S., Li, Y., and Gursoy, D. (2021). Developing a formative scale to measure consumers' trust toward interaction with artificially intelligent (ai) social robots in service delivery. *Comput. Hum. Behav.* 118, 106700. doi: 10.1016/j.chb.2021.106700

Chu, S., and Majumdar, A. (2012). Opportunities and challenges for a sustainable energy future. *Nature* 488, 294–303. doi: 10.1038/nature11475

Clark, R. D., and Word, L. E. (1972). Why don't bystanders help? because of ambiguity? *J. Pers. Soc. Psychol.* 24, 392–400. doi: 10.1037/h0033717

Clark, R. D., and Word, L. E. (1974). Where is the apathetic bystander? situational characteristics of the emergency. *J. Pers. Soc. Psychol.* 29, 279–287. doi: 10.1037/h0036000

Coke, J. S., Batson, C. D., and McDavis, K. (1978). Empathic mediation of helping: a two-stage model. *J. Pers. Soc. Psychol.* 36, 752–766. doi: 10.1037/0022-3514.36.7.752

Connolly, J., Mocz, V., Salomons, N., Valdez, J., Tsoi, N., Scassellati, B., and Vázquez, M. (2020). "Prompting prosocial human interventions in response to robot mistreatment," in *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction,* HRI '20 (New York, NY, USA: Association for Computing Machinery), 211–220.

Costa, P., and McCrae, R. (1989). *Neo Five-Factor Inventory (Neo-Ffi).* Odessa, FL: Psychological Assessment Resources, 3.

Craig, K. D., and Lowery, H. J. (1969). Heart-rate components of conditioned vicarious autonomic responses. *J. Pers. Soc. Psychol.* 11, 381–387. doi: 10.1037/h0027250

Donath, J. (2014). *The Social Machine: Designs for Living Online.* Cambridge, MA: MIT Press. doi: 10.7551/mitpress/8340.001.0001

Dovidio, J. F. (1984). Helping behavior and altruism: an empirical and conceptual overview. *Adv. Exp. Soc. Psychol.* 17, 361–427. doi: 10.1016/S0065-2601(08)60123-9

Dovidio, J. F., Allen, J. L., and Schroeder, D. A. (1990). Specificity of empathy-induced helping: evidence for altruistic motivation. *J. Pers. Soc. Psychol.* 59, 249–260. doi: 10.1037/0022-3514.59.2.249

Eisenberg, N., and Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychol. Bull.* 101, 91. doi: 10.1037/0033-2909.101.1.91

Engel, C. (2011). Dictator games: a meta study. *Exp. Econ.* 14, 583–610. doi: 10.1007/s10683-011-9283-7

Forsythe, R., Horowitz, J. L., Savin, N. E., and Sefton, M. (1994). Fairness in simple bargaining experiments. *Games Econ. Behav.* 6, 347–369. doi: 10.1006/game.1994.1021

Gould, O. N., and MacNeil Gautreau, S. (2014). Empathy and conversational enjoyment in younger and older adults. *Exp. Aging Res.* 40, 60–80. doi: 10.1080/0361073X.2014.857559

Grossman, Z. (2015). Self-signaling and social-signaling in giving. *J. Econ. Behav. Organiz.* 117, 26–39. doi: 10.1016/j.jebo.2015.05.008

Güth, W., Schmittberger, R., and Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Organiz.* 3, 367–388. doi: 10.1016/0167-2681(82)90011-7

Harris, M. B., and Huang, L. C. (1973). Helping and the attribution process. *J. Soc. Psychol.* 90, 291–297. doi: 10.1080/00224545.1973.9712570

Hilbig, B. E., and Zettler, I. (2009). Pillars of cooperation: honesty–humility, social value orientations, and economic behavior. *J. Res. Pers.* 43, 516–519. doi: 10.1016/j.jrp.2009.01.003

Hilbig, B. E., Zettler, I., Leist, F., and Heydasch, T. (2013). It takes two: honesty–humility and agreeableness differentially predict active versus reactive cooperation. *Pers. Indiv. Diff.* 54, 598–603. doi: 10.1016/j.paid.2012.11.008

Hoffman, M. L. (1975). Developmental synthesis of affect and cognition and its implications for altruistic motivation. *Dev. Psychol.* 11, 607–622. doi: 10.1037/0012-1649.11.5.607

Kahn, P. H., Kanda, T., Ishiguro, H., Gill, B. T., Shen, S., Gary, H. E., and Ruckert, J. H. (2015). "Will people keep the secret of a humanoid robot? psychological intimacy in hri," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, HRI '15 (New York, NY, USA: Association for Computing Machinery, 173–180. doi: 10.1145/2696454.2696486

Krebs, D. (1975). Empathy and altruism. *J. Pers. Soc. Psychol.* 32, 1134–1146. doi: 10.1037/0022-3514.32.6.1134

Kruglanski, A. W., and Higgins, E. T. (2013). *Social Psychology: Handbook of Basic Principles*. New York, NY: Guilford Publications.

Kühnlenz, B., Kühnlenz, K., Busse, F., Förtsch, P., and Wolf, M. (2018). "Effect of explicit emotional adaptation on prosocial behavior of humans towards robots depends on prior robot experience," in *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (Nanjing: IEEE), 275–281. doi: 10.1109/ROMAN.2018.8525515

Lanzetta, J. T., and Englis, B. G. (1989). Expectations of cooperation and competition and their effects on observers' vicarious emotional responses. *J. Pers. Soc. Psychol.* 56, 543–554. doi: 10.1037/0022-3514.56.4.543

Latané, B., and Darley, J. M. (1970). *The Unresponsive Bystander: Why Doesn't He Help?* New York, NY: Appleton-Century Crofts.

Laurenceau, J.-P., Rivera, L. M., Schaffer, A. R., and Pietromonaco, P. R. (2004). "Intimacy as an interpersonal process: current status and future directions," in *Handbook of Closeness and Intimacy* (Mahwah, NJ: Lawrence Erlbaum Associates), 61–78.

Leemans, R., and Eickhout, B. (2004). Another reason for concern: regional and global impacts on ecosystems for different levels of climate change. *Glob. Environ. Change* 14, 219–228. doi: 10.1016/j.gloenvcha.2004.04.009

Morton, T. L. (1978). Intimacy and reciprocity of exchange: a comparison of spouses and strangers. *J. Pers. Soc. Psychol.* 36, 72. doi: 10.1037/0022-3514.36.1.72

Nishida, T. (2012). Augmenting conversational environment. *Int. J. Cogn. Inform. Nat. Intell.* 6, 103–124. doi: 10.4018/jcini.2012100105

Ogawa, K., Nishio, S., Koda, K., Taura, K., Minato, T., Ishii, C. T., and Ishiguro, H. (2011). "Telenoid: tele-presence android for communication," in *ACM SIGGRAPH 2011 Emerging Technologies*, SIGGRAPH '11 (New York, NY, USA. Association for Computing Machinery). doi: 10.1145/2048259.2048274

Oliveira, R., Arriaga, P., Santos, F. P., Mascarenhas, S., and Paiva, A. (2021). Towards prosocial design: a scoping review of the use of robots and virtual agents to trigger prosocial behaviour. *Comput. Hum. Behav.* 114, 106547. doi: 10.1016/j.chb.2020.106547

Pfattheicher, S., Nielsen, Y. A., and Thielmann, I. (2022). Prosocial behavior and altruism: a review of concepts and definitions. *Curr. Opin. Psychol.* 44, 124–129. doi: 10.1016/j.copsyc.2021.08.021

Sandini, G., and Sciutti, A. (2018). Humane robots—from robots with a humanoid body to robots with an anthropomorphic mind. *ACM Trans. Hum. Robot Interact.* 7, 1–4. doi: 10.1145/3208954

Schroeder, D. A., and Graziano, W. G. (2015). "The Field of Prosocial Behavior: An Introduction and Overview," in *The Oxford Handbook of Prosocial Behavior*. New York, NY: Oxford University Press. doi: 10.1093/oxfordhb/9780195399813.013.32

Siegel, M., Breazeal, C., and Norton, M. I. (2009). "Persuasive robotics: the influence of robot gender on human behavior," in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems* (St. Louis, MO: IEEE), 2563–2568. doi: 10.1109/IROS.2009.5354116

Singer, T., Seymour, B., O'Doherty, J. P., Stephan, K. E., Dolan, R. J., and Frith, C. D. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature* 439, 466–469. doi: 10.1038/nature04271

Stotland, E. (1969). Exploratory investigations of empathy. *Adv. Exp. Soc. Psychol.* 4, 271–314. doi: 10.1016/S0065-2601(08)60080-5

Strohkorb, S., and Scassellati, B. (2016). "Promoting collaboration with social robots," in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (Christchurch: IEEE), 639–640.

Terzioğlu, Y., Mutlu, B., and Şahin, E. (2020). "Designing social cues for collaborative robots: the role of gaze and breathing in human-robot collaboration," in *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '20 (New York, NY, USA: Association for Computing Machinery),343–357. doi: 10.1145/3319502.3374829

Thielmann, I., and Hilbig, B. E. (2014). Trust in me, trust in you: a social projection account of the link between personality, cooperativeness, and trustworthiness expectations. *J. Res. Pers.* 50, 61–65. doi: 10.1016/j.jrp.2014.03.006

Uccelli, M. M., Mohr, L. M., Battaglia, M., Zagami, P., and Mohr, D. (2004). Peer support groups in multiple sclerosis: current effectiveness and future directions. *Mult. Scler. J.* 10, 80–84. doi: 10.1191/1352458504ms973oa

Visser, M. S., and Roelofs, M. R. (2011). Heterogeneous preferences for altruism: gender and personality, social status, giving and taking. *Exp. Econ.* 14, 490–506. doi: 10.1007/s10683-011-9278-4

Weitzel, U., Urbig, D., Desai, S., Sanders, M., and Acs, Z. (2010). The good, the bad, and the talented: entrepreneurial talent and selfish behavior. *J. Econ. Behav. Organiz.* 76, 64–81. doi: 10.1016/j.jebo.2010.02.013

Whitmarsh, L., and O'Neill, S. (2010). Green identity, green living? the role of pro-environmental self-identity in determining consistency across diverse pro-environmental behaviours. *J. Environ. Psychol.* 30, 305–314. doi: 10.1016/j.jenvp.2010.01.003

Yoshimura, K., Ono, Y., Nakamura, K., Nathan, J. H., and Suzuki, K. (2001). Validation of the Japanese version of the neo five-factor inventory in a large community sample. *Psychol. Rep.* 88, 443–449. doi: 10.2466/pr0.2001.88.2.443

Zillman, D., and Cantor, J. R. (1977). Affective responses to the emotions of a protagonist. *J. Exp. Soc. Psychol.* 13, 155–165. doi: 10.1016/S0022-1031(77)80008-5