



<https://openaccess.leidenuniv.nl>

License: Article 25fa pilot End User Agreement

This publication is distributed under the terms of Article 25fa of the Dutch Copyright Act (Auteurswet) with explicit consent by the author. Dutch law entitles the maker of a short scientific work funded either wholly or partially by Dutch public funds to make that work publicly available for no consideration following a reasonable period of time after the work was first published, provided that clear reference is made to the source of the first publication of the work.

This publication is distributed under The Association of Universities in the Netherlands (VSNU) 'Article 25fa implementation' pilot project. In this pilot research outputs of researchers employed by Dutch Universities that comply with the legal requirements of Article 25fa of the Dutch Copyright Act are distributed online and free of cost or other barriers in institutional repositories. Research outputs are distributed six months after their first online publication in the original published version and with proper attribution to the source of the original publication.

You are permitted to download and use the publication for personal purposes. All rights remain with the author(s) and/or copyrights owner(s) of this work. Any use of the publication other than authorised under this licence or copyright law is prohibited.

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please contact the Library through email: OpenAccess@library.leidenuniv.nl

Article details

Kleijn R. de, Es L. van, Kachergis G. & Hommel B. (2019), Anthropomorphization of artificial agents leads to fair and strategic, but not altruistic behavior, INTERNATIONAL JOURNAL OF HUMAN-COMPUTER STUDIES 122: 168-173.

Doi: 10.1016/j.ijhcs.2018.09.008



Anthropomorphization of artificial agents leads to fair and strategic, but not altruistic behavior

Roy de Kleijn^{*,a}, Lisa van Es^a, George Kachergis^b, Bernhard Hommel^a

^a Cognitive Psychology Unit, Leiden University, The Netherlands

^b Department of Psychology, Stanford University

ARTICLE INFO

Keywords:

Strategic behavior
Fairness preference
Human-robot interaction
Anthropomorphism

ABSTRACT

With robots playing an increasing role in our daily lives, our emotional responses to them have become an active subject of study. The process of anthropomorphization, ascribing human affordances to non-human objects, is thought to play a large role in human-robot interaction. However, earlier studies have relied largely on experimenter's manipulation of anthropomorphism, and the use of virtual robots. The aim of this study was to investigate people's fairness preference and strategic and altruistic behavior toward different opponents (a human, a semi-humanoid and a spider-like robot, and a laptop) in two economic games. Anthropomorphization questionnaires and mood measures were also administered. Our findings suggest that fairness preference and strategic behavior are not predicted by the opponent's physical appearance, but instead predicted by individual differences in the tendency to anthropomorphize others. Altruistic behavior, on the other hand, is affected by the opponent's physical appearance.

1. Introduction

In a society where humans and robots increasingly interact with each other, it will become important for robots to be accepted by humans as members of a shared society. Humans have been shaped by both natural selection and cultural tradition to cooperate and interact with other humans and animals. In fact, some theories even state that human intelligence is the evolutionary result of complex social interactions, requiring the ability to predict the behavior of others (Dautenhahn, 1998; Dunbar, 1998). In contrast, robots are a relatively novel addition to our social environment and, as such, we are not shaped by evolution to accurately predict their behavior. Without such predictive ability, we believe it is important to investigate what would be critical design features to build up trust and collaboration between humans and robots. Increasing our insight into such features is likely to ease our transition to a more automated society.

The physical design of modern robots is rather heterogeneous. At the moment, consumer robots take the form of vacuum cleaners (e.g., Roomba) or self-driving vehicles, but as we expect robots to perform more and more everyday human action, we could expect robots to increasingly look like us. Many research robots that are used to study everyday action (e.g., Willow Garage's PR2) are equipped with two arms and a binocular camera system arranged similarly to the anatomy of humans (albeit with more wheels and fewer legs). This physical

similarity to humans has sparked the interest of human-robot interaction researchers, and some have suggested the existence of a non-linear relationship between physical similarity to human appearance and likeability of robots, nicknamed the *uncanny valley* (Mori, 2012; Pollick, 2010). The uncanny valley theory states that there is a positive relationship between the likeability of a robot and its human-likeness. However, at very high levels of human-likeness there is a sharp decrease in likeability, which is coined the uncanny valley.

1.1. Anthropomorphization

Another consequence of a robot's physical similarity to humans is an increase in potential *anthropomorphization*: the tendency to attribute human characteristics to non-human agents or even objects, such as animals or computers (Bartneck et al., 2009). It has been proposed (e.g. Epley et al., 2007) that the extent to which we anthropomorphize an agent is dependent on its physical similarity due to the inaccessibility of the phenomenological experience of others. While we are unable to imagine what it would be like to be, let's say, a bat (Nagel, 1974), it is easier for us to imagine what it would be like to be another person, and humanoid robots would potentially fall somewhere in between.

There has been an increasing interest in studying anthropomorphism, both on a psychological and a neuroscientific level. The perhaps most comprehensive psychological framework for

* Corresponding author.

E-mail address: kleijnrde@fsw.leidenuniv.nl (R. de Kleijn).

<https://doi.org/10.1016/j.ijhcs.2018.09.008>

Received 4 February 2017; Received in revised form 11 September 2018; Accepted 20 September 2018

Available online 21 September 2018

1071-5819/© 2018 Elsevier Ltd. All rights reserved.

anthropomorphism, described by Epley et al. (2007), predicts that people tend to anthropomorphize agents when they are motivated to be effective social agents, when they lack a connection to other humans, and when anthropocentric knowledge is accessible and applicable.

Other studies have instead argued for the importance of goal-directed, meaningful action. In as early as 1944, Heider and Simmel (1944) showed that people tend to attribute human states such as intent to elementary geometric shapes moving in a seemingly meaningful way. Similarly, neuroscientific studies on anthropomorphism have shown that the human mirror neuron system responds to observed actions performed even by industrial robots (Gazzola et al., 2007), especially when the action seems to be goal-directed. When anthropomorphizing, the superior temporal sulcus, which is also involved in dispositional attribution to people, and amygdala, which is involved in social categorization, seem to play an important role (Harris and Fiske, 2008). And indeed, amygdala-damaged patients, as well as patients with autism seem to exhibit impaired anthropomorphization (Castelli et al., 2002; Heberlein and Adolphs, 2004).

From a psychological viewpoint, it is interesting to investigate the social consequences of such anthropomorphization. Robots that look like humans, or for another reason are attributed with human-like characteristics, might be expected to elicit a higher empathic response than non-humanoid robots. However, as the uncanny valley may predict, this anthropomorphization could also lead to negative emotional responses, depending on the similarity to humans.

1.2. Altruistic and strategic behavior toward robots

Riek et al. (2009) have investigated the influence of anthropomorphization on empathic behavior. Subjects were presented with a film clip featuring one of five protagonists, ranging in physical appearance from a Roomba to a human. Film clips were either neutral or emotionally evocative in which the protagonist was being treated cruelly. After the film clip, subjects were asked which one of the four robots they would save in the event of an earthquake. More human-like protagonists induced higher empathy in subjects, feeling more sorry for them and reporting taking higher risks to save them.

Another paradigm in the study of empathy uses economic games to measure altruistic and strategic behavior. Underlying this design is the idea that altruistic behavior is necessarily preceded by empathic concern for others, known as the *empathy-altruism hypothesis* (Batson, 1991; Cialdini et al., 1997). In such economic games, some amount of money is given to a human participant, the proposer, who is subsequently asked to offer a stake of this amount to another player, the receiver. The dictator and ultimatum games are among the most widely-used economic games in the social sciences (Andersen et al., 2011; Engel, 2011; Güth et al., 1982). The premise of these games is that the amount of money given away by the proposer is an indicator of altruistic or strategic behavior, and reflects a preference for fairness.

In the ultimatum game, the proportion of the stake offered by the proposer is thought to reflect both an altruistic “taste for fairness” as well as the strategic anticipation that small offers may be turned down (Oosterbeek et al., 2004). Earlier research has shown that the amount proposed is dependent on the information given to both proposer and receiver, i.e. the proposer is more likely to make a fair offer if the proposer knows that the receiver is aware of the amounts to be divided (Pillutla and Murnighan, 1995). Also, the proposer is thought to reflect on the mental state of the responder (Campbell-Meiklejohn and Frith, 2012). In other words, the amount offered to the receiver is a function of the perceived capability of the receiver to know and reason with the proposed amount; it is hypothesized that “smart” receivers will be offered a larger stake due to (1) being perceived as able to reason with the proposed amount, and (2) the expectation that smart receivers will keep track of reciprocity, rejecting low offers to punish the proposer.

In the dictator game, the amount given away is considered a “more pure” measure of altruism (Eckel and Grossman, 1996; Fehr and Schmidt, 2006), as the receiver does not have the option of turning down an offer, removing the fear of rejection. Although the dictator game can be considered to measure a more pure form of altruism, factors such as experimental demand characteristics and social norms play a role as well (Bardsley, 2008).

Torta et al. (2013) investigated rejection rates in an ultimatum game in which human participants played as a receiver against a (virtual) proposer that was either human, a humanoid robot, or a computer. In their study, participants rejected offers made by a computer more often than offers made by a human or humanoid robot, although this effect was only marginally significant. However, these findings seem to contradict earlier studies, which have more consistently shown that when offers are made by a computer rather than a human player rejection rates are much lower (Moretti and di Pellegrino, 2010; Sanfey et al., 2003). Sanfey et al. (2003) showed that this is reflected in neural activity, and found weaker activation of the anterior insula when unfair offers were randomly generated by a computer instead of a human opponent in an fMRI study investigating ultimatum game rejection behavior.

In a similar paradigm, van Dijk (2013) also had participants play as the receiver in an ultimatum game. In addition, participants completed anthropomorphism questionnaires in which they rated how much they anthropomorphized their (virtual) opponent, which could either be a human, a robot, or a computer. While this study did not find an effect of opponent type on rejection behavior, a correlation was found between anthropomorphization and rejection behavior, where offers being made by proposers who were anthropomorphized more were less likely to be rejected.

This finding suggests that it is not the opponent *type*, which is often manipulated by experimenters to manipulate different levels of anthropomorphism, but individual *differences in anthropomorphization* that determine rejection behavior.

1.3. The current study

So far, several human-robot interaction studies looking at ultimatum game behavior have focused on rejection rate behavior (Moretti and di Pellegrino, 2010; Sanfey et al., 2003; Torta et al., 2013). However, relatively few studies have investigated altruism using proposer behavior and individual differences in the tendency to anthropomorphize (van Dijk, 2013).

In the current study, we investigated the role of both physical human-robot similarity and the individual degree of anthropomorphization on altruistic and strategic behavior. To assess altruistic and strategic behavior, we used the dictator and ultimatum games. In this study, human participants were proposers, and we used different types of robots as well as a human confederate as receivers. The manipulation of the type of opponent was thought to tap into the physical similarity between proposer and receiver, with the other human being the most similar and the laptop being the least similar opponent.

Due to the criticism toward the use of virtual robots or avatars (Bainbridge et al., 2011; Li, 2015), we used physical, co-present robots. Importantly, we also assessed the individual degree of anthropomorphization. Using this design, we could more carefully investigate the effect of anthropomorphization on altruistic and strategic behavior. We were particularly interested in comparing the impact of physical similarity and individual anthropomorphization on altruistic and strategic behavior. From the viewpoint of physical similarity, the latter should merely reflect the former, which should be the main factor accounting for the degree of altruistic behavior—which in turn should be most pronounced for the human opponent and least pronounced for the laptop. From an anthropomorphization point of view, however, it might be mainly the individual tendency to perceive an opponent as human-like that determines the proposer’s altruistic and strategic behavior.

Based on the theoretical considerations above, we hypothesized that (1) participants will offer a larger proportion of their stake in the dictator and ultimatum games to opponents they anthropomorphize more, and (2) individual differences in anthropomorphization, rather than physical appearance, predict altruistic behavior.

2. Method

2.1. Participants

We recruited 136 participants (16 males, 120 females; age $M = 22.4$ years old, $SD = 3.22$) from the Leiden University online participant database. Almost all participants were Western European undergraduate psychology students at Leiden University and—as per institutional requirements—all were paid 3.50 euro or 1 participation credit. In addition, they were paid 18.75% of profits made during the experiment. The study was approved by the institutional review board of the Institute for Psychology at Leiden University.

2.2. Materials

2.2.1. Opponent types

We manipulated opponent type using one of four game opponents. The first one was a regular Dell laptop computer. The second opponent was a spider-like Hermes II hexapod robot (IS Robotics, Somerville, MA). The third one was a semi-humanoid Q.bo robot (TheCorpora Inc, Madrid, Spain) with a head with two degrees of freedom and two large eyes. The fourth was a female human opponent, who was a confederate of the experimenters. All three artificial opponents can be seen in Fig. 1.

2.2.2. Economic and social games

Dictator game. In a dictator game (Kahneman et al., 1986), the participant (the *proposer*) is given a certain amount of money or points by the experimenter, and is asked to divide this amount between themselves and their opponent by giving away a proportion of it. Participants are told that this is a one-shot game, so there will be no further interaction with the opponent after the game. From a *homo economicus* perspective, a person should not give anything to the receiver; after all, any amount given away would be a loss to the proposer. However, meta-analyses suggest that most people in fact give away at least part of their stake, across diverse demographic groups (Engel, 2011). As such, the dictator game is considered to be a suitable instrument for measuring altruistic, unselfish behavior (Eckel and Grossman, 1996; Fehr and Schmidt, 2006), with good external validity (Franzen and Pointner, 2013). However, it should be noted that some authors have argued that dictator game behavior is not a measure of *pure* altruism,

but reflects experimental demand characteristics and social norms (Bardsley, 2008).

Ultimatum game. The ultimatum game (Güth et al., 1982) has a structure similar to the dictator game, but here the receiver is given the opportunity to either accept or decline the offer. If the receiver accepts, the amount is divided according to the offer made by the proposer. If the receiver declines, neither player receives anything. Assuming economically-rational agents, the receiver will accept any non-zero offer, while the proposer will offer the lowest possible non-zero amount. However, earlier research has shown that offers seen as “unfair” are often rejected. On average, proposers offer as much as 40% of the stake to the receiver (Oosterbeek et al., 2004). It is thought that proposer behavior in the ultimatum game is a measure of both an altruistic preference for fairness and strategic behavior, as a too low offer could lead to rejection. In the current study, the participant will act as a proposer against one of the four opponents.

2.2.3. Affect grid

The affect grid (Russell et al., 1989) is a single-item scale on which participants rate their current affect on the dimensions pleasure-displeasure and arousal-sleepiness. It is presented as a 9×9 grid on which the participant selects the position that best reflects his or her current affect, and has been shown to have moderate to good reliability, convergent validity, and discriminant validity (Killgore, 1998; Russell et al., 1989).

2.2.4. Anthropomorphism questionnaires

Anthropomorphization is a difficult concept to define, as it encompasses many (indeed perhaps all) aspects of human behavior. We chose to use two operationalizations, and we will refer to these two measures as *cognitive anthropomorphism* and *general anthropomorphism*.

To assess cognitive anthropomorphism, we used the Epley questionnaire (Epley et al., 2007), which was inspired by three psychological determinants. It is a five-item questionnaire using a 7-point Likert scale for all items, and is used to assess anthropomorphism of robots based on the accessibility and applicability of anthropocentric knowledge, the motivation to explain and understand the behavior of other agents, and the desire for social contact and affiliation. The questions asked specifically concerned the assigned opponent, e.g. “This opponent can feel emotions”. It has been shown to have excellent reliability, $\alpha = 0.938$ (Torta et al., 2013).

In addition to the Epley questionnaire, we also administered the Van ’t Sant questionnaire (van Dijk, 2013) to assess general anthropomorphism. This is a 25-item dichotomous questionnaire to assess not only cognitive aspects of anthropomorphization, but also more general human affordances, such as language ability, physical capability, and emotional experience, e.g. “This opponent can perceive objects”, “This opponent can understand language”, and “This opponent can talk”. Reliability data for this relatively new measure of anthropomorphism is being collected.

2.3. Design and procedure

After having signed the informed consent form, participants were randomly assigned to one of four types of opponent: (1) a human opponent, (2) a semi-humanoid robot, (3) a hexapod robot, or (4) a laptop computer (see Fig. 1).

Participants performed the computer tasks and questionnaires on a desktop computer placed opposite the opponent. They were told that they would be playing against the opponent, while in fact, all interactions with the opponent were mock interactions, and preprogrammed by the experimenter. None of the artificial opponents made any movements or verbalizations. The human opponent (confederate) was instructed to minimize unnecessary conversation and movement. Therefore, all “interactions” with opponents were completely moderated by the participant’s computer.

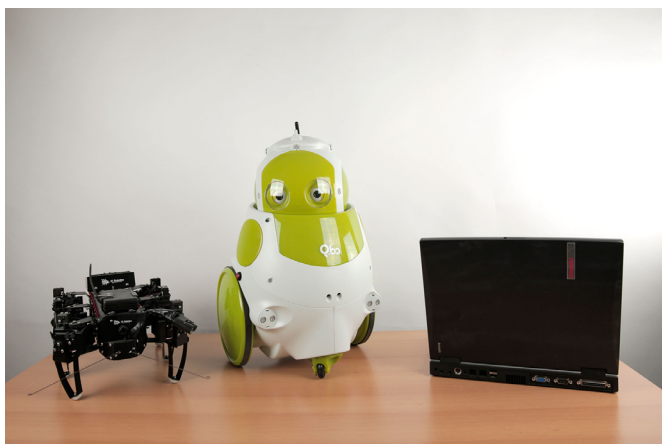


Fig. 1. Robot types used in this study. From left to right: the hexapod robot, the humanoid robot, and the laptop computer.

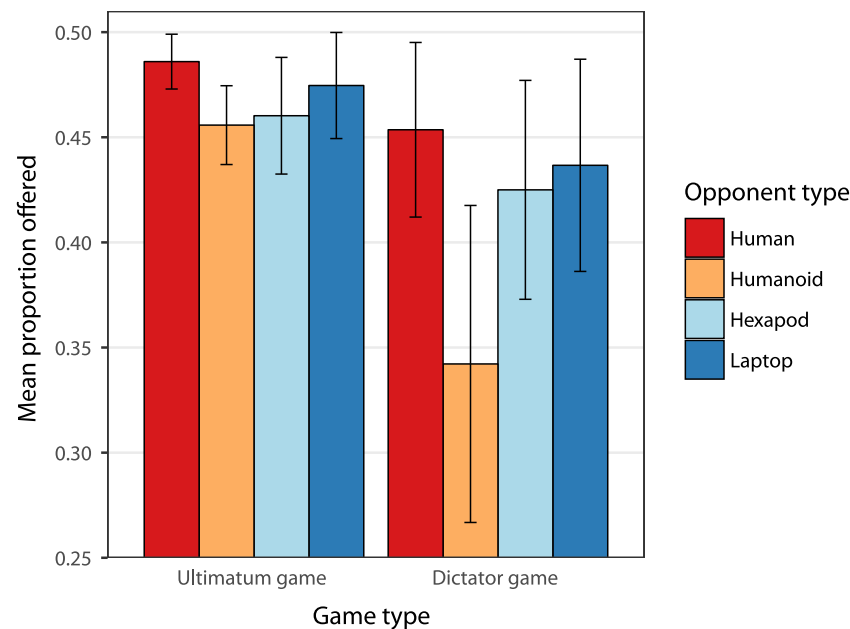


Fig. 2. The relation between opponent type and proportion of stake offered to the opponent for both the ultimatum game and the dictator game. Error bars indicate 95% CI.

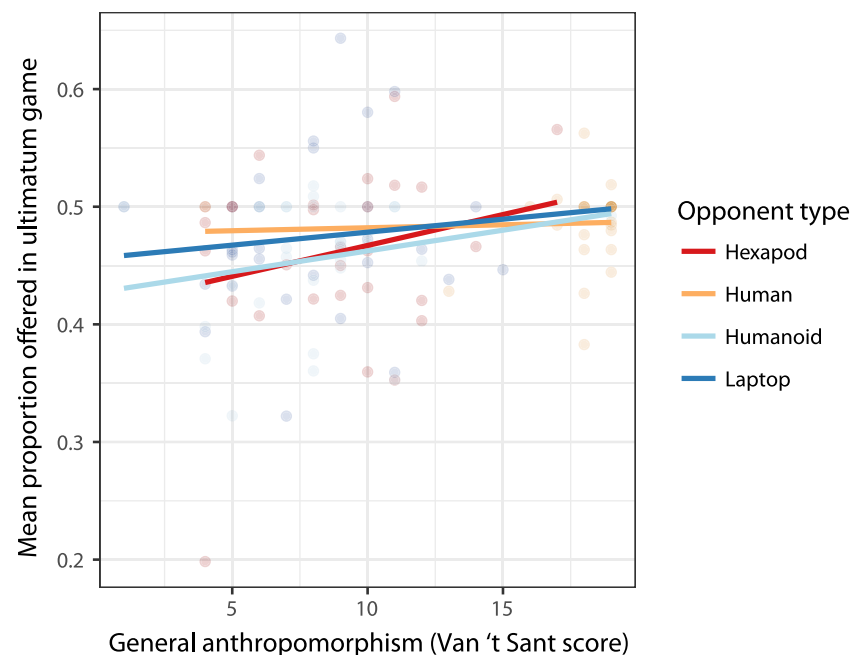


Fig. 3. The relation between general anthropomorphism as measured by the Van 't Sant questionnaire and proportion of stake offered to the opponent in the ultimatum game.

Before starting, all participants were informed that they would be playing for real money, and that 18.75% of their profits would be paid out at the end of the experiment. The experiment started with the affect grid to determine a baseline measurement for mood.

Next, participants played the dictator game as the proposer. Participants were given 10 euro, and were instructed that they could give away a proportion of this money to their opponent. They were also informed that the opponent had no say in this, and that the money would be divided as proposed by the participant.

After this, participants played 8 rounds of the ultimatum game, in the role of proposer, against the opponent. Each round, a different amount was given to the participant, in random order. Amounts given were 1, 2, 4, 8, 10, 20, 30, and 50 euro. The opponent was programmed

(or, in the case of the confederate, instructed) to decline offers lower than 30% of the stake, similar to human behavior (Fehr and Fischbacher, 2003).

Finally, the Epley and Van 't Sant anthropomorphization questionnaires were administered to assess cognitive and general anthropomorphization, respectively, after which the affect grid was presented for the last time. The total duration of the experiment was approximately 30 min.

3. Results

Due to the skewed gender distribution of the sample, our analyses were performed over female participants only. No significant

differences were found between males and females on measures of anthropomorphism or ultimatum or dictator game performance, $t_s < 0.707$, $p_s > 0.481$.

3.1. Anthropomorphization of robots

Both measures of anthropomorphization correlated strongly, $r(118) = 0.830$, $p < .001$. An analysis of variance with Bonferroni-corrected post-hoc tests revealed no significant differences between robot types on either measure of anthropomorphization, $p_s > 0.562$. Unsurprisingly, the human confederate was anthropomorphized more than the robot opponents on both measures, $p < .001$.

3.2. Dictator and ultimatum game

A one-way analysis of variance revealed a significant effect of opponent type on dictator game behavior, $F(3, 116) = 3.191$, $p = .026$, but not on ultimatum game behavior, $F(3, 116) = 1.607$, $p = .192$. Figure 2 shows the overall results. This suggests that in the ultimatum game, participants offered an equally large proportion of the stake to all types of opponents. In the dictator game, Bonferroni-corrected post-hoc tests revealed that the human opponent was offered a larger proportion of the stake ($M = 0.454$) than the humanoid opponent ($M = 0.342$), $p = .037$.

Using linear regression, we did not find a relationship between anthropomorphization and dictator game behavior, $R^2_{adj} = 0.007$, $F(1, 118) = 1.798$, $p = .183$. However, we did find a small but significant linear relationship between general anthropomorphization and ultimatum game behavior, $R^2_{adj} = 0.043$, $F(1, 118) = 6.334$, $p = .013$ (see Fig. 3). Adding polynomial terms did not significantly improve the model. Regression slopes did not differ between opponent types, $F(3, 112) = 0.320$, $p = .811$, suggesting that this relationship is equally large for all opponent types.

To investigate which items in the general anthropomorphism questionnaire have the most influence on ultimatum game behavior, we trained a random forest model with 50 trees on the individual items and ultimatum game behavior. The items “This opponent is ambitious”, “This opponent has a goal”, and “This opponent understands language” have the strongest influence on the random forest output. These results show that—although the type of opponent does not affect ultimatum game behavior—the amount to which participants anthropomorphize the opponent does.

Summarizing, the physical appearance of the opponent did not influence ultimatum game behavior, but did influence dictator game behavior. In contrast, anthropomorphization of the opponent did not influence dictator game behavior, but did influence ultimatum game behavior.

4. Discussion

In this study, we investigated strategic and altruistic behavior as a function of anthropomorphization. The current design was motivated by two concerns with earlier studies: (1) the manipulation of anthropomorphism by experimenters instead of participants, and (2) the widespread use of virtual robots. In this approach, we not only analyzed the effect of robot type on strategic and altruistic behavior, but also the effect of the participants’ anthropomorphization of the opponent. In addition, opponents were not virtual, but physically present in the same room as the participants. The large amount of variance in anthropomorphization for all robot types is clearly visible in Fig. 3.

We did not find an effect of opponent type on ultimatum game behavior. Ultimatum game behavior measures motivation for fairness, but is also influenced by the risk of rejection. More humanlike opponents are both more likely to enforce fairness norms and should therefore be offered a larger proportion than non-humanlike opponents, but are also likely to stimulate an empathic willingness to be fair. This is

visible in our data as a relationship between opponent anthropomorphization and the proportion of the stake offered to the opponent in the ultimatum game. Also, the items in the anthropomorphism questionnaire related to agent autonomy and goal-directedness seem to have the most influence in this relationship. In this sense, anthropomorphization of an opponent seems to affect the intrinsic motivation for fairness as well as strategic motivations.

The results from the dictator game are somewhat more difficult to interpret. The dictator game is thought to measure a somewhat more pure form of altruistic behavior (although interpretations differ, see our discussion in Section 1.2). We found that the semi-humanoid robot was offered the smallest amount of money. We posit that this may reflect the discrepancy between expected and actual opponent behavior. None of the non-human opponents physically interacted with the participant, which is perhaps not surprising in the case of the laptop (which do not normally move) and the hexapod (due to unfamiliarity). However, the semi-humanoid is a robot that is normally displayed in the media and even around the university as a moving, autonomous agent, possibly causing an “uncanny” feeling due to this discrepancy. However, this hypothesis cannot be tested in the context of the current study but requires follow-up research.

In conclusion, our findings highlight that anthropomorphism is not merely the result of an object’s physical appearance, but rather reflects individual differences in the tendency to anthropomorphize, as has been proposed by Waytz et al. (2014). Our results show that, in order to make social robots, it does not suffice to study robot design from an industrial design viewpoint.

Limitations of the current study are the rather noisy nature of the one-shot dictator game, as well as the limited interaction with the opponents. It is quite possible that an iterated dictator game using various stakes would provide a more stable measure of altruistic behavior. Also, it would be interesting to see how prolonged, and more physical interaction with robots would affect anthropomorphization as well as altruistic behavior. While the complete moderation of participant–opponent interaction by a computer reduces unwanted differences between the different opponent conditions, it could also have reduced believability of true interaction with the opponents.

It is quite possible that the process of anthropomorphization is dependent on task-relevant feature overlap between human and robot, or overlap of task-relevant affordances. Future research could focus on the individual determinants of anthropomorphization and how they can be influenced.

Funding

The preparation of this work was supported by the European Commission (EU Cognitive Systems project ROBOHOW.COG; FP7-ICT-2011; grant no. 288533).

Acknowledgment

We gratefully thank Vera Mekern for her assistance with data collection, Rosanne van den Berg for photography, and Roderik Gerritsen for help with interpreting our data.

References

- Andersen, S., Ertaç, S., Gneezy, U., Hoffman, M., List, J.A., 2011. Stakes matter in ultimatum games. *Am. Econ. Rev.* 101, 3427–3439.
- Bainbridge, W.A., Hart, J.W., Kim, E.S., Scassellati, B., 2011. The benefits of interactions with physically present robots over video-displayed agents. *Int. J. Soc. Robot.* 3, 41–52.
- Bardsley, N., 2008. Dictator game giving: altruism or artefact? *Exp. Econ.* 11, 122–133.
- Bartneck, C., Croft, E., Kulic, D., 2009. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *Int. J. Soc. Robot.* 1 (1), 71–81.
- Batson, C.D., 1991. *The Altruism Question: Toward a Social-Psychological Answer*. Hillsdale, NJ: Erlbaum.

- Campbell-Meiklejohn, D., Frith, C.D., 2012. Social factors and preference change. In: Dolan, R., Sharot, T. (Eds.), *Neuroscience of Preference and Choice*. Academic Press, San Diego, pp. 177–206.
- Castelli, F., Frith, C., Happé, F., Frith, U., 2002. Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain* 125, 1839–1849.
- Cialdini, R.B., Brown, S.L., Lewis, B.P., Luce, C., Neuberg, S.L., 1997. Reinterpreting the empathy–altruism relationship: when one into one equals oneness. *J. Personal. Soc. Psychol.* 73, 481–494.
- Dautenhahn, K., 1998. The art of designing socially intelligent agents: science, fiction, and the human in the loop. *Appl. Artif. Intell.* 12, 573–617.
- van Dijk, E., 2013. Investigating Rejection Behavior in the Ultimatum Game as a Measure of Anthropomorphism. Master's thesis.
- Dunbar, R.I.M., 1998. The social brain hypothesis. *Evol. Anthr. Issues News Rev.* 6, 178–190.
- Eckel, C.C., Grossman, P.J., 1996. Altruism in anonymous dictator games. *Games Econ. Behav.* 16, 181–191.
- Engel, C., 2011. Dictator games: a meta-study. *Exp. Econ.* 14, 583–610.
- Epley, N., Waytz, A., Cacioppo, J.T., 2007. On seeing human: a three-factor theory of anthropomorphism. *Psychol. Rev.* 114, 864–886.
- Fehr, E., Fischbacher, U., 2003. The nature of human altruism. *Nature* 425, 785–791.
- Fehr, E., Schmidt, K.M., 2006. The economics of fairness, reciprocity and altruism – experimental evidence and new theories. In: Kolm, S.-C. (Ed.), *Handbook of the Economics of Giving, Altruism and Reciprocity*. Elsevier, Amsterdam, pp. 615–691.
- Franzen, A., Pointner, S., 2013. The external validity of giving in the dictator game: a field experiment using the misdirected letter technique. *Exp. Econ.* 16, 2.
- Gazzola, V., Rizzolatti, G., Wicker, B., Keysers, C., 2007. The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. *NeuroImage* 35, 1674–1684.
- Güth, W., Schmittberger, R., Schwarze, B., 1982. An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Organ.* 3, 367–388.
- Harris, L.T., Fiske, S.T., 2008. The brooms in Fantasia: neural correlates of anthropomorphizing objects. *Soc. Cogn.* 26, 210–223.
- Heberlein, A.S., Adolphs, R., 2004. Impaired spontaneous anthropomorphizing despite intact perception and social knowledge. *Proc. Natl. Acad. Sci. U.S.A* 101, 7487–7491.
- Heider, F., Simmel, M., 1944. An experimental study of apparent behavior. *Am. J. Psychol.* 57, 243–259.
- Kahneman, D., Knetsch, J.L., Thaler, R., 1986. Fairness and the assumptions of economics. *J. Bus.* 59, S285–S300.
- Killgore, W.D., 1998. The affect grid: a moderately valid, nonspecific measure of pleasure and arousal. *Psychol. Rep.* 83, 639–642.
- Li, J., 2015. The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *Int. J. Hum. Comput. Stud.* 77, 23–37.
- Moretti, L., di Pellegrino, G., 2010. Disgust selectively modulates reciprocal fairness in economic interactions. *Emotion* 10, 169–180.
- Mori, M., 2012. The uncanny valley. *IEEE Robot. Autom. Mag.* 98–100.
- Nagel, T., 1974. What is it like to be a bat? *Philos. Rev.* 83, 435–450.
- Oosterbeek, H., Sloof, R., van de Kuilen, G., 2004. Cultural differences in ultimatum game experiments: evidence from a meta-analysis. *Exp. Econ.* 7, 171–188.
- Pillutla, M.M., Murnighan, J.K., 1995. Being fair or appearing fair: strategic behavior in ultimatum bargaining. *Acad. Manag. J.* 38, 1408–1426.
- Pollick, F.E., 2010. In search of the uncanny valley. In: Daras, P., Ibarra, O.M. (Eds.), *User Centric Media*. Springer, pp. 69–78.
- Riek, L.D., Rabinowitch, T., Chakrabarti, B., Robinson, P., 2009. How anthropomorphism affects empathy toward robots. *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction*. New York, NY. pp. 245–246.
- Russell, J.A., Weiss, A., Mendelsohn, G.A., 1989. Affect grid: a single-item scale of pleasure and arousal. *J. Personal. Soc. Psychol.* 57, 493–502.
- Sanfey, A.G., Rilling, J.K., Aronson, J.A., Nystrom, L.E., Cohen, J.D., 2003. The neural basis of economic decision-making in the ultimatum game. *Science* 300, 1755–1758.
- Torta, E., van Dijk, E., Ruijten, P.A.M., Cuijpers, R.H., 2013. The ultimatum game as a measurement tool for anthropomorphism in human-robot interaction. *Lect. Notes Comput. Sci.* 8239, 209–217.
- Waytz, A., Cacioppo, J., Epley, N., 2014. Who sees human? The stability and importance of individual differences in anthropomorphism. *Perspect. Psychol. Sci.* 5, 219–232.