Scammed By Robots: Can Deception by Two Simple Non-Humanoid Robots Shape General Attitudes toward Robots?

Nevo Heimann Saadon
Media Innovation Lab
Reichman University, Israel
nevo.hesa@milab.idc.ac.il

Benny Megidish
Media Innovation Lab
Reichman University, Israel
benny.megidish@post.runi.ac.il

Hadas Erel
Media Innovation Lab
Reichman University, Israel
hadas.erel@milab.idc.ac.il

***Abstract*—As multi-robot environments become more common, it is important to assess the robots' impact on the humans interacting with them and its magnitude. We investigated whether an interaction with two simple non-humanoid robots can be perceived as an intentional collusion aimed to deceive the person interacting with them. We evaluated if such a negative interaction would result in shaping people's perception of the robots in the interaction and extend to shaping their attitudes towards robots in general. In the study, participants played a gambling game with two simple non-humanoid robots: a "Dealer" responsible for card shuffling and an "Advisor" providing recommendations. We tested participants' experience when the "Advisor" robot gave a false recommendation in the final round, causing them to lose all their winnings. Our findings indicated that this single instance of perceived deception led to a negative perception of the "Advisor" robot, and also negatively influenced people's attitudes toward robots in general, at least temporarily. This study highlights the importance of carefully designing interactions in multi-robot environments, even when the interaction involves very simple robots. Careful consideration of ingroup-outgroup effects and verifying that the robots' collaboration is not interpreted as intentional deception is essential for avoiding powerful negative effects that reach beyond the specific interaction with the robots.**

***Index Terms*—Robot-Robot-Human Interaction, Multi-Robot, HRI, Deception, Non-humanoid robots**

I. Introduction

As robots increasingly become part of our daily lives, we are predicted to interact with more than one robot at the same time [1]. This highlights the need to map the possible impact of multi-robot interactions [2], [3]. Interactions between a human and several robots (even just two robots) introduce a complex social context involving various interpersonal relationships. Such multi-robot scenarios can potentially increase the interaction's social influence due to the dynamics between the robots [1], [4]–[6]. Several studies have begun to look into the implications of these socially complex interactions, highlighting significant outcomes, such as experiencing social exclusion by robots [7], triggering emotions and stereotypes [8], conforming to the robots' opinion [9], enhancing human compliance [10], and shaping the robots' perception [11].

[1]The person in Fig. 1.A is simulating a participant.

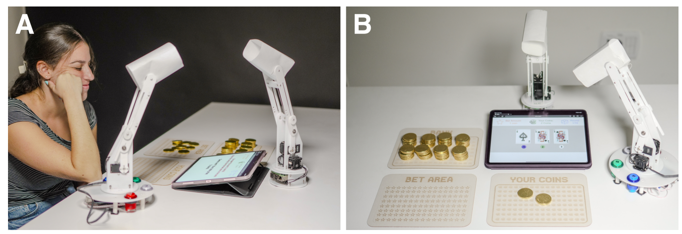


Fig. 1. A participant playing with two robots (A)[1]; Experimental settings (B)

When a person interacts with two robots, different collaboration dynamics may take place. Among these, the specific case where robots collaborate with each other while not collaborating with the person has been identified as a sensitive context that can lead to negative effects. For example, people may interpret the robots' collaboration as an act of ostracism and feel left out [2], [10]. Another possibility that may lead to even more drastic effects concerns people's interpretation of the robots' collaboration as an intentional act of colluding against them (i.e., collaborating dishonestly to cheat or deceive others [12]). The social context presented by such interactions may lead to powerful effects as it involves the two factors that are considered to be primary predictors of robots' social influence [6]: (1) Belonging - in this context, belonging to a different group than the robots who collaborate with each other but not with the person; (2) Violating expectations - as robots are not expected to collude against humans. Such interactions may, therefore, lead to powerful social influences, including changes in people's beliefs, behaviors, or attitudes [6], [13].

Previous studies have already indicated that an interaction with a robot can heavily impact people's sense of belonging due to ingroup-outgroup effects. It is well established that ingroup-outgroup effects are not exclusive to human interactions and also occur in groups that involve humans and robots [14], [15]. These studies, which typically involve several humans and a single robot, indicated that humans often perceive robots as outgroup members, which can lead to competitive and less cooperative behaviors when interacting with them [16], [17]. In the case of two robots collaborating against a person, several aspects of robots' ingroup-outgroup effects take place. First, robots inherently belong to a different group than humans due to their mechanical nature and appearance [14], [16], [18]. Furthermore, the robots' behavior is known to influence the perception of their group belonging [14], [17]. In the case of collaborating against a human, the robots' behavior involves a collaboration that does not include the human. Lastly, interpreting the robots' behavior as collusion would likely trigger a competitive context where they have a mutual goal that competes with the person's goal, further enhancing the person's low belonging to their group [6], [19].

The perception of robots as capable of intentional deception was also previously tested in the context of HRI [20]–[29]. These studies typically focused on interactions with a single robot [5], [23] and indicated that deception by robots is rather similar to deception by humans [4], [5], leading to considerable effects on trust [27], [30] and triggering negative emotions [31]. However, deception by robots was also indicated to have some unique effects related to the specific context of HRI [6], [13]. For example, dishonest robotic behaviors were perceived as an indication that the robot is intelligent and capable [5], [23]. This effect was associated with the perception of the robot as violating expectations [6]. When a robot deceives during a human-robot interaction, it violates the expectation that its behavior would have a positive impact on people and that it should comply with acceptable norms [13], [29], [32]. In the case of two robots collaborating against a person, the robots' collusion is predicted to violate expectations even further. Since violating people's expectations is one of the main factors leading to intense social impact [6], colluding robots are predicted to have a powerful influence over the humans interacting with them.

Taken together, these aspects of robotic collusion against a person imply that such interactions may have a profound negative influence. While we do not expect robots to be intentionally designed with colluding behavior (but see [21], [23], [25], [26], [33], [34]), it is possible that the designers of multi-robot environments would fail to consider the possibility that robots collaboration would be perceived as intentional collusion. This may be the case when robot designers and practitioners develop multi-robot environments that involve communication between non-humanoid robots that do not speak. If not carefully designed, the behavior of robots with limited communication modalities (e.g., robotic arms, robotic dogs, and simpler abstract robots) may be misinterpreted as a collaborative effort to collude against the person they interact with. This is a concern even when interacting with very simple non-humanoid robots, since despite their simplicity, they are known to be perceived as social entities leading to consistent social experiences [35]–[37]. It is therefore important to explore whether people may interpret the collaboration between simple robotic objects as intentional collusion. Moreover, it is possible that, due to participants' low sense of belonging and perception of the robots as violating expectations, the intensity of such effect would reach beyond the specific interaction and impact robots' perception in general.

In an effort to better understand the possible negative impact of interacting with two simple non-humanoid robots that also interact with each other, we designed a gambling game where a participant plays a computerized "Find The Lady" game and gambles on chocolate coins. Participants were asked to identify the location of a specific card among three faced-down cards after they were shuffled. The game interaction involved two simple non-humanoid robots that communicated via non-verbal gestures: (1) a "Dealer" robot that was presented as responsible for shuffling the cards at each gambling round and competed against the participant; (2) an "Advisor" robot that recommended an answer on every gambling round (see Figure 3A). The perceived collaboration between the robots against the participant was manipulated in the last round of gambling. In the *Collusion* condition, the "Advisor" robot gave a wrong recommendation that led to losing all previous winnings to the "Dealer" robot. In the *Non-Collusion* condition, the "Advisor" robot gave accurate recommendations throughout the interaction. We measured whether, in such a multi-robot context, participants would interpret the "Advisor" robot's wrong recommendation as representing the two robots' attempt to collude against the participant instead of as a simple mistake. We additionally evaluated the extent of such negative impact and tested if beyond its impact on participants' perception of the robots in the interaction it would also shape their negative attitudes toward robots in general.

II. Related Work

Related research includes interactions with more than one robot, ingroup-outgroup dynamics, and deception in HRI.

A. Interactions between humans and more than one robot

Several studies indicated that interactions with multiple robots can form a powerful social context that impacts humans' emotions and behavior [8], [38]–[44]. For example, Salomons and Scassellati [45] tested whether participants would change their opinions to align with those of the robots. Participants interacted with a group of three robots and were instructed to associate a word with one out of six potential images. Participants selected a preliminary response, followed by a display of the robots' responses, and were subsequently asked to choose a final answer. The robots' responses were either uniform or differed among themselves. It was found that when robots had uniform responses, participants modified their answers to conform to the robots' selections. Other studies tested people's perception of two robots based on the type of interaction between them. Tan et al. [11] tested if social communication between robots would enhance their perception. In the study, participants were asked to request assistance from a stationary robot that summoned a mobile robot to help them navigate. The study showed that when the interaction between the robots involved social cues, people perceived the robots as more likeable and warm. Interactions with more than one robot were also shown to have highly negative effects. For example, Erel et al. [7] demonstrated how an interaction with two non-humanoid robots could lead to the experience of ostracism. In a physical ball-tossing activity adapted from the Cyberball paradigm [46], participants interacted with the robots under three conditions: "Exclusion" - where the robots tossed the ball mostly between each other, "Over-inclusion" - where the robots tossed the ball mostly to the participants, and "Inclusion" - where participants tossed the ball equally to each other and the participant. Results indicated that the exclusion condition led to an ostracism experience and a decrease in psychological well-being. Expanding on this, Erel et al. [10] showed that robotic ostracism has a carryover effect, leading to an increase in compliance in a subsequent Human-Human interaction. The exclusion by the robots resulted in an effort to compensate for the threat to participants' need to belong, which led to higher compliance in subsequent interaction with another human.

While the studies demonstrated both positive and negative aspects of interactions with more than one robot, they did not explore the impact of perceiving the robots as colluding against the person interacting with them. We extend these studies and test if a single negative experience of perceiving robots as colluding against a human would lead to a negative perception of the robots and would extend to negative attitudes towards robots in general.

B. Ingroup-Outgroup Dynamics in HRI

Several studies indicated that ingroup and outgroup effects can impact interactions with robots [1], [14], [47]–[51]. For example, Fraune [51] tested how people perceive and morally behave toward robots and humans that either belonged to their group or to a competing group. Participants engaged in a price guessing competition against another group. The researchers tested participants' moral behavior towards humans and robots, which were either presented as playing with them or competing against them. Results showed that the participants favoured ingroup members over outgroup members and humans over robots. Yet, ingroup robots were treated more favourably than outgroup humans, emphasising that group membership had a greater impact than agent type. Ingroup-outgroup effects were also indicated in groups that involved two robots and one human. Häring et al. [50] investigated how social group membership influences human perceptions of robots. Participants played a card game with two robots, one that cooperated with them (ingroup) and the other that competed against them (outgroup). Results showed that participants rated ingroup robots as more competent, attributed more wisdom to them, and preferred interacting with them over outgroup robots. Previous studies have also indicated that different dynamics between humans and robots in a group activity can lead to ingroup-outgroup effects. Sebo et al. [52] tested the impact of assigning a specialized role of "communicator with a robot" to one of the group members. They tested groups that included three human participants and a robot and asked them to perform a collaborative task. When choosing one of the participants to be the only team member who can communicate with a robot, ingroup-outgroup effects emerged, making that person feel less included and making it less likely that their ideas would be incorporated into the group's final decision.

This line of studies indicates that formal roles and group assignments can lead to ingroup-outgroup effects. We extend it by indicating that robots' collaboration and its interpretation as related to deception can trigger a feeling of the robots teaming against the person, despite being simple non-humanoid robots.

C. Deception in Human-Robot Interaction

Studies have also explored deception in HRI [20]–[28]. While some of these studies indicated that deception can have negative effects like a decrease in trust [22], [27], others pointed towards some less negative influences, including increased engagement in the interactions [5], [23]. For example, Short et al. [5] tested participants' perception of a robot that cheats in a rock-paper-scissor game. Participants played 20 rounds against a humanoid robot that declared itself the winner on specific rounds despite losing. Their results indicated that participants attributed greater mental abilities to the robot that cheated and reported higher engagement compared to the interaction with the fair robot. Ullman et al. [53] also tested the perception of a cheating robot that presented either two humans or a human and a robot playing a competitive game. They introduced either honest behavior by the players or cheating behavior. Their findings showed that participants perceived the cheating robot (but not the cheating human) as more intelligent. They also perceived both the cheating robot and the cheating human as less trustworthy. Another line of studies explored if robots facilitate honest human behavior. For example, Hoffman et al. [54] explored if placing a robot next to a person who performs a task would decrease people's tendency to cheat. Participants performed a task with the opportunity to cheat for personal gain, either when being in the presence of another person, a robot, or alone. Results showed that both the robot and the human reduced the tendency to cheat. Interestingly, participants who cheated reported feeling less guilty in the robot's presence.

Taken together, these studies indicate different effects related to deception in HRI. We add to these effects the context of a multi-robot environment and the possibility that communication between two robots would be perceived as an attempt to deceive a person. We further evaluate whether such an impact would extend to the perception of robots in general.

III. Method

To evaluate whether an interaction with two simple non-humanoid robots would be interpreted as colluding against the human interacting with them, we designed a gambling task in which participants interacted with two robots. One robot was presented as the "Dealer", shuffling cards and competing against the participant. The other robot was presented as the "Advisor", offering recommendations on which card to choose. In the *Collusion* condition, the "Advisor" robot provided accurate recommendations to the participant during the first nine rounds but a wrong recommendation in the final double-or-nothing round, leading to the loss of all winnings. In

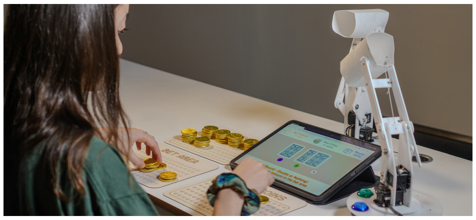


Fig. 2. The card game and a participant betting on chocolate coins.

the *Non-Collusion* condition, the “Advisor” robot consistently provided accurate recommendations throughout all ten rounds. Using this task, we could evaluate if participants would perceive the two simple non-humanoid robots as collaborating and colluding against them and the impact on their attitudes towards robots in general.

A. Gambling task and the robots

1) Gambling task: We used a computerized version of the “Find the Lady” gambling game. In a series of 10 rounds, participants were asked to identify a target card (the “Ace”) among three shuffled cards (an additional two “Kings”) presented on a tablet. Participants received two chocolate coins at the beginning of the experiment for their initial bets. In each round, they were asked to bet on one coin. If they guessed accurately, they won an additional coin, which they took from the “Bank Area” on the table (see Figure 1B). If their guess for the location of the “Ace” was wrong, they lost the coin they were betting on and placed it in the bank area. The participants were informed that they would keep any chocolate coins they won, adding a tangible incentive to the task.

At the beginning of each round, participants were instructed to place one of their chocolate coins in the “Bet Area”, indicating their bet on the outcome of that round (see Figure 2). During the round, the “Dealer” robot performed non-verbal gestures indicating that the cards on the screen should be shuffled, this triggered an animation of cards being shuffled on the tablet. The “Advisor” robot gave recommendations throughout the task, offering its opinion on where the target card was located. It indicated the location by pressing one of three buttons next to it, using a small plastic arm. Each of the buttons corresponded to one of the cards on the tablet (marked by the same number as the button).

Participants were asked to indicate their choice after the “Advisor” robot indicated its recommendation. However, they were explicitly informed that they were not obliged to take the “Advisor” robot’s recommendation and should choose according to their own judgment. After the participant chose a card, all three cards were flipped, revealing whether the participant won or lost. In most rounds (i.e., all apart from the 10th round in the *Collusion* condition), the winning card (the “Ace”) was in the position indicated by the “Advisor” robot. Thus, participants won the round if their choice corresponded with the “Advisor” robot’s recommendation. If the participants won, they were instructed to take their coin back from the “Bet Area” and take another coin from the “Bank Area”. If they lost, they were instructed to move their coin from the “Bet Area” to the “Bank Area”.

The game was custom-built using HTML, CSS, JavaScript, and Node.js. The game’s app was deployed on a server and communicated with the robots via a Node-RED plugin, utilizing the Butter Composer software from Butter Robotics [55] to create and manage the robots’ movements. The game consisted of two main components: the game interface for participants and an admin panel for the researcher.

2) The robots and gesture design: We used two simple non-humanoid robots resembling desk lamps [56], each designed for specific roles in the gambling task. The robots could communicate only via non-verbal cues. The “Advisor” robot was modified with three buttons and a small plastic arm that allowed it to indicate one of the three possible locations of the target card (the “Ace”) by pressing a button, simulating an advisory role in the task (see Figure 3). The robotic gestures required for the study were designed by an experienced animator and an HRI expert. 6 participants were recruited for the pilot study (3 women, 3 men; mean age = 35.4). The participants observed the robots’ gestures in a counterbalanced order and accurately described them in their own words, validating the gestures’ design. The final interaction involved the following gestures (see Table I for elaboration):

1) *Shuffling Gesture*: To indicate that the “Dealer” robot was responsible for shuffling the cards, it inclined forward towards the tablet and initiated a lateral rotation while the cards were shuffled.
2) *Recommendation Gesture*: To recommend a location, the “Advisor” robot leaned forward towards the tablet. Once the shuffling ended, it used its small plastic arm to press one of three buttons corresponding to the card it recommended and turned to face the participant.
3) *Feedback Gestures*:
   Feedback for a participant’s win: The “Advisor” robot performed an up-and-down motion, mimicking a cheerful *Winning gesture*. The “Dealer” robot lowered itself in disappointment, presenting a *Losing gesture*.
   Feedback for a participant’s loss: If the participant ignored the “Advisor” robot suggestion and lost, the “Advisor” robot performed a *Losing gesture*. The “Dealer” robot performed a *Winning gesture*.
4) *Final round gesture in the Collusion condition*: In the final round of the *Collusion* condition, both robots turned toward each other and performed the *Winning gesture*.

Fig. 3. The two simple non-humanoid robots.

B. Participants

48 participants were recruited for the study either from the university or via social media (31 women and 17 men; mean age = 23.25, SD = 2.86). All participants signed a consent form and received a 15 USD gift card or course credits. Additionally, all participants received 10 chocolate coins regardless of their performance in the gambling task.

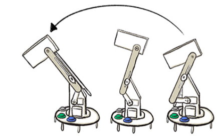
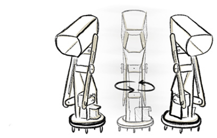
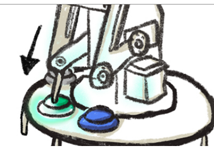
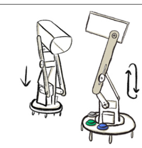
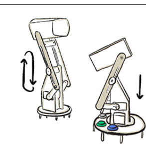
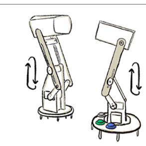
C. Experimental Design

In a between-subjects experimental design, we compared two conditions: (1) *Collusion* condition - the “Advisor” robot provided accurate recommendations in the first nine rounds but provided an incorrect recommendation in the last (10th) round that involved a bet of double-or-nothing, leading to the loss of all chocolate coins won until that round; (2) *Non-Collusion* condition - the “Advisor” robot provided accurate recommendations in all 10 rounds, leading to maximizing wins (i.e., chocolate coins). Participants were *randomly* assigned to one of the conditions using a matching technique that balanced gender and General Attitudes Towards Robots (GAToRS) [57].

D. Experimental Settings

The experiment was conducted in a quiet room, where the participant sat at a table with two robots, a tablet displaying the computerized gambling task, and a pile of chocolate coins in a designated “Bank Area”. The tablet was positioned in front of the participant, with the “Dealer” robot situated on the opposite side. The “Advisor” robot was positioned on the table to the front-right of the participant. The table was divided into three areas: “Your Coins” area was where the participants’ initial and won coins were placed; the “Bet Area” for placing bets (chocolate coins) in each round; and the “Bank Area” for the game’s chocolate coins (see Figure 1B).

TABLE I
THE ROBOTS’ GESTURES

Gesture Name	Description	Figure
Watch Gesture	“Advisor” Robot leans forward toward the tablet direction, indicating it is watching the tablet.	
Shuffling Gesture	The “Dealer” robot’s repetitive lateral rotation movement synchronized with the cards’ shuffling on the tablet.	
Recommendation Gesture	The “Advisor” robot recommends a card by pressing a button corresponding with the card’s location using a small plastic arm.	
Participant Win Gesture	Cheerful up-and-down movement by the “Advisor” Robot, the “Dealer” robot is down.	
Participant Lose Gesture	The “Advisor” robot is down, the “Dealer” Robot is cheerful.	
Final round Gesture (*Collusion-condition)*	The two robots perform a synchronized cheerful gesture.	

E. Dependent Measures

We used qualitative and quantitative measures to evaluate whether participants would perceive simple non-humanoid robots as collaborating against them and the impact on robot perception in general. Notably, we did not include a direct questionnaire for assessing collusion to avoid bias when evaluating the participants’ perception of the robots’ behavior, leaving room for the possibility that they would interpret the “Advisor” robot’s wrong recommendation as an error.

1) Qualitative Measure - semi-structured interview: A semi-structured interview was conducted to better understand the participants’ perceptions and reactions to the robots and the task [58]. The interview aimed to gather insights about the participants’ overall experience during the gambling task and their thoughts about the robots (e.g., “Describe your experience”, “What did you think of the robots?”).

2) Quantitative Measures: To evaluate if participants perceived the behavior of simple non-humanoid robots as collaborating against them, we applied two measures:

Interactive IOS Scale for Multiparty Interactions (IIMI) [59] - the IIMI scale assessed perceived closeness between the three parties of the interaction. This measure required participants to drag three circles presented on a screen into the arrangement that best represents their relationship with the robots and the relationship between the robots. The distance between the three circles represented the interpersonal closeness between the parties in the interaction (i.e., participant and the two robots). Therefore, robots that are perceived as colluding against the participant would be represented in circles positioned closer to each other.

Robotic Social Attributes Scale (RoSAS) [60] - A 9-point Likert scale for evaluating the perception of a robot’s warmth, competence, and discomfort. Since the “Dealer” robot was explicitly playing against the participant in both conditions, we focused on the perception of the “Advisor” robot.

To evaluate if the participants’ perception of the robots’ deception led to a broader effect and impacted their general perception of robots, we applied an additional scale:

Negative Attitudes Towards Robots Scale (NARS) [61]: A 5-point Likert scale evaluating broader perspectives of robots. It includes three sub-scales - Negative Attitudes toward Situations and Interactions with Robots; Negative Attitudes toward the Social Influence of Robots; and Negative Attitudes toward Emotions in Interaction with Robots.

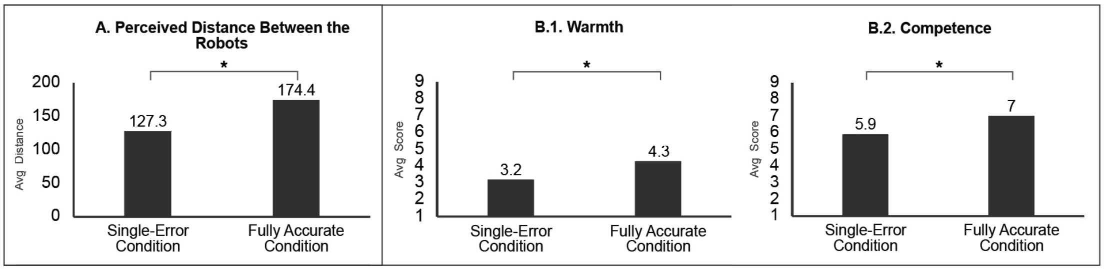


Fig. 4. The impact of the robots' behavior on the: A. IIMI distance between the two robots; B. the RoSAS warmth (1) and competence (2) ratings.

F. Procedure

A few days before the experiment, participants received an email with the general attitudes towards robots questionnaire (GAToRS; [57]) and a demographics questionnaire. Upon arrival, participants were given a brief overview and instructions for the gambling task. Participants were seated at the experiment table with the "Advisor" robot and the "Dealer" robot. The researcher explained that the "Advisor" robot would offer recommendations concerning the location of the target card but that the participants are free to follow or disregard the "Advisor" robot suggestion. The researcher left the room and activated the robots using a Wizard-of-Oz (WoZ) technique [62] (we used WoZ to minimize latency in the robot's responses). The robots performed their roles, with the "Advisor" robot providing correct recommendations in the first nine rounds. In the *Collusion* condition, the "Advisor" robot provided an incorrect recommendation at the final round, leading to a loss, while in the *Non-Collusion* condition, its recommendations were accurate at all rounds. Upon completing the gambling task, participants were asked to fill out the questionnaires and participate in the semi-structured interview. At the end of the experiment, the researcher debriefed the participants and verified that they had a positive experience.

IV. Analysis

To verify the lack of early differences between groups, we conducted Bayesian analyses for the pre-test (as it allows for testing NULL hypotheses). To evaluate the impact of the robots on perception and general attitudes towards robots we analyzed the IIMI, RoSAS, and NARS ratings using separated t-tests. To analyze the interviews, we performed thematic coding [58]: (1) Three coders transcribed and read all interviews. (2) The coders extracted initial themes. (3) The themes were discussed with another researcher.Inconsistencies were resolved. (4) Two coders analyzed part of the interviews and verified inter-rater reliability (kappa=0.89). (5) The coders analyzed the rest of the interviews.

V. Findings

The Bayesian analysis indicated no early differences between groups ($BF_{10} = 0.19$). The main analyses indicated the negative impact of the interaction with the robots. We note that only one participant in each condition tested the "Advisor" robot recommendation, yet no participants ran out of coins.

A. Interactive IOS Scale for Multiparty Interactions (IIMI)

The analysis of the IIMI distances revealed different perceptions of the relationship between the robots in the two conditions $t_{(46)} = 1.9$, $p = .05$ (see Figure 4A). In the *Collusion* condition, participants placed the robots in a closer distance compared to the *Non-Collusion* condition. This supports the possibility that participants in the *Collusion* condition perceived the two robots as collaborating in the game.

B. Robotic Social Attributes Scale (RoSAS) [60]

The RoSAS analysis revealed a significant difference between groups for the Warmth scale $t_{(46)} = 1.8$, $p = .05$ (see Figure 4B.1) and the competence scale $t_{(46)} = 2.23$, $p = 0.02$ (see Figure 4B.2). In the *Collusion* condition, participants rated the "Advisor" robot as less warm in comparison to the *Non-Collusion* condition and perceived it as having lower competence. This finding indicates that the robots' perceived collusion led to a negative perception of the "Advisor" robot.

C. Negative Attitudes Towards Robots Scale (NARS)

The analysis of the NARS ratings revealed that the robots' behaviour during the game also impacted participants' attitudes towards robots in general. In all three sub-scales participants indicated higher general negative attitudes towards robots in the *Collusion* condition in comparison to the *Non-Collusion* condition: Negative Attitudes toward Situations and Interactions with Robots ($t_{(46)} = 3.7, p < .001$; see Figure 5C.1); Negative Attitudes toward the Social Influence of Robots ($t_{(46)} = 3.6$, $p < .001$; see Figure 5C.2); and Negative Attitudes toward Emotions in Interaction with Robots ($t_{(46)} = 2.1$, $p = .03$, see Figure 5C.3). This indicates that perceiving the robots as colluding against the participant led to a negative perception of robots in general.

D. Qualitative Measure - semi-structured interview

The interviews' analysis revealed three themes: (1) Interpretation of the robots' intentions; (2) "Advisor" robot ingroup-outgroup affiliation; and (3) General attitudes towards robots

1) Interpretation of the robots' intentions: 18/48 participants discussed the robots' intentions. All of these participants were from the **Collusion** condition (18/24). Most participants explicitly stated that the "Advisor" robot was collaborating with the "Dealer" robot in an effort to deceive them (16/18). They reported a strong sense of betrayal: *"he kind of back-stabbed me.. it was their plan all along"* (P.3, F); *"He cheated,*

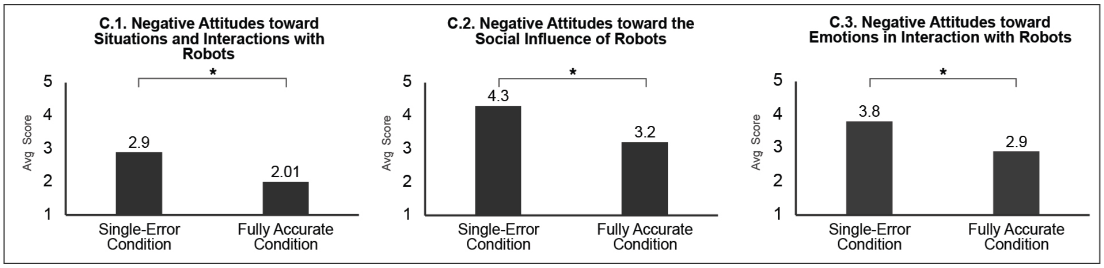


Fig. 5. The impact of the robots' behaviour on the three sub-scales of the NARS questionnaire.

*that son of a b*tch. I trusted him. I listened to him"* (P.47, F). Some participants considered the interaction between the robots as a collaborative scam: *"It's a scam. They, you know, scammed me at the end."* (P.23, M); *"I was betrayed. They set me up."* (P.17, F). Only 2/18 participants did not interpret the behavior of the "Advisor" robot as cheating and instead considered it as a simple mistake: *"It just made a mistake at the end."* (P.15, F). In the **Non-Collusion** condition, none of the participants discussed the robots' intentions.

2) "Advisor" robot ingroup-outgroup affiliation: 31/48 participants discussed the robots' affiliation. In the **Collusion** condition, 18/24 mentioned the perception of the "Advisor" as either on their side or the "Dealer" robot side. 14/18 participants felt that the "Advisor" robot with the "Dealer" robot and *against* them: *"I felt like they were scheming together against me"* (P.5, F); *"They were celebrating together, that I lost."* (P.33, F). Only 4/18 participants perceived the "Advisor" robot as being on their side and even as a companion: *"It was like I was not playing alone in the room."* (P.31, F). In the **Non-Collusion** condition, 13/24 discussed the "Advisor" robot's affiliation. The results pattern was opposite to that observed in the **Collusion** condition. 3/13 participants perceived the "Advisor" robot as playing against them: *"I felt that he was a little condescending like it was obvious that he is right."* (P.32, F). 10/13 participants reported feeling that the robot was trying to help them: *"The robot basically, like, helped me, it increased my chances to win."* (P.12, M). Some participants explicitly described the robot as becoming their team member: *"I felt that he is part of my team. It was me and the robot, so that's a memorable moment for me."* (P.42, F).

3) General attitudes towards robots: 42/48 participants talked about their general attitudes towards robots, either their current attitudes, or the ways they expect to behave around robots in the future. In the **Collusion** condition, 20/24 participants talked about their general attitudes towards robots. From them, 15/20 participants demonstrated negative general attitudes. Some participants showed concern about the future role of robots in significant tasks: *"If robots would start doing, like more serious tasks, I don't know if I would be okay with it."* (P.17, F); *"I feel like it's coming too soon [robots in our society], and the fact that it's like gaining knowledge... is freaking me out."* (P.7, F). Other participants reported being more cautious and suspicious in their next interactions with robots: *"I probably would trust robots less after this. I don't think I would trust them right off the bat."*, (P.3, F); *"Next time interacting with robots I'll keep in mind that I might get played."* (P.11, F). The other 5/20 participants in this condition indicated positive general attitudes towards robots: *"I think I will trust them more, stuff like army, natural disasters.. they could do great things that save time"*, (P.43, M). In the **Non-Collusion** condition, 22/24 participants talked about their general feelings and attitudes towards robots, exhibiting almost only (21/22) positive attitudes: *"I would integrate robots wherever they can help"* (P.36, M). Other participants specifically referred to how this specific interaction enhanced their general attitudes towards robots: *"Maybe I will trust them a little bit more. I will assume that they're capable."* (P.24, M). Only 1/22 participants in this condition reported negative general attitudes: *"It's better for robots to do like dangerous tasks rather than humans. but I also think that it will make us more lazy and less willing to do things."* (P.16, F).

VI. Discussion

In this study, we indicated the possibility that two simple robots would be perceived as colluding against a human. Participants interpreted the robots as interacting with one another and collaborating in an intentional attempt to deceive them. Both the IIMI results and the interviews indicated that the non-verbal gestures of the simple non-humanoid robots were interpreted as indicating the robots' *"plan"* " and *"scam"* intended to lead to the participant's loss in the game. This interpretation resulted in a negative perception of the robots (specifically their warmth and competence). Moreover, the participants' experience was so intense that it had a negative impact on their attitudes towards robots in general. Apart from the attitude ratings, in the interview, participants stated that in the future, they will *"not trust robots"* and will be *"more cautious"* when interacting with them. This intense negative impact highlights how subtle gestures performed by two non-humanoid robots can have profound effects that generalize beyond a specific interaction and influence (at least temporarily) attitudes towards robots in general.

Our findings suggest that people's tendency to assign complex social interpretations to interactions with robots should be carefully considered and addressed even when designing interactions with very simple robots. The robots used in this study were simple in form and limited in communication, using only simple non-verbal gestures throughout the entire interaction. Even so, participants perceived them as engaging in a complex social interaction that involved an intentional attempt to cheat and deceive them. In the interview, participants described explicit social intent attributed to the robots and stated that the "Advisor" robot *"kind of back-stabbed me"* and *"betrayed me"*. This finding is in line with previous studies, which indicated that people automatically perceive interactions with robots (even very simple ones) as social experiences due to human's inherent tendency to perceive the world through a social lens [35]. Therefore, we highlight the risk of dismissing simple robots as technologies that do not trigger social experiences, especially if more than one robot is involved. Overlooking the possible interpretation of their behavior as a colluding against a person may lead to intense negative effects that go beyond the specific interaction.

Our findings also suggest that perceived deception in a multi-robot interaction can lead to a negative experience of great intensity. It is possible that the negative effect of being deceived in a human-robot interaction was intensified by the multi-robot context, which positioned the participant as an outgroup member against which the robots collaborate. This possibility aligns with previous studies that indicated humans' sensitivity to perceived collaboration between robots, especially when this collaboration does not include them [2], [10]. Together with the negative effect of the robotic perceived collusion, a multi-robot interaction can become an extremely adverse experience. This explanation for the intensity of the experience is further supported by the RoSI model [6] that highlighted social belonging and violation of expectations as the main factors determining the intensity of robots' social impact. Deception by robots involves both factors. First, the robots' collaboration against the person positions the person as an outgroup member, decreasing their sense of belonging. Second, the cheating behavior violates the person's expectations as robots are expected to benefit humans and not deceive them [6]. According to the RoSI model, this type of interaction would result in a highly intense influence of the robots' behavior, which can explain participants' tendency to generalize their negativity from the specific interaction with the robots in our experiment to robots' perception in general.

The intense negative impact of the interaction with the robots may also explain people's ratings of the "Advisor" robot as having low competence. This finding is not in line with previous studies, which indicated that robots who cheat are perceived as intelligent and highly competent [53]. It is possible that in our study, the robotic collusion led to such a strong emotional response that resulted in an overall negative perception of the robots and an inability to acknowledge any positive aspects of their behavior. This idea is supported by participants' descriptions of the robots in the interview: *"It's worse than humans. Like I trusted it, I feel cheated."* (P.47, F), *"I was kind of betrayed, they set me up."* (P.17, F). The overwhelming sense of betrayal and being *'played'* by the robots may have overshadowed any potential recognition of the robots' intelligence or competence.

Taken together, our findings underscore the importance of carefully designing interactions with multiple robots, even when those involve very simple robots. We highlight the need to ensure that robots are not perceived as colluding against humans, even when the robots do not use verbal communication or complex non-verbal cues. If not well considered and controlled, the interaction between the robots can lead to extremely negative experiences where people feel betrayed and deceived. We suggest that the implications of such experiences may (temporarily) extend beyond the immediate interaction and negatively shape people's general attitudes toward robots. Designers of future public environments involving humans and robots should carefully consider the interaction between the robots in those environments. Whether designing interactions between delivery robots sharing our streets or robots in a store assisting human shoppers and organizing the store's products. If the communication between the robots is not designed carefully, people may perceive them as colluding against them to get priority over humans in busy streets or try to make people buy the shop's products even if they do not need them.

VII. Limitations

We used a specific non-humanoid robot. Future research should explore various robotic morphologies and robots that are not similar. Additionally, the artificial setup and specific context of a gambling task mediate our effects. As this is the first study focusing on the general impact of deception, we designed a clean lab study. Future studies should test different types of tasks in different contexts and meaningful consequences (i.e., risks or gains) to capture real-life complexity (preferably also with quantitative measures). Our finding may also reflect a temporary effect that should be further tested after participants are no longer emotionally aroused. Participants' capabilities may also mediate our effects. Future studies should evaluate the effect with different task complexities. Future work should also test the impact of ambiguous robotic behaviors and recommendation patterns (partial recommendations, number of wrong recommendations, and when they are presented). Lastly, interviews may be biased by the interviewers' expectations and the "good subject effect" [63].

VIII. Conclusion

This study highlights the importance of carefully designing interactions in multi-robot environments, even when the interaction involves very simple robots that do not use verbal language. Careful consideration of ingroup-outgroup effects and verifying that the robots' collaboration is not interpreted as collusion are essential for avoiding powerful negative effects that reach beyond the specific interaction with the robots.

Acknowledgement

We thank Ohad Shwartz, Hila Zohar, Ziv Keidar, May Bini, and Andrey Grishko for their invaluable help in this project.

REFERENCES

[1] A. Dahiya, A. M. Aroyo, K. Dautenhahn, and S. L. Smith, "A survey of multi-agent Human–Robot Interaction systems," *Robotics and Autonomous Systems*, vol. 161, p. 104335, Mar. 2023. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S092188902200224X

[2] H. Erel, Y. Cohen, K. Shafrir, S. D. Levy, I. D. Vidra, T. Shem Tov, and O. Zuckerman, "Excluded by robots: Can robot-robot-human interaction lead to ostracism?" in *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, 2021, pp. 312–321.

[3] E. Schneiders, E. Cheon, J. Kjeldskov, M. Rehm, and M. B. Skov, "Non-dyadic interaction: A literature review of 15 years of human-robot interaction conference publications," vol. 11, no. 2, pp. 1–32. [Online]. Available: https://dl.acm.org/doi/10.1145/3488242

[4] A. Litoiu, D. Ullman, J. Kim, and B. Scassellati, "Evidence that Robots Trigger a Cheating Detector in Humans," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. Portland Oregon USA: ACM, Mar. 2015, pp. 165–172. [Online]. Available: https://dl.acm.org/doi/10.1145/2696454.2696456

[5] E. Short, J. Hart, M. Vu, and B. Scassellati, "No fair!!: an interaction with a cheating robot."

[6] H. Erel, M. Vázquez, S. Sebo, N. Salomons, S. Gillet, and B. Scassellati, "RoSI: A model for predicting robot social influence," vol. 13, no. 2, pp. 1–22. [Online]. Available: https://dl.acm.org/doi/10.1145/3641515

[7] H. Erel, Y. Cohen, K. Shafrir, S. D. Levy, I. D. Vidra, T. Shem Tov, and O. Zuckerman, "Excluded by robots: Can robot-robot-human interaction lead to ostracism?" in *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, pp. 312–321. [Online]. Available: https://dl.acm.org/doi/10.1145/3434073.3444648

[8] M. R. Fraune, S. Sherrin, S. Sabanović, and E. R. Smith, "Rabble of robots effects: Number and type of robots modulates attitudes, emotions, and stereotypes," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. ACM, pp. 109–116. [Online]. Available: https://dl.acm.org/doi/10.1145/2696454.2696483

[9] M. Shiomi and N. Hagita, "Do synchronized multiple robots exert peer pressure?" in *Proceedings of the Fourth International Conference on Human Agent Interaction*, ser. HAI '16. Association for Computing Machinery, pp. 27–33. [Online]. Available: https://doi.org/10.1145/2974804.2974808

[10] H. Erel, E. Carsenti, and O. Zuckerman, "A Carryover Effect in HRI: Beyond Direct Social Effects in Human-Robot Interaction," in *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. Sapporo, Japan: IEEE, Mar. 2022, pp. 342–352. [Online]. Available: https://ieeexplore.ieee.org/document/9889554/

[11] X. Z. Tan, S. Reig, E. J. Carter, and A. Steinfeld, "From One to Another: How Robot-Robot Interaction Affects Users' Perceptions Following a Transition Between Robots," in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. Daegu, Korea (South): IEEE, Mar. 2019, pp. 114–122. [Online]. Available: https://ieeexplore.ieee.org/document/8673304/

[12] Merriam-Webster, "Collude." [Online]. Available: https://www.merriam-webster.com/dictionary/collude

[13] A. Rosero, "Using Justifications to Mitigate Loss in Human Trust when Robots Perform Norm - Violating and Deceptive Behaviors," in *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. Stockholm Sweden: ACM, Mar. 2023, pp. 766–768. [Online]. Available: https://dl.acm.org/doi/10.1145/3568294.3579979

[14] M. R. Fraune, S. Sabanovic, and E. R. Smith, "Teammates first: Favoring ingroup robots over outgroup humans," in *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. Lisbon: IEEE, Aug. 2017, pp. 1432–1437. [Online]. Available: http://ieeexplore.ieee.org/document/8172492/

[15] M. R. Fraune, S. Šabanović, and E. R. Smith, "Some are more equal than others: Ingroup robots gain some but not all benefits of team membership," *Interaction Studies. Social Behaviour and Communication in Biological and Artificial Systems*, vol. 21, no. 3, pp. 303–328, Dec. 2020. [Online]. Available: http://www.jbe-platform.com/content/journals/10.1075/is.18043.fra

[16] M. B. Brewer, "The Psychology of Prejudice: Ingroup Love and Outgroup Hate?" *Journal of Social Issues*, vol. 55, no. 3, pp. 429–444, Jan. 1999. [Online]. Available: https://spssi.onlinelibrary.wiley.com/doi/10.1111/0022-4537.00126

[17] S. Sebo, B. Stoll, B. Scassellati, and M. F. Jung, "Robots in groups and teams: A literature review," vol. 4, pp. 1–36. [Online]. Available: https://dl.acm.org/doi/10.1145/3415247

[18] J. F. Dovidio, S. L. Gaertner, and T. Saguy, "Another view of "we": Majority and minority group perspectives on a common ingroup identity," vol. 18, no. 1, pp. 296–330. [Online]. Available: http://www.tandfonline.com/doi/full/10.1080/10463280701726132

[19] J. C. Turner, R. J. Brown, and H. Tajfel, "Social comparison and group interest in ingroup favouritism," *European journal of social psychology*, vol. 9, no. 2, pp. 187–204, 1979.

[20] J. Danaher, "Robot Betrayal: a guide to the ethics of robotic deception," *Ethics and Information Technology*, vol. 22, no. 2, pp. 117–128, Jun. 2020. [Online]. Available: http://link.springer.com/10.1007/s10676-019-09520-3

[21] A. M. C. Isaac and W. Bridewell, *White Lies on Silver Tongues*. Oxford University Press, Oct. 2017, vol. 1. [Online]. Available: https://academic.oup.com/book/2320/chapter/142468698

[22] H. S. Sætra, "Social robot deception and the culture of trust," vol. 12, no. 1, pp. 276–286. [Online]. Available: https://www.degruyter.com/document/doi/10.1515/pjbr-2021-0021/html

[23] J. Shim and R. C. Arkin, "A Taxonomy of Robot Deception and Its Benefits in HRI," in *2013 IEEE International Conference on Systems, Man, and Cybernetics*. Manchester: IEEE, Oct. 2013, pp. 2328–2335. [Online]. Available: http://ieeexplore.ieee.org/document/6722151/

[24] K. Terada and A. Ito, *Can a robot deceive humans?*, journal Abbreviation: 5th ACM/IEEE International Conference on Human-Robot Interaction, HRI 2010 Pages: 192 Publication Title: 5th ACM/IEEE International Conference on Human-Robot Interaction, HRI 2010.

[25] A. R. Wagner and R. C. Arkin, "Robot deception: Recognizing when a robot should deceive," in *2009 IEEE International Symposium on Computational Intelligence in Robotics and Automation - (CIRA)*. Daejeon, Korea (South): IEEE, Dec. 2009, pp. 46–54. [Online]. Available: http://ieeexplore.ieee.org/document/5423160/

[26] ——, "Acting Deceptively: Providing Robots with the Capacity for Deception," *International Journal of Social Robotics*, vol. 3, no. 1, pp. 5–26, Jan. 2011. [Online]. Available: http://link.springer.com/10.1007/s12369-010-0073-8

[27] K. Winkle, P. Caleb-Solly, U. Leonards, A. Turton, and P. Bremner, "Assessing and Addressing Ethical Risk from Anthropomorphism and Deception in Socially Assistive Robots," in *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. Boulder CO USA: ACM, Mar. 2021, pp. 101–109. [Online]. Available: https://dl.acm.org/doi/10.1145/3434073.3444666

[28] A. Sharkey and N. Sharkey, "We need to talk about deception in social robotics!" *Ethics and Information Technology*, vol. 23, no. 3, pp. 309–316, Sep. 2021. [Online]. Available: https://link.springer.com/10.1007/s10676-020-09573-9

[29] A. Rosero, E. Dula, H. Kelly, B. F. Malle, and E. K. Phillips, "Human perceptions of social robot deception behaviors: an exploratory analysis," *Frontiers in robotics and AI*, vol. 11, p. 1409712, 2024.

[30] M. Vermue, R. Meleady, and C. R. Seger, "Member-to-member generalisation in trust behaviour: How do prior experiences inform prosocial behaviour towards novel ingroup and outgroup members?" *Current Psychology*, vol. 38, pp. 1003–1020, 2019.

[31] K. Haring, K. Nye, R. Darby, E. Phillips, E. de Visser, and C. Tossell, "I'm not playing anymore! a study comparing perceptions of robot and human cheating behavior," in *International Conference on Social Robotics*. Springer, 2019, pp. 410–419.

[32] Merriam-Webster, "Deception." [Online]. Available: https://www.merriam-webster.com/dictionary/deception

[33] J. Danaher, "Robot betrayal: a guide to the ethics of robotic deception," *Ethics and Information Technology*, vol. 22, no. 2, pp. 117–128, 2020.

[34] J. Shim and R. C. Arkin, "Robot deception and squirrel behavior: A case study in bio-inspired robotics," *Georgia Institute of Technology, Tech. Rep. ADA608845*, 2014.

[35] H. Erel, T. Shem Tov, Y. Kessler, and O. Zuckerman, "Robots are Always Social: Robotic Movements are Automatically Interpreted as Social Cues," in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. Glasgow Scotland Uk: ACM, May 2019, pp. 1–6. [Online]. Available: https://dl.acm.org/doi/10.1145/3290607.3312758

[36] H. Erel, G. Hoffman, and O. Zuckerman, "Interpreting non-anthropomorphic robots' social gestures," in *Proceedings of the 2018 ACM/IEEE international conference on Human-Robot Interaction, Chicago, Illinois. ACM*, 2018.

[37] L. Anderson-Bashan, B. Megidish, H. Erel, I. Wald, G. Hoffman, O. Zuckerman, and A. Grishko, "The greeting machine: an abstract robotic object for opening encounters," in *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2018, pp. 595–602.

[38] J. Kaduk, F. Weilbeer, and H. Hamann, "Emotional tandem robots: How different robot behaviors affect human perception while controlling a mobile robot." [Online]. Available: http://arxiv.org/abs/2403.03746

[39] A. Dahiya, A. M. Aroyo, K. Dautenhahn, and S. L. Smith, "A survey of multi-agent human-robot interaction systems." [Online]. Available: http://arxiv.org/abs/2212.05286

[40] T. V. Pham, T. H. Weisswange, and M. Hassenzahl, "Embodied mediation in group ideation – a gestural robot can facilitate consensus-building," in *Proceedings of the 2024 ACM Designing Interactive Systems Conference*, ser. DIS '24. Association for Computing Machinery, pp. 2611–2632. [Online]. Available: https://doi.org/10.1145/3643834.3660696

[41] A. Shamekhi and T. Bickmore, "A multimodal robot-driven meeting facilitation system for group decision-making sessions," in *2019 International Conference on Multimodal Interaction*, ser. ICMI '19. Association for Computing Machinery, pp. 279–290. [Online]. Available: https://doi.org/10.1145/3340555.3353756

[42] M. Shiomi and N. Hagita, "Do the number of robots and the participant's gender influence conformity effect from multiple robots?" vol. 33, no. 15, pp. 756–763. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/01691864.2019.1617781

[43] N. Salomons, M. Van Der Linden, S. Strohkorb Sebo, and B. Scassellati, "Humans conform to robots: Disambiguating trust, truth, and conformity," in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, pp. 187–195. [Online]. Available: https://dl.acm.org/doi/10.1145/3171221.3171282

[44] N. Salomons, S. S. Sebo, M. Qin, and B. Scassellati, "A minority of one against a majority of robots: Robots cause normative and informational conformity," vol. 10, no. 2, pp. 1–22. [Online]. Available: https://dl.acm.org/doi/10.1145/3442627

[45] N. Salomons and B. Scassellati, "Trust and conformity when interacting with a group of robots," in *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '18. Association for Computing Machinery, pp. 315–316. [Online]. Available: https://doi.org/10.1145/3173386.3176919

[46] K. D. Williams and B. Jarvis, "Cyberball: A program for use in research on interpersonal ostracism and acceptance," *Behavior research methods*, vol. 38, pp. 174–180, 2006.

[47] W.-L. Chang, J. P. White, J. Park, A. Holm, and S. Sabanovic, "The effect of group size on people's attitudes and cooperative behaviors toward robots in interactive gameplay," in *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. Paris, France: IEEE, Sep. 2012, pp. 845–850. [Online]. Available: http://ieeexplore.ieee.org/document/6343857/

[48] F. Eyssel and D. Kuchenbrandt, "Social categorization of social robots: Anthropomorphism as a function of robot group membership," *British Journal of Social Psychology*, vol. 51, no. 4, pp. 724–731, 2012.

[49] D. Kuchenbrandt, F. Eyssel, S. Bobinger, and M. Neufeld, "When a robot's group membership matters," vol. 5.

[50] M. Häring, D. Kuchenbrandt, and E. Andre, "Would you like to play with me? how robots' group membership and task features influence human-robot interaction," pp. 9–16.

[51] M. R. Fraune, "Our Robots, Our Team: Robot Anthropomorphism Moderates Group Effects in Human–Robot Teams," *Frontiers in Psychology*, vol. 11, p. 1275, Jul. 2020. [Online]. Available: https://www.frontiersin.org/article/10.3389/fpsyg.2020.01275/full

[52] S. Strohkorb Sebo, L. L. Dong, N. Chang, and B. Scassellati, "Strategies for the inclusion of human members within human-robot teams," in *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, 2020, pp. 309–317.

[53] D. Ullman, I. Leite, J. Phillips, J. Kim-Cohen, and B. Scassellati, "Smart human, smarter robot: How cheating affects perceptions of social agency."

[54] G. Hoffman, J. Forlizzi, S. Ayal, A. Steinfeld, J. Antanitis, G. Hochman, E. Hochendoner, and J. Finkenaur, "Robot presence and human honesty: Experimental evidence," in *proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction*, 2015, pp. 181–188.

[55] B. Megidish, O. Zuckerman, and G. Hoffman, "Animating mechanisms: A pipeline for authoring robot gestures," in *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, 2017, pp. 45–45.

[56] G. Hoffman, O. Zuckerman, G. Hirschberger, M. Luria, and T. Shani Sherman, "Design and Evaluation of a Peripheral Robotic Conversation Companion," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. Portland Oregon USA: ACM, Mar. 2015, pp. 3–10. [Online]. Available: https://dl.acm.org/doi/10.1145/2696454.2696495

[57] M. Koverola, A. Kunnari, J. Sundvall, and M. Laakasuo, "General Attitudes Towards Robots Scale (GAToRS): A New Instrument for Social Surveys," *International Journal of Social Robotics*, vol. 14, no. 7, pp. 1559–1581, Sep. 2022. [Online]. Available: https://link.springer.com/10.1007/s12369-022-00880-3

[58] R. E. Boyatzis, *Transforming qualitative information: Thematic analysis and code development*. Sage, 1998.

[59] A. W. Zhang, T.-H. Lin, X. Zhao, and S. Sebo, "Ice-Breaking Technology: Robots and Computers Can Foster Meaningful Connections between Strangers through In-Person Conversations," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. Hamburg Germany: ACM, Apr. 2023, pp. 1–14. [Online]. Available: https://dl.acm.org/doi/10.1145/3544548.3581135

[60] C. M. Carpinella, A. B. Wyman, M. A. Perez, and S. J. Stroessner, "The Robotic Social Attributes Scale (RoSAS): Development and Validation," in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. Vienna Austria: ACM, Mar. 2017, pp. 254–262. [Online]. Available: https://dl.acm.org/doi/10.1145/2909824.3020208

[61] D. S. Syrdal, K. Dautenhahn, K. L. Koay, and M. L. Walters, "The Negative Attitudes towards Robots Scale and Reactions to Robot Behaviour in a Live Human-Robot Interaction Study."

[62] L. D. Riek, "Wizard of oz studies in hri: a systematic review and new reporting guidelines," *Journal of Human-Robot Interaction*, vol. 1, no. 1, pp. 119–136, 2012.

[63] R. Opdenakker *et al.*, "Advantages and disadvantages of four interview techniques in qualitative research," in *Forum qualitative sozialforschung/forum: Qualitative social research*, vol. 7, no. 4, 2006.