Should Robots Chicken? How Anthropomorphism and Perceived Autonomy Influence Trajectories in a Game-Theoretic Problem

Ilaria Torre
ilariat@kth.se
KTH Royal Institute of Technology
Stockholm, Sweden

Alexis Linard
linard@kth.se
KTH Royal Institute of Technology
Stockholm, Sweden

Anders Steen
astee@kth.se
KTH Royal Institute of Technology
Stockholm, Sweden

Jana Tumová
tumova@kth.se
KTH Royal Institute of Technology
Stockholm, Sweden

Iolanda Leite
iolanda@kth.se
KTH Royal Institute of Technology
Stockholm, Sweden

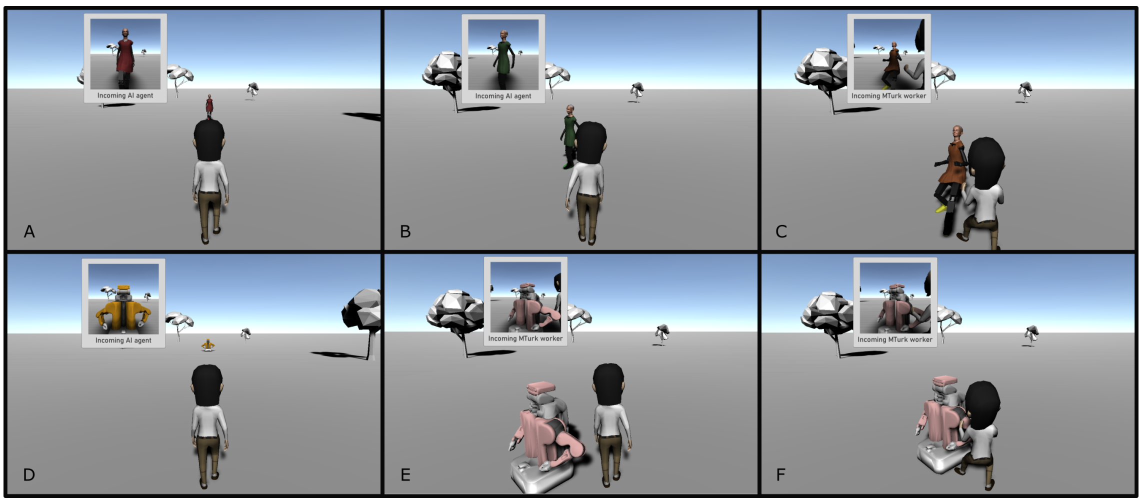

Figure 1: Screenshots from the chicken game experiment. A: Robot Sophia and participant’s avatar walking towards each other. B: Sophia swerving. C: Sophia and participant’s avatar colliding. D: Robot PR2 and participant’s avatar walking towards each other. E: PR2 swerving. F: PR2 and participant’s avatar colliding. The zoomed up window in the top left corners indicates whether the robot is ‘teleoperated’ (C, E, F, labelled as ’Mturk worker’) or ‘autonomous’ (A, B, D, labelled as ’AI agent’).

ABSTRACT

Two people walking towards each other in a colliding course is an everyday problem of human-human interaction. In spite of the different environmental and individual factors that might jeopardise successful human trajectories, people are generally skilled at avoiding crashing into each other. However, it is not clear if the same strategies will apply when a human is in a colliding course with a robot, nor which (if any) robot-related factors will influence the human’s decision to swerve or not. In this work, we present the results of an online study where participants walked towards a virtual robot that differed in terms of anthropomorphism and perceived autonomy, and had to decide whether to swerve, or continue straight. The experiment was inspired by the game-theoretic game of chicken. We found that people performed more swerving actions when they believed the robot to be teleoperated by another participant. When they swerved, they also swerved closer to the robot with high levels of human-likeness, and farther away from the robot with low anthropomorphism score, suggesting a higher uncertainty about the mechanical-looking robot’s intentions. These results are discussed in the context of socially-aware robot navigation, and will be used to design novel algorithms for robot trajectories that take robot-related differences into account.

HRI ’21, March 8–11, 2021, Boulder, CO, USA

ACM ISBN 978-1-4503-8289-2/21/03…$15.00
https://doi.org/10.1145/3434073.3444687

KEYWORDS

Human-aware navigation, anthropomorphism, autonomy, game theory

ACM Reference Format:

Ilaria Torre, Alexis Linard, Anders Steen, Jana Tumová, and Iolanda Leite. 2021. Should Robots Chicken? How Anthropomorphism and Perceived Autonomy Influence Trajectories in a Game-Theoretic Problem. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction (HRI ’21), March 8–11, 2021, Boulder, CO, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3434073.3444687

1 INTRODUCTION

Consider the following situation, which is often encountered in everyday human-human interactions: as you are walking, you notice that there is another person on the same path as you, coming from the opposite direction. Do you continue straight, or do you move to the side and let the other person continue straight? And if you move, how far from the other person do you perform this swerving action? In this paper, we look at how people behave when the incoming agent is a robot.

From the human-human interaction example just mentioned, we can imagine that a person’s behaviour will be dictated by many different individual and environmental factors. For example, people who are in a hurry might be less likely to let other people pass. Characteristics of the incoming individual might also play a role, for example if they look distracted, or angry. Certain cultures might also favour giving way to elderly people; and so on. In other words, if there are people involved, the seemingly simple action of going from point A to point B will be regulated by a series of unwritten specifications that influence people’s trajectories. It is not just a coordination problem, but a social coordination problem [37]. Yet, so far this problem has mostly been approached from a purely computational perspective, for example, by modelling individual trajectories in a crowd with fluid dynamics [e.g. 31], or as repulsively interacting physical particles [e.g. 88].

However, for social robots to step out into society, this algorithmically optimal approach – consisting of quantifiably successful actions, such as reaching the target destination in time, avoiding collisions, etc. – will have to be integrated with a ‘social’ approach – consisting of the robot’s ability to behave based on subjective human preferences and concerns. Indeed, socially-aware robot navigation has already received some attention in the HRI community [e.g. 43, 70]. Most previous studies focused on defining robot navigation in social spaces based on environmental features – e.g. corridors vs. open spaces [65] – and features of nearby humans – e.g. groups vs. individuals [81]. However, what we are still missing is developing models for robot navigation based on robot-related features, such as the robot’s physical appearance and attributed autonomy. A similar gap is present also in the human-human navigation literature, where very few studies investigated how first impressions of an incoming person influence another person’s navigation and collision avoidance behaviour [39].

While previous studies assumed that the same models would be suitable for different robots and different contexts, with the current work we contribute to the topic of socially-aware navigation by showing that people’s behaviour changes based on the anthropomorphic level of the robot, and on whether the robot is believed to be teleoperated or autonomous.

2 RELATED WORK

A useful concept when dealing with human-robot negotiation of space in shared environments is proxemics – the personal space that people maintain around themselves [24]. Proxemics has been widely studied in HRI, mostly in the context of finding the right distance from which people feel comfortable when interacting with a robot [e.g. 59]. Similarly to human-human interactions, people tend to maintain their personal space when interacting with robots [59, 72, 78], including virtual ones [2, 3, 49, 66], although these spaces tend to be smaller than human-human ones [84]. Also, the actual distance might vary depending on the type of task being performed [34]. This comfort distance also tends to decrease for repeated interactions with the same robot [26, 40, 84]. Measuring people’s preferred distances allows researchers to program robots to keep the same distance when approaching humans, thus taking into account people’s comfort, as well as safety [13, 57, 73, 75, 82].

Most of these studies on human-robot distances focused on modelling robot behaviour based on different human and environmental characteristics, such as creating different approach behaviours based on people’s age [e.g. 83] or on how crowded a space is [e.g. 10]. However, it is reasonable to assume that robot-related characteristics will also influence a socially appropriate human-robot distance. Indeed, studies of human-human navigation and collision avoidance suggest that people anticipate the future motion of dynamic objects and possible collisions; this means that humans include prediction into their own motion planning and do not solely react [39, 64]. This predictive behaviour is likely going to be used when anticipating robots’ trajectories, and this is where the perception of robots’ social capabilities will come into play. Specifically, people’s behaviour might be modulated by whether they think the robot ‘understands’ concepts such as personal space and comfort zone.

A characteristic that is likely to influence people’s perceptions of a robot’s social capabilities is anthropomorphism, i.e. the degree to which a robot resembles a human. Previous studies have shown that people attribute more social capabilities to human-like robots, compared to machine-like robots, which in turn could influence people’s expectations of how the robot will move in a shared space. For example, Krach et al. [41] had participants play a Prisoner’s Dilemma game with a humanoid robot, a robot made of Lego blocks, a computer and a human; they found that participants had increasingly more fun and thought that the agents were increasingly more intelligent the more the agent was human-like. The authors speculate that participants attributed increased agency to increased human-likeness. In Walters et al. [84], participants kept more distance from a mechanical-looking robot with a synthetic voice than a mechanical-looking robot with a pre-recorded human voice, suggesting that machine-like characteristics evoked a feeling of ‘threat’. In another study, adding a human-like body to a mobile box-shaped robot caused participants’ comfort level to drop, and, as a result, they kept a bigger distance from it [9]. However, robotic bodies have greatly improved since this study was conducted, and these results might not be applicable to the increasing acceptability and familiarity with robots in the wider population. A few studies also suggest that different levels of anthropomorphism elicit different expectations regarding rational behaviour. For example, Malle et al. [53] highlighted that there is an asymmetry in moral judgments: humans are blamed more for action in a moral dilemma (derailing the train so that it kills the worker but saves the 4 passengers), while robots are blamed more for inaction (not derailing the train and letting the 4 passengers crash). In a follow-up study, Malle et al. [54] added a pictorial representation of a robot to the same experiment: the story was depicted either with a human making a decision, a human-like robot, a mechanical robot, or a generic AI. In 2 out of 3 experiments, they found that the asymmetry only held for a mechanical robot, while the human-like robot and the AI were judged similarly to the human. They suggest that people have a mental model of mechanical robots as cold and calculating, while human-like robots were ascribed more human-like features and judgments, consistent with making a moral choice.

This leads to the formulation of our first research question:

*RQ*1: Does robot anthropomorphism influence human trajectories, specifically in whether people will swerve to let the robot pass, and how much distance they will keep from it?

From previous literature, we can speculate that, given expectations of rational behaviour from more mechanical-looking robots, as well as the potential unease at seeing one walking towards you, people might keep a bigger distance from them; conversely, given expectations of social capabilities from human-like robots, people might keep a smaller distance from them, expecting them to give way first.

Another aspect, related to anthropomorphism, that can contribute to how people behave with robots, is perceived autonomy. How much a robot is perceived to act based on its own artificial intelligence, or based on explicit human guidance [5], contributes to a robot's perceived human-likeness in the sense that it implies capabilities for "self-governed movement, understanding, and decision-making" [74, p. 207]. While humans are granted autonomy and agency by default, artificial agents' (in)dependence from humans is taken into account and influences social decision-making [15, 16, 46]. Due to this difference in agency, (perceived) teleoperated avatars and autonomous agents often yield different levels of social influence. Note that we take the following definitions of the terms 'avatar' and 'agent': we use 'agent' to define artificial agents that are not controlled by the human whose behaviour is being assessed – although they might be controlled by an experimenter – while 'avatar' refers to an artificial character that serves as a proxy for a human user [1].

For example, Hoyt et al. [33] conducted a study wherein participants were asked to perform a learned or novel task in front of virtual humans, and found that participants' performance of the novel task was impaired when they were led to believe that the virtual audience consisted of avatars, but not when they were framed as agents. In a meta-analysis, Fox et al. [23] also concluded that avatars exert stronger levels of social influence compared to agents. Similarly, people cooperated more with a (cooperative) virtual confederate if they thought it was controlled by a human rather than a computer [16, 18]. In general, people are likely to allocate more goods in social dilemma games to agents that are supposed to have a human-like mind, based on perceived agency and patiency [17]. The authors suggest that there are differences in how people behave with a virtual agent, based on whether they think the virtual agent is being directed by a human, or has a "mind of its own". Also, people felt angrier when they were faced with competitive agents than competitive avatars. De Melo et al. [16] suggest that this difference in behaviour could either be due to people adhering more to social rules with avatars (such as keeping harmony and professionalism), or because they are genuinely feeling angrier at agents, consistent with theories of out-group prejudice [77]. Thus, an agent's mind and a human player's perception of an agent's mind are crucial to how their social decision-making and negotiation unfold [17].

For robots specifically, there are few studies on the effect of perceived autonomy on human behaviour, although some conclusions can be inferred from studies on human perception. For example, autonomous robots were rated as more intelligent than teleoperated ones, but people felt more social presence, empathy, and companionship towards teleoperated robots [12, 20, 44, 45]. Also, people rated feeling more secure when working with teleoperated robots [86], and more optimistic about potentially getting rescued by a robot that was controlled by a human operator [20]. Similarly, robots were perceived as posing a bigger threat if they were presented as autonomous, rather than only capable of following human commands [89]. Also, people were more willing to follow advice from an autonomous robot to conduct objective tasks, but not subjective tasks [45]. These findings, together with the previously-discussed asymmetry in moral judgments [53, 54], suggests that people expect certain types of robots to behave in a 'rational' way.

These findings are consistent with the idea that people behave more socially with entities that are perceived as having more agency [7, 8], as well as identify more with entities that share a social group with them, such as other humans, as opposed to machines [77]. Still, with the advent of autonomous machines, such as self-driving cars, it is important to investigate how perception of this autonomy will influence human behaviour. The concept of autonomy is important also regarding attribution of blame and social responsibility: as robots become more and more autonomous, it will be necessary, from a moral standpoint, to come to an agreement as to whose fault it will be when things go wrong [30].

From this, we can formulate our second research question:

*RQ*2: Does perceived robot autonomy influence human trajectories, specifically in whether people will swerve to let the robot pass, and how much distance they will keep from it?

Framing this question in the Social Psychology literature, we can expect people to behave more 'politely', such as giving way first, with artificial agents that are perceived to be controlled by other humans, consistent with Social Identity Theory [77].

In this section, we have presented a series of previous works demonstrating that people behave differently when an artificial agent exhibits different levels of anthropomorphism and autonomy. However, none of these studies dealt specifically with how people adjust their walking trajectories to moving robots on this basis. Our paper addresses this gap by presenting a novel study based on Game Theory, whereby participants were asked to decide if, and when, to swerve when a robot with different levels of anthropomorphism (human-like vs. machine-like) and perceived autonomy (autonomous vs. controlled by a human) walked towards them in a virtual environment.

3 METHOD

In this section, we briefly describe the game of chicken, which was the basis of our experiment, in its original, game-theoretic formulation, followed by how we implemented it in an experiment. This experiment was conducted online due to COVID-19 restrictions at the time. However, this allowed us to have a greater degree of control on the displayed robots – e.g. changing robot colour at each trial, thus reinforcing participants' beliefs that they were interacting with a new robot every time – and on participants' behaviour – e.g. reducing participants' action space to only swerving (or not), thus removing many potential sources of variability, such as approach speed, swerving angle, etc. These will be addressed in follow-up studies that we are planning to run (see Section 5.2). As a first experiment, we wanted to see if robot-related characteristics had any effect on participants' behaviour at all, before proceeding with more narrowed down research questions.

3.1 The game of chicken

Participants played a simplified version of the game of chicken[1]. In the classic Game Theory formulation, two drivers, A and B, drive towards each other on a collision course: one must swerve, or they will crash, but if one driver swerves and the other does not, the one who swerved will be called a 'chicken', while the other will be a hero. The preferred outcome for A is thus for B to swerve, thus ensuring that A survives and keeps its face; and the preferred outcome for B is for A to swerve.

	swerve	straight
swerve	0, 0	+2, −2
straight	−2, +2	−1000, −1000

Table 1: A payoff matrix of the game of chicken. Numerical payoffs are arbitrarily set (the benefit of winning is +2, the cost of chickening is -2, and the cost of colliding is -1000).

Game theory [60] provides mathematical tools to study the interaction between decision makers, where the decisions taken by one may impact the other's interests. A game is a mathematical object consisting of *players*, *strategies* and *payoffs* [62]. Formally (we reuse here the notations in [22]), the game of chicken is a two-player game with two actions for player A and two actions for player B (go straight or swerve). A pair of matrices $P, Q \in \mathbb{R}^4$ can represent the payoffs of the two players, where p_{ij} and q_{ij} are the payoffs of players A and B respectively when player A plays action i and player B plays action j. Such a payoff matrix is depicted in Table 1 representing both matrices P and Q, where row player is player A, column player is player B, and where $p_{11} = 0$, $q_{11} = 0$, $p_{12} = +2, \dots,$ $q_{22} = -1000$, action 1 and 2 are to swerve and go straight, respectively. The set of possible strategies of the two players is denoted by X and Y, respectively, i.e., $X = \{x \in \mathbb{R}^2_{\geq 0} : \sum_{i=1}^{2} x_i = 1\}$ and $Y = \{y \in \mathbb{R}^2_{\geq 0} : \sum_{i=1}^{2} y_i = 1\}$, where x_1 denotes the (stochastic) strategy for player A to swerve, x_2 the (stochastic) strategy for player A to go straight, y_1 and y_2 the (stochastic) strategies for player B to swerve and go straight respectively.

The interest of game theoretical methods is to find optimal strategies for both players. Formally, strategy $x \in X$ of the row player is a best response to strategy $y \in Y$ of the column player if for all $x' \in X$, $p(x, y) \geq p(x', y)$. The concept of the best response for the column player is defined analogously. A pair of strategies $(x, y) \in X \times Y$ such that x is the best response to y and y is the best response to x called a (Nash) equilibrium. In the case of the game of chicken, there are 2 Nash equilibria: for player A to go straight and player B to swerve, or for player B to go straight and player A to swerve.

However, similar game-theoretic optimal solutions have been shown not to hold in real social situations. One such example is the Dictator Game. In this game, a participant is offered a sum of money by a dictator, and only if he/she accepts it they will both get the corresponding amount; therefore, the rational strategy would be to accept any sum of money being offered, since a little money is always better than no money [58]. However, multiple experiments showed that people would rather reject unfair offers, even though this means that they will not get any money [11, 32]. It is, therefore, possible that people will not behave following the optimal strategy when presented with a similar dilemma. Therefore, for our version, we did not give participants the traditional chicken payoff matrix (Table 1), and simply observed participants' unprompted behaviour when walking towards a virtual robot.

3.2 Procedure

Participants were recruited on the online platform Amazon Mechanical Turk (AMT). When they accepted to take part in the study, they were redirected to a webpage that contained a digital consent form; then they were asked to read the game instructions and press a SPACE key, to make sure that the device they were using contained one. Then they could personalise their own avatar in terms of gender and skin tone. This was done to ensure that people belonging to different demographic groups could feel equally represented in the game. Then there was one test trial, where both participants and the incoming agent were represented as geometric shapes, which allowed participants to get acquainted with the swerving movement and with the game environment. Finally, they completed 20 trials of the game.

Each trial (including the test trials) proceeded as follows: participants walked on the same path as a robot, but in the opposite direction, by means of a virtual avatar (Figure 1). They could swerve by pressing the SPACE key. Whenever a swerving action was made (either by the participant or the robot), the direction of the swerving (to the left or right) was chosen at random. Participants' decision to swerve or not, as well as the distance from the robot at the time of swerving, were the dependent variables in our analysis. Participants interacted with two different robots for 20 trials (10 per robot); the robots were chosen based on their anthropomorphism level in the ABOT database [67]: PR2 (which has an overall anthropomorphism score of 20/100) and Sophia (78.88/100). Furthermore, for half the

[1]This game was famously played on the big screen by James Dean in "Rebel without a cause".

trials participants were told that the incoming robot was controlled by an AI, and for the other half that the robot was controlled by another Mechanical Turk worker, who was playing in real time. Since the human and robot avatars start from a large distance from each other, we added a zoomed-in view of the robot so that participants had enough time to see the robot before the decision to swerve had to be made (Figure 1). Thus, the experiment followed a *2 (anthropomorphism: PR2 or Sophia) by 2 (perceived autonomy: autonomous or teleoperated)* within-subject design. The trial order was randomised across participants, and each robot appeared with a new colour at each trial, in order to increase the impression that participants were interacting with a new agent at each trial.

After the last trial, participants answered some questions about their gender, age, country of origin and experience with robots, and were given the unique survey code that they could use to be paid in AMT. The whole experiment lasted approximately 10 minutes, and workers were compensated $2, which is higher than US minimum wage and than average Mechanical Turk pay rate [25].

It is worth mentioning why we did not implement a 'control' human condition. Firstly, having participants interact with incoming human agents/avatars would have created several issues, in that we would have had to arbitrarily decide on the appearance (e.g. in terms of gender expression and skin tone) of this agent/avatar, which would have added a layer of unpredictability to participants' behaviour. Secondly, whilst it would be interesting to compare how people behave with robotic and human agents/avatars, it would not have contributed to our final goal of generating robot trajectories.

3.3 Task implementation

The game was developed in Unity version 2019.4.1f1 and exported to Webgl. The 3D model for Sophia was purchased from TurboSquid, while that for PR2 was community-sourced from The Modular OpenRobots Simulation Engine [47]. Screenshots from the game are shown in Figure 1.

3.3.1 Robot behaviour generation. The robot agent was designed to make a decision in each trial about when to swerve. At what point the robot decided to swerve was made to depend on the distance d between the two agents. The distance was sampled from a normal distribution:

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2 \big/ 2\sigma^2} \tag{1}$$

where the mean μ of the distribution was the minimum possible distance between the agents where an agent could still swerve to avoid a collision assuming the other agent did not swerve. The standard deviation σ was chosen so that if the player did not swerve, approximately one third of all trials would result in a collision (Figure 2). The distance at which the robot would swerve was sampled once for each trial, resulting in unpredictability in the robot's behaviour in regard to swerving. This, together with having the robots appear with different colours at each trial, was done to avoid a learning effect in participants.

Each of the two robot 3D models (PR2 and Sophia), as well as each of the participants' avatars, were encapsulated in an invisible cylinder-shaped collision object that stayed consistent across the trials. This object was used in Unity to calculate the occurrence of a collision. The collision was considered in the horizontal plane.

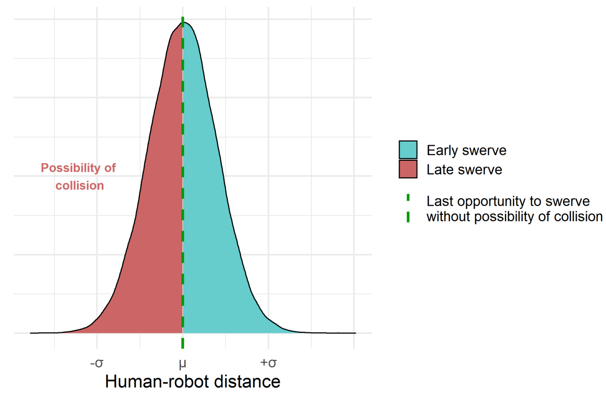

Figure 2: Distribution of the robot swerving distance behaviour.

3.4 Pilots

Prior to the main experiment, we conducted two pilot studies, with 20 participants each. The first pilot was a manipulation check for the perceived autonomy condition; specifically, after people played the 20 rounds of the game of chicken, we asked them if during the game it was clear to them which robots were controlled by an AI, and which by another player. All participants answered that this information was clear, and that they paid attention to it.

The second pilot was meant to design the most appropriate collision dynamics. In the first drafts of the experiment, we noticed that the visual representation of the collision between the two avatars was not very accurate, in that it looked like the two avatars were not physically touching. To see if this was impacting participants' reactions, we ran a pilot with a bird's eye view of the scene, instead of a frontal view. We hypothesised that, if the collision dynamics were good enough, people would behave similarly in both views – in other words, they would behave based on their internal representation of the robot's movements, and therefore the collision dynamics were good enough; instead, if people behaved differently in the two conditions, it might mean that they were behaving based on what they were seeing, on the actual visual representation of the robot. If this was the case, then we would need to make sure that the visual representation of the collision was more accurate. Indeed, we found that people were behaving differently when they had a frontal view, and a view from the top; specifically, they were swerving closer to the two robots in the view from the top (linear regression: $\chi^2(1) = 15.48, p < .001, M_{\text{top}} = 6.39, M_{\text{frontal}} = 14.23$). Therefore, we updated the collision dynamics so that the two avatars would appear to be actually touching in the case of a crash (Figure 1C and 1F), and we kept the frontal view for our main experiment.

We also used these two pilots to conduct an *a priori* power analysis to determine the sample size of the main experiment, using the R package `sjstats`, function `smpsize_lmm` [51]. We tested a linear mixed-effects model with 20 repeated trials, a medium effect size ($R^2 = .54$), the desired power (0.8) and an alpha of .05; we used distance from robot as dependent variable, robot type and autonomy level as predictors, and participant id as random intercept. Since we have two different outcome variables – the binary decision to swerve or not, and the distance from the robot when swerving – we computed the minimal power, i.e. the probability of detecting effects of at least a particular size on at least one outcome. Thus, we only ran one power analysis on the latter outcome. The result showed that a total sample of 110 participants would be required to achieve an 80% power.

3.5 Participants

We recruited 110 participants from Amazon Mechanical Turk. Of these, 4 reported technical issues and were therefore excluded from the analyses. Of the remaining 106, there were 34 participants who identified as females, 71 as males, and 1 as non-binary; their age ranged from 18 to 73 years old, with a median of 35 years old. The majority of participants reported having little to no experience with robots, or only having seen robots in the media (N = 63); 42 people reported having interacted with a robot before, and 1 reported interacting with robots on a regular basis. However, none of the participants reported having seen the robots used in the study before. Given the participant imbalance on previous experience with robots, we grouped them into two categories: "some experience" (people who interacted with a robot before, or who interact with robots on a regular basis, N = 43) and "no experience" (people who had never seen a robot before, or who had only seen a robot in the media, N = 63). Finally, since the experiment took place in a virtual environment, we also asked participants about their experience with games: 25 reported playing games every day, 33 often, 34 sometimes, 12 rarely and 2 never. All these individual differences were used as covariates in the statistical analyses.

4 RESULTS

All analyses were conducted in R version 4.0.1 [68]. Our outcomes were (i) whether participants swerved or not and (ii) if so, the distance from the robot when the player swerved. The independent variables were robot type (PR2 or Sophia) and perceived autonomy (autonomous or teleoperated). For the first outcome, we ran a mixed-effects logistic regression (R function `glmer`) with swerving as dependent variable, robot type and perceived autonomy as predictors, and participant id as random intercept. For the second outcome, we ran a mixed-effects linear regression (R function `lmer`) with distance from the robot as a dependent variable, robot type (Sophia or PR2) and perceived autonomy (autonomous or teleoperated) as predictors, and participant id as a random intercept. We added random intercepts to account for within-subject variability over repeated measures. The ICC (from the R package `performance`) was 0.31 for the linear model and 0.53 for the logistic model, suggesting that the inclusion of random effects was appropriate. We added participants' demographics characteristics (age, gender, experience with robots, experience with games) as covariates in both models, after establishing the presence of main effects. Significant effects, where present, were maintained after inclusion of these covariates in the models.

4.1 Decision to swerve

We found a main effect of perceived autonomy, with people swerving more often when they thought the robot was controlled by a human ($OR = 1.44, z = 2.16, p = .03$, Figure 3). There was no main effect of robot type, however ($OR = 0.84, z = -1.05, p = .30$). There was also no interaction between robot type and perceived autonomy ($OR = 0.83, z = -0.82, p = .41$), and no additional effects of participants' demographics.

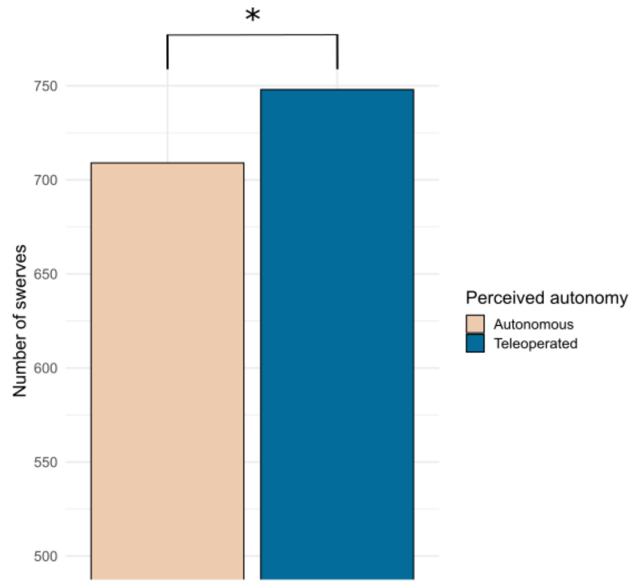

Figure 3: Number of swerves by level of perceived autonomy. '*' indicates $P < .05$.

4.2 Distance from robot when swerving

We found a main effect of robot type ($\chi^2(1) = 22.26, p < .001$): as can be seen from Figure 4, participants swerved farther away from PR2 than Sophia ($M_{\text{PR2}} = 6.41, M_{\text{Sophia}} = 5.24$). There was no effect of perceived autonomy ($\chi^2(1) = 1.37, p = .24$) and no interaction between robot type and perceived autonomy ($\chi^2(1) = 3.42, p = .18$). Adding participants' demographics characteristics as covariates in the model revealed no additional main effects.

5 DISCUSSION

In the current study we designed an experiment based on the game of chicken to observe how participants behave in relation to a robot that is walking on the same path, but in the opposite direction. We found that participants' behaviour changed based on the robot's anthropomorphism and perceived autonomy. While participants' behaviour is interesting from a human psychology point of view, we can also use these observations to model socially-appropriate navigation behaviour for robots [e.g. 13, 57, 73, 82].

Answering RQ2, we found that participants decided to swerve more often when the incoming robot was perceived to be controlled by another human, rather than being autonomous (Section 4.1). As hypothesised, since swerving to give way can be seen as an act of politeness, it might be that people were more socially compliant when they thought they were actually interacting with another human. The current results, together with findings from previous studies in Human-Agent Interaction [14, 16–18, 23, 33], support

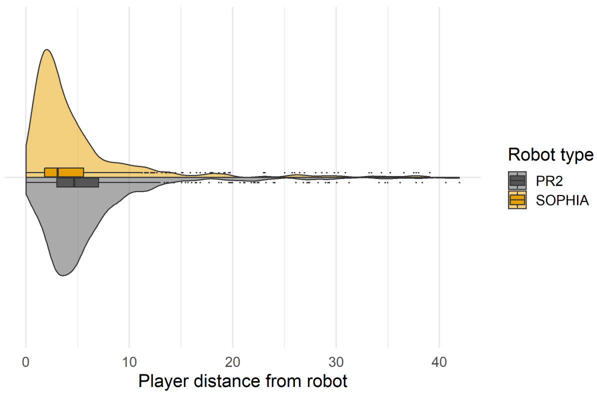

Figure 4: Boxplots and distributions of the distances from the robot when the participant decided to swerve. Vertical bars within a box indicate the median, horizontal bars indicate standard errors, dots indicate outliers.

Blascovich's "Threshold model of social influence" [7] over "Computers As Social Actors" (CASA) [63]. The former posits that social influence will be greater the higher the perceived agency of an artificial agent, while the latter suggests that people start applying automatic social behaviour as soon an interaction with an artificial agent includes social cues.

This finding has practical implications for the field of HRI: robots that are currently being tested and deployed in human-populated environments are often operated from a safe distance, in a Wizard-of-Oz fashion [29, 85], or allow remote telepresence of people located elsewhere [42]. In our simulated encounters with robots, people behaved differently if they thought the robot was controlled by another human, or had a 'mind' of its own. Thus, teleoperated robots might need to follow different navigation plans than autonomous robots. This is consistent with previous observations that people were more trusting and behaved more recklessly with autonomous vehicles than with human-driven cars, because there is no human error involved in the former [71]. Similarly, when there were errors in a human-robot collaborative task, people attributed less blame to themselves, and more to the robot, when their robotic partner was autonomous [38].

Our results seem to contrast with previous findings that people perceive autonomous and teleoperated robots in a similar fashion [6]. However, while in the study by Bennett et al. the robot was always operated by an experimenter, in the current study we told participants that the robot was controlled by another Mechanical Turk worker. Consistent with Social Identity Theory [77], it is possible that our participants behaved more politely with people who they believed belonged to their same social group, while participants in Bennett et al. [6] might have considered the experimenter and autonomous robot to be equally outside of their identity group.

While robot type did not have an effect on participants' decision to swerve or not, when participants decided to swerve, they swerved closer to Sophia than PR2, thus answering RQ1 (Section 4.2). Sophia is more human-like than PR2 – according to the ABOT database, Sophia is in the top 25% most human-like robots, while PR2 is in the bottom 25% [67]. People generally attribute more social competencies to more human-like agents [e.g. 41, 76]. Therefore, supporting our hypothesis, it is possible that participants were expecting Sophia to swerve first, signalling politeness towards them as a participant. Only when it became apparent that Sophia was not going to swerve (or was going to swerve at an uncomfortable distance) did participants decide to swerve themselves. On the other hand, people might have had lower expectations of PR2 [cf. 28], thus deciding to swerve earlier themselves in case the robot was not capable of doing it. This is consistent with previous studies showing that people keep a bigger distance from more mechanical-looking and -sounding robots [84]. Also, since this effect of robot type only appeared after the decision to swerve had been made, it is possible that people who decided to behave 'conservatively' (by swerving) were also 'conservative' in swerving earlier with a robot that could have evoked feelings of 'threat' [84, 89]. Another interpretation is that mechanical-looking robots elicit expectations of rationality in their behaviour [cf. 54]; therefore, participants might have predicted the mechanical-looking robot to make a more rational choice, which in this case could be to swerve at the last possible moment before collision. As a robot swerving at the last possible moment will likely produce a breach in personal space, participants might have opted for swerving earlier themselves.

Similarly, another interesting finding can be drawn from observing people's distances when they choose to swerve: as can be seen from Figure 4, the distribution of distances from PR2 is wider than from Sophia, showing more variability in when people chose to swerve with PR2. The fact that people were more consistent in their swerving behaviour with Sophia suggests that they were more certain about their actions with this robot. It might be that Sophia's level of anthropomorphism communicated a first impression of human-likeness, which prompted participants to only swerve when they were getting too close to the robot. On the other hand, participants might have been less sure about PR2's intentions, hence the more widely varied behaviour from the participants. Since none of the participants had seen the two robots used in the study before, we can assume that this familiarity effect could have been due to Sophia's human-likeness, whereas participants were not sure how a mechanical robot such as PR2 would behave.

It is also interesting that we did not find any effect of individual differences in participants' behaviour, suggesting that these results can be extended to a wide population. This is consistent with previous studies in human-human interaction, which found no effect of individual differences in collision avoidance strategies [39]. However, our results contrast with other studies showing that, for example, gender and previous experience with robots play a role in interpersonal human-robot distance, with men and people with less experience keeping a bigger distance [59, 78]. As discussed in Section 2, previous studies on social models for robot navigation focused mostly on human-related and environment-related factors. Thus, the interplay between these and robot-related factors remains to be explored. We plan to use the current results as a baseline human behaviour, that we can further examine in combination with other manipulations, such as type of environment (corridor vs. open space) and agent motivations (e.g. robot or participant needing to prioritise speed vs. accuracy).

Finally, of the 106 analysed participants, 77 (equalling to 73% of participants) collided with the incoming robots at least once, with a median of 3 collisions per participant. This provides yet further evidence that people are not rational agents, and that mathematical models of human-human collaboration and conflict often do not work in real life [11, 32]. In this case, the game-theoretic model of the game of chicken would have predicted that people try to avoid collision at all costs, as it generates the worst payoff for both agents (Table 1) [69]. It is possible that these collisions were a by-product of the interaction taking place in a virtual environment, and that, had the experiment taken place in person, people would not have crashed into a real robot. It is also possible that participants tried to explicitly see what would happen in case of collision. Or, they simply decided to swerve when it was already too late. Whichever the reason for this behaviour, it highlights the importance of taking into account people's (at times) irrational behaviour when designing machines that need to interact with them.

5.1 Limitations

There are a couple of limitations that should be mentioned. First and foremost, while we had originally intended to run this study in person, this was not possible due to restrictions in light of the COVID-19 pandemic. While we still plan to continue this line of research and apply our findings to in-person studies (see Section 5.2), it might not be possible to fully replicate the current design, due to ethical concerns about the possibility of people physically colliding with a robot. Furthermore, it remains to be seen whether robot-related characteristics will influence people's behaviour in the same manner when interacting with a real robot. In fact, it has been suggested that people behave differently when interacting with a robot in-presence and in telepresence, for example in terms of proxemics [4], despite more general evidence that participants' online behaviour is comparable to offline [e.g. 36, 48]. Nonetheless, we still believe that our results are an important first step in considering robot-related factors in social robot navigation scenarios.

Secondly, it is possible that the different distances we observed with Sophia and PR2 were due to features other than anthropomorphism, namely size, sophistication or gender expression. Regarding size, PR2 is bigger than Sophia (Fig. 1), so it is possible that the earlier swerves with PR2 might have been due to people affording it more personal space. Also, Hegel [27] suggests that the perception of social capabilities might not be dependant on anthropomorphism, but rather on sophistication. The differences we found might also have been due to perceived robot gender. Although PR2 is not explicitly gendered, previous studies suggest that people consistently consider it to be male [56, 80]. Thus, it is possible that participants might have swerved closer to Sophia than PR2 because they were expecting Sophia to swerve first, consistent with gender-based attributions of submission. This would be in line with previous studies showing that people keep a smaller distance from females than males [35]. However, we did not find any interaction between participants' gender and robot type, so the potential presence of gender stereotypes would be independent of participants' own gender. Future studies will need to discriminate this potential confound by examining people's swerving behaviour with a wider range of robots, covering different levels of anthropomorphism, size, sophistication, and gender.

Finally, the human avatar appearance was not realistic (Fig. 1). This could have affected participants' behaviour a) in terms of judging distances from the robots and b) in terms of not fully identifying with the avatar, resulting in behaviour that they would not necessarily perform in real life.

5.2 Future work

As future work, we will implement the results of this study into robot controllers. Game theory is popular in control of robot systems, especially since controllers can be viewed as decision-making entities [55, 79]. In the context of robots completing tasks in shared environments with humans, game theory appears to be an interesting choice to deal with uncertainty concerning human behaviours and has been successfully implemented in automated vehicle motion planning applications [19, 61, 87] and interactive motion behaviour for robots in physical contact with human users [50].

As shown in Figure 4, the ideal distance for a robot to swerve can be determined depending on the robot's embodiment and on whether the user thinks the robot is autonomous. The ideal swerve distance can be implemented in the robot controller to swerve optimally using constrained motion planning [21], and encapsulating the ideal distance to swerve into e.g. temporal logic constraints [52]. This way, we will develop more socially sophisticated robot motion planning algorithms.

Further, we will also investigate to what extent a robot can make the human swerve first, and with a different distance. Indeed, in some cases, the robot may have to accomplish an urgent task (transportation of medicines in a hospital, urgent delivery, etc.). Therefore, some ways of communicating such an emergency (employing warning sounds or red light, for instance) can be used to notify humans in the robot's environment of the priority of the robot's task. We plan to investigate ways of communicating such emergencies in future work. We will also focus on the perceived safety of robots in human-crowded environments.

6 CONCLUSION

By studying how people swerve to avoid crashing into a robot, or to avoid entering an uncomfortable distance from it, we can design novel robot trajectories that anticipate the human need for safety and comfort. In particular, in the current study we have shown that robots' 'chicken' behaviour should also depend on their human-likeliness and perceived autonomy. In other words, these robot-related factors seem to play a role in how people navigate a shared space with a robot. These factors should be integrated into robot navigation models that account for human-related and environment-related factors, in order to generate truly socially-aware robots that navigate in society.

ACKNOWLEDGMENTS

The work was funded by the WASP Expedition Project on Correct-by-design and Socially Acceptable Autonomy (CorSA).

REFERENCES

[1] Jeremy N. Bailenson and Jim Blascovich. 2004. *Avatars. Encyclopedia of human-computer interaction.* Berkshire Publishing Group.

[2] Jeremy N. Bailenson, Jim Blascovich, Andrew C. Beall, and Jack M. Loomis. 2001. Equilibrium theory revisited: Mutual gaze and personal space in virtual environments. *Presence: Teleoperators & Virtual Environments* 10, 6 (2001), 583–598.

[3] Jeremy N. Bailenson, Jim Blascovich, Andrew C. Beall, and Jack M. Loomis. 2003. Interpersonal distance in immersive virtual environments. *Personality and social psychology bulletin* 29, 7 (2003), 819–833.

[4] Wilma A. Bainbridge, Justin W. Hart, Elizabeth S. Kim, and Brian Scassellati. 2011. The benefits of interactions with physically present robots over video-displayed agents. *International Journal of Social Robotics* 3, 1 (2011), 41–52.

[5] Jenay M. Beer, Arthur D. Fisk, and Wendy A. Rogers. 2014. Toward a framework for levels of robot autonomy in human-robot interaction. *Journal of Human-Robot Interaction* 3, 2 (2014), 74.

[6] Maxwell Bennett, Tom Williams, Daria Thames, and Matthias Scheutz. 2017. Differences in interaction patterns and perception for teleoperated and autonomous humanoid robots. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '17)*. IEEE, 6589–6594.

[7] Jim Blascovich, Jack M. Loomis, Andrew C. Beall, Kimberly R. Swinth, Crystal L. Hoyt, and Jeremy N. Bailenson. 2002. Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry* 13, 2 (2002), 103–124.

[8] Jim Blascovich and Cade McCall. 2013. *Social influence in virtual environments.* New York: Oxford University Press, 305–315.

[9] John Travis Butler and Arvin Agah. 2001. Psychological effects of behavior patterns of a mobile personal robot. *Autonomous Robots* 10, 2 (2001), 185–202.

[10] Zachary Byers, Michael Dixon, Kevin Goodier, Cindy M. Grimm, and William D. Smart. 2003. An autonomous robot photographer. In *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '03, Vol. 3)*. IEEE, 2636–2641.

[11] Colin F. Camerer. 2011. *Behavioral Game Theory: Experiments in Strategic Interaction.* Princeton University Press.

[12] Jung Ju Choi, Yunkyung Kim, and Sonya S. Kwak. 2014. The autonomy levels and the human intervention levels of robots: The impact of robot types in human-robot interaction. In *Proceedings of the 23rd IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN '14)*. IEEE, 1069–1074.

[13] Kerstin Dautenhahn, Michael Walters, Sarah Woods, Kheng Lee Koay, Chrystopher L. Nehaniv, E. A. Sisbot, Rachid Alami, and Thierry Siméon. 2006. How may I serve you? A robot companion approaching a seated person in a helping context. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-Robot Interaction (HRI '06)*. 172–179.

[14] Celso M. De Melo, Peter J. Carnevale, and Jonathan Gratch. 2013. Agent or Avatar? Using Virtual Confederates in Conflict Management Research. In *Academy of Management Proceedings*, Vol. 2013. Academy of Management Briarcliff Manor, NY 10510, 10859.

[15] Celso M. De Melo, Peter J. Carnevale, Stephen J. Read, and Jonathan Gratch. 2014. Reading people's minds from emotion expressions in interdependent decision making. *Journal of Personality and Social Psychology* 106, 1 (2014), 73.

[16] Celso M. De Melo, Jonathan Gratch, and Peter J. Carnevale. 2014. Humans versus computers: Impact of emotion expressions on people's decision making. *IEEE Transactions on Affective Computing* 6, 2 (2014), 127–136.

[17] Celso M. De Melo, Jonathan Gratch, and Peter J Carnevale. 2014. The Importance of Cognition and Affect for Artificially Intelligent Decision Makers.. In *28th AAAI Conference on Artificial Intelligence*. 336–342.

[18] Celso M. De Melo, Peter Khooshabeh, Ori Amir, and Jonathan Gratch. 2018. Shaping Cooperation between Humans and Agents with Emotion Expressions and Framing. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems.* 2224–2226.

[19] Clément Dextreit and Ilya V Kolmanovsky. 2013. Game theory controller for hybrid electric vehicles. *IEEE Transactions on Control Systems Technology* 22, 2 (2013), 652–663.

[20] Lorin D. Dole, David M. Sirkin, Rebecca M. Currano, Robin R. Murphy, and Clifford I. Nass. 2013. Where to look and who to be designing attention and identity for search-and-rescue robots. In *Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI '13)*. IEEE, 119–120.

[21] Cameron Finucane, Gangyuan Jing, and Hadas Kress-Gazit. 2010. LTLMoP: Experimenting with language, temporal logic and robot control. In *International Conference on Intelligent Robots and Systems*. IEEE, 1988–1993.

[22] Felix Fisher. 2015. Lecture notes in Optimization.

[23] Jesse Fox, Sun Joo Ahn, Joris H. Janssen, Leo Yeykelis, Kathryn Y. Segovia, and Jeremy N. Bailenson. 2015. Avatars versus agents: A meta-analysis quantifying the effect of agency on social influence. *Human–Computer Interaction* 30, 5 (2015), 401–432.

[24] Edward Twitchell Hall. 1966. *The hidden dimension*. Vol. 609. Garden City, NY: Doubleday.

[25] Kotaro Hara, Abigail Adams, Kristy Milland, Saiph Savage, Chris Callison-Burch, and Jeffrey P. Bigham. 2018. A data-driven analysis of workers' earnings on Amazon Mechanical Turk. In *Proceedings of the 2018 SIGCHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, 1–14.

[26] Kerstin Sophie Haring, Yoshio Matsumoto, and Katsumi Watanabe. 2013. How do people perceive and trust a lifelike robot. In *Proceedings of the World Congress on Engineering and Computer Science*, Vol. 1.

[27] Frank Hegel. 2012. Effects of a robot's aesthetic design on the attribution of social capabilities. In *Proceedings of the 21st IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN '12)*. IEEE, 469–475.

[28] Frank Hegel, Sören Krach, Tilo Kircher, Britta Wrede, and Gerhard Sagerer. 2008. Theory of mind (ToM) on robots: A functional neuroimaging study. In *Proceedings of the 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI '08)*. IEEE, 335–342.

[29] Päivi Heikkilä, Hanna Lammi, Marketta Niemelä, Kathleen Belhassein, Guillaume Sarthou, Antti Tammela, Aurélie Clodic, and Rachid Alami. 2019. Should a robot guide like a human? A qualitative four-phase study of a shopping mall robot. In *International Conference on Social Robotics*. Springer, 548–557.

[30] Thomas Hellström. 2013. On the moral responsibility of military robots. *Ethics and Information Technology* 15, 2 (2013), 99–107.

[31] L. F. Henderson. 1971. The statistics of crowd fluids. *Nature* 229, 5284 (1971), 381–383.

[32] Johannes Hewig, Nora Kretschmer, Ralf H. Trippe, Holger Hecht, Michael G. H. Coles, Clay B. Holroyd, and Wolfgang H. R. Miltner. 2011. Why humans deviate from rational choice. *Psychophysiology* 48, 4 (2011), 507–514.

[33] Crystal L. Hoyt, Jim Blascovich, and Kimberly R. Swinth. 2003. Social inhibition in immersive virtual environments. *Presence: Teleoperators & Virtual Environments* 12, 2 (2003), 183–195.

[34] Helge Hüttenrauch, Kerstin Severinson Eklundh, Anders Green, and Elin A Topp. 2006. Investigating spatial relationships in human-robot interaction. In *Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 5052–5059.

[35] Tina Iachini, Yann Coello, Francesca Frassinetti, Vincenzo Paolo Senese, Francesco Galante, and Gennaro Ruggiero. 2016. Peripersonal and interpersonal space in virtual and real environments: Effects of gender and age. *Journal of Environmental Psychology* 45 (2016), 154–164.

[36] Patrik Jonell, Taras Kucherenko, Ilaria Torre, and Jonas Beskow. 2020. Can we trust online crowdworkers? Comparing online and offline participants in a preference test of virtual agents. In *International Conference on Intelligent Virtual Agents (IVA '20)*. ACM.

[37] Harmish Khambhaita and Rachid Alami. 2020. Viewing robot navigation in human environment as a cooperative activity. In *Robotics Research*. Springer, 285–300.

[38] Taemie Kim and Pamela Hinds. 2006. Who should I blame? Effects of autonomy and transparency on attributions in human-robot interaction. In *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN '06)*. IEEE, 80–85.

[39] Alexander G. Knorr, Lina Willacker, Joachim Hermsdörfer, Stefan Glasauer, and Melanie Krüger. 2016. Influence of person-and situation-specific characteristics on collision avoidance behavior in human locomotion. *Journal of Experimental Psychology: Human Perception and Performance* 42, 9 (2016), 1332.

[40] Kheng Lee Koay, Dag Sverre Syrdal, Michael L. Walters, and Kerstin Dautenhahn. 2007. Living with robots: Investigating the habituation effect in participants' preferences during a longitudinal human-robot interaction study. In *Proceedings of the 16th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN '07)*. IEEE, 564–569.

[41] Sören Krach, Frank Hegel, Britta Wrede, Gerhard Sagerer, Ferdinand Binkofski, and Tilo Kircher. 2008. Can machines think? Interaction and perspective taking with robots investigated via fMRI. *PloS one* 3, 7 (2008), e2597.

[42] Annica Kristoffersson, Silvia Coradeschi, and Amy Loutfi. 2013. A review of mobile robotic telepresence. *Advances in Human-Computer Interaction* 2013 (2013).

[43] Thibault Kruse, Amit Kumar Pandey, Rachid Alami, and Alexandra Kirsch. 2013. Human-aware robot navigation: A survey. *Robotics and Autonomous Systems* 61, 12 (2013), 1726–1743.

[44] Sonya S. Kwak, Yunkyung Kim, Eunho Kim, Christine Shin, and Kwangsu Cho. 2013. What makes people empathize with an emotional robot?: The impact of agency and physical embodiment on human empathy for a robot. In *Proceedings of the 22nd IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN '13)*. IEEE, 180–185.

[45] Hyewon Lee, Jung Ju Choi, and Sonya S. Kwak. 2015. Can human jobs be taken by robots?: The appropriate match between robot types and task types. *Archives of Design Research* 28, 3 (2015), 49–59.

[46] Minha Lee, Gale Lucas, Johnathan Mell, Emmanuel Johnson, and Jonathan Gratch. 2019. What's on Your Virtual Mind? Mind Perception in Human-Agent Negotiations. In *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents (IVA '19)*. 38–45.

[47] Séverin Lemaignan, Gilberto Echeverria, Michael Karg, Jim Mainprice, Alexandra Kirsch, and Rachid Alami. 2012. Human-robot interaction in the MORSE simulator. In *Proceedings of the 2012 ACM/IEEE International Conference on Human-Robot Interaction (HRI ’12)*. ACM, 181–182.

[48] Jamy Li. 2015. The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *International Journal of Human-Computer Studies* 77 (2015), 23–37.

[49] Rui Li, Marc van Almkerk, Sanne van Waveren, Elizabeth Carter, and Iolanda Leite. 2019. Comparing human-robot proxemics between virtual reality and the real world. In *Proceedings of the 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI ’19)*. IEEE, 431–439.

[50] Y Li, G Carboni, F Gonzalez, D Campolo, and E Burdet. 2019. Differential game theory for versatile physical human–robot interaction. *Nature Machine Intelligence* 1, 1 (2019), 36–43.

[51] Daniel Lüdecke. 2020. *sjstats: Statistical Functions for Regression Models (Version 0.18.0)*. https://doi.org/10.5281/zenodo.1284472

[52] Oded Maler and Dejan Nickovic. 2004. Monitoring temporal properties of continuous signals. In *Formal Techniques, Modelling and Analysis of Timed and Fault-Tolerant Systems*. Springer, 152–166.

[53] Bertram F. Malle, Matthias Scheutz, Thomas Arnold, John Voiklis, and Corey Cusimano. 2015. Sacrifice one for the good of many? People apply different moral norms to human and robot agents. In *Proceedings of the 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI ’15)*. IEEE, 117–124.

[54] Bertram F. Malle, Matthias Scheutz, Jodi Forlizzi, and John Voiklis. 2016. Which robot am I thinking about? The impact of action and appearance on people’s evaluations of a moral robot. In *Proceedings of the 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI ’16)*. IEEE, 125–132.

[55] Jason R Marden and Jeff S Shamma. 2018. Game theory and control. *Annual Review of Control, Robotics, and Autonomous Systems* 1 (2018), 105–134.

[56] Conor McGinn and Ilaria Torre. 2019. Can you Tell the Robot by the Voice? An Exploratory Study on the Role of Voice in the Perception of Robots. In *Proceedings of the 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI ’19)*. 211–221.

[57] Marek P. Michalowski, Selma Sabanovic, and Reid Simmons. 2006. A spatial model of engagement for a social robot. In *9th IEEE International Workshop on Advanced Motion Control*. IEEE, 762–767.

[58] Oskar Morgenstern and John Von Neumann. 1953. *Theory of games and economic behavior*. Princeton University Press.

[59] Jonathan Mumm and Bilge Mutlu. 2011. Human-robot proxemics: physical and psychological distancing in human-robot interaction. In *Proceedings of the 6th ACM/IEEE International conference on Human-Robot Interaction (HRI ’11)*. 331–338.

[60] Roger B Myerson. 1991. Game theory: analysis of conflict. (1991).

[61] Xiaoxiang Na and David J Cole. 2014. Game-theoretic modeling of the steering interaction between a human driver and a vehicle collision avoidance controller. *IEEE Transactions on Human-Machine Systems* 45, 1 (2014), 25–38.

[62] John Nash. 1951. Non-cooperative games. *Annals of mathematics* (1951), 286–295.

[63] Clifford I. Nass and Youngme Moon. 2000. Machines and mindlessness: Social responses to computers. *Journal of Social Issues* 56, 1 (2000), 81–103.

[64] Anne-Hélène Olivier, Antoine Marin, Armel Crétual, Alain Berthoz, and Julien Pettré. 2013. Collision avoidance between two walkers: Role-dependent strategies. *Gait & Posture* 38, 4 (2013), 751–756.

[65] Amit Kumar Pandey and Rachid Alami. 2010. A framework towards a socially aware mobile robot motion in human-centered dynamic environment. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 5855–5860.

[66] Christopher Peters, Fangkai Yang, Himangshu Saikia, Chengjie Li, and Gabriel Skantze. 2018. Towards the use of mixed reality for HRI design via virtual robots. In *Proceedings of the 1st International Workshop on Virtual, Augmented, and Mixed Reality for HRI (VAM-HRI)*.

[67] Elizabeth Phillips, Xuan Zhao, Daniel Ullman, and Bertram F. Malle. 2018. What is Human-like? Decomposing Robots’ Human-like Appearance Using the Anthropomorphic roBOT (ABOT) Database. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction (HRI ’18)*. 105–113.

[68] R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

[69] Anatol Rapoport and Albert M. Chammah. 1966. The game of chicken. *American Behavioral Scientist* 10, 3 (1966), 10–28.

[70] Jorge Rios-Martinez, Anne Spalanzani, and Christian Laugier. 2015. From proxemics theory to socially-aware navigation: A survey. *International Journal of Social Robotics* 7, 2 (2015), 137–153.

[71] Dirk Rothenbücher, Jamy Li, David Sirkin, Brian Mok, and Wendy Ju. 2016. Ghost driver: A field study investigating the interaction between pedestrians and driverless vehicles. In *Proceedings of the 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN ’16)*. IEEE, 795–802.

[72] Aziez Sardar, Michiel Joosse, Astrid Weiss, and Vanessa Evers. 2012. Don’t stand so close to me: users’ attitudinal and behavioral responses to personal space invasion by robots. In *Proceedings of the 7th annual ACM/IEEE International Conference on Human-Robot Interaction (HRI ’12)*. 229–230.

[73] Satoru Satake, Takayuki Kanda, Dylan F. Glas, Michita Imai, Hiroshi Ishiguro, and Norihiro Hagita. 2009. How to approach humans? Strategies for social robots to initiate interaction. In *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction (HRI ’09)*. 109–116.

[74] Matthias Scheutz. 2011. *The Inherent Dangers of Unidirectional Emotional Bonds between Humans and Social Robots*. MIT Press, 205–221.

[75] Emmanuel Senft, Satoru Satake, and Takayuki Kanda. 2020. Would You Mind Me if I Pass by You? Socially-Appropriate Behaviour for an Omni-based Social Robot in Narrow Environment. In *Proceedings of the 15th ACM/IEEE International Conference on Human-Robot Interaction (HRI ’20)*. 539–547.

[76] Valerie K. Sims, Matthew G. Chin, Heather C. Lum, Linda Upham-Ellis, Tatiana Ballion, and Nicholas C. Lagattuta. 2009. Robots’ auditory cues are subject to anthropomorphism. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 53. SAGE Publications Sage CA: Los Angeles, CA, 1418–1421.

[77] Henri Tajfel. 1970. Experiments in intergroup discrimination. *Scientific American* 223, 5 (1970), 96–103.

[78] Leila Takayama and Caroline Pantofaru. 2009. Influences on proxemic behaviors in human-robot interaction. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 5495–5502.

[79] Claire J Tomlin, John Lygeros, and S Shankar Sastry. 2000. A game theoretic approach to controller design for hybrid systems. *Proc. IEEE* 88, 7 (2000), 949–970.

[80] Ilaria Torre, Adrian Benigno Latupeirissa, and Conor McGinn. 2020. How context shapes the appropriateness of a robot’s voice. In *Proceedings of the 29th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN ’20)*. IEEE.

[81] Xuan-Tung Truong and Trung-Dung Ngo. 2016. Dynamic social zone based mobile robot navigation for human comfortable safety in social environments. *International Journal of Social Robotics* 8, 5 (2016), 663–684.

[82] Xuan-Tung Truong and Trung-Dung Ngo. 2017. “To approach humans?”: A unified framework for approaching pose prediction and socially aware robot navigation. *IEEE Transactions on Cognitive and Developmental Systems* 10, 3 (2017), 557–572.

[83] Michael L. Walters, Kerstin Dautenhahn, Kheng Lee Koay, Christina Kaouri, R. Boekhorst, Chrystopher Nehaniv, Iain Werry, and David Lee. 2005. Close encounters: Spatial distances between people and a robot of mechanistic appearance. In *Proceedings of the 5th IEEE-RAS International Conference on Humanoid Robots*. IEEE, 450–455.

[84] Michael L. Walters, Dag Sverre Syrdal, Kheng Lee Koay, Kerstin Dautenhahn, and R. Te Boekhorst. 2008. Human approach distances to a mechanical-looking robot with different robot voice styles. In *Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN ’08)*. IEEE, 707–712.

[85] Astrid Weiss, Regina Bernhaupt, Manfred Tscheligi, Dirk Wollherr, Kolja Kuhnlenz, and Martin Buss. 2008. A methodological variation for acceptance evaluation of human-robot interaction in public places. In *RO-MAN 2008-The 17th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 713–718.

[86] Astrid Weiss, Daniela Wurhofer, Michael Lankes, and Manfred Tscheligi. 2009. Autonomous vs. tele-operated: How people perceive human-robot collaboration with HRP-2. In *Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI ’09)*. IEEE, 257–258.

[87] Hongtao Yu, H. Eric Tseng, and Reza Langari. 2018. A human-like game theory-based controller for automatic lane changing. *Transportation Research Part C: Emerging Technologies* 88 (2018), 140 – 158. https://doi.org/10.1016/j.trc.2018.01.016

[88] Wenjian Yu and Anders Johansson. 2007. Modeling crowd turbulence by many-particle simulations. *Physical review E* 76, 4 (2007), 046105.

[89] Jakub Złotowski, Kumar Yogeeswaran, and Christoph Bartneck. 2017. Can we control it? Autonomous robots threaten human identity, uniqueness, safety, and resources. *International Journal of Human-Computer Studies* 100 (2017), 48–54.