



# A Study on Social Inclusion of Humanoid Robots: A Novel Embodied Adaptation of the Cyberball Paradigm

Cecilia Roselli<sup>1</sup> · Serena Marchesi<sup>1</sup> · Nicola Severino Russi<sup>1</sup> · Davide De Tommaso<sup>1</sup> · Agnieszka Wykowska<sup>1</sup>

Accepted: 19 March 2024 / Published online: 5 April 2024  
© The Author(s) 2024

## Abstract

As social robots are being built with the aim of employing them in our social environments, it is crucial to understand whether we are inclined to include them in our social ingroups. Social inclusion might depend on various factors. To understand if people have the tendency to treat robots as their in-group members, we adapted a classical social psychology paradigm, namely the “Cyberball game”, to a 3-D experimental protocol involving an embodied humanoid robot. In our experiment, participants played the ball-tossing game with the iCub robot and another human confederate. In our version, the human confederate was instructed to exclude the robot from the game. This was done to investigate whether participants would re-include the robot in the game. In addition, we examined if acquired technical knowledge about robots would affect social inclusion. To this aim, participants performed the Cyberball twice, namely before and after a familiarization phase when they were provided with technical knowledge about the mechanics and software related to the functionality of the robot. Results showed that participants socially re-included the robot during the task, equally before and after the familiarization session. The familiarization phase did not affect the frequency of social inclusion, suggesting that humans tend to socially include robots, independent of the knowledge they have about their inner functioning.

**Keywords** Social inclusion · Cyberball · Human-robot interaction · Intentional stance · Embodiment

## 1 Introduction

In everyday life, people usually engage with others, generally expecting to be included in social interactions [1, 2]. Humans tend to establish and maintain social connections with others to promote both physical and psychological well-being [3]. Indeed, social connections provide numerous benefits: for instance, they facilitate a quick detection of potential threats [1, 4], access to important resources such as food and protection, and constitute a source of social support by one’s peers [5, 6].

In a nutshell, social connections are crucial for the survival of human species. Recognizing someone as part of their in-group [7] is a consequence of establishing a social

connection. It can be that also the reverse direction holds, namely, identifying someone as an in-group member facilitates social connections. Interestingly, the relationship between social inclusion (in-group membership) and social connection might apply not only to other humans, but also to artificial agents such as robots. Indeed, human users prefer to interact with robots that are categorized as in-group members [8], based on certain characteristics such as sex, physical appearance, or country of origin. Given the increasing presence of a variety of robots in human social spaces, such as workplaces, hospitals, or education settings [1], researchers have started addressing the question of what conditions might be relevant for the acceptance of robots in human social spaces; in other words, what factors might facilitate the perception of robots as in-group members, and thus their social inclusion. Being perceived as a social agent might be the consequence of activating similar socio-cognitive mechanisms in response to robots as those that would be activated in the presence of humans. Our brain has developed a plethora of such mechanisms, and they all serve specific functions that allow us to efficiently navigate in social environments. For example, attention orienting

---

Cecilia Roselli and Serena Marchesi contributed equally to this work.

---

✉ Agnieszka Wykowska  
Agnieszka.Wykowska@iit.it

<sup>1</sup> Social Cognition in Human-Robot Interaction, Italian Institute of Technology, 16152 Genoa, Italy

based on robots' gaze direction [9–11], or the mirror neuron system similarly activated by the sight of both human and robotic actions [12, 13], are mechanisms that enable proper functioning with others. When such mechanisms are activated also in response to robots, it might result in perceiving the robot as an in-group social companion (since the brain reacts to it in a similar manner as it does towards another human) [14].

A better comprehension of what factors are relevant for the perception of robots as in-group members might be beneficial to the field of social robotics, by demonstrating that social robots can in fact be included into human spaces as social peers- even if human users are aware that robots are just mechanical artifacts. In the present paper, we decided to focus on one specific factor that might play a role in the social inclusion of robots, namely, familiarity with their inner workings. Indeed, understanding how robots work and what are their internal mechanisms might affect treating them as in-group members, as it becomes clear to the users that robots are very different from us. On the other hand, it is not trivial to experimentally measure whether a person socially includes or excludes another person (or a robot). In order to address social inclusion experimentally, we adapted to a human-robot interaction setting an experimental protocol that has been widely used in social psychology to address social inclusion, namely the Cyberball ball-tossing game paradigm [15]. Our aim was to examine the propensity of human participants to include the humanoid robot iCub [16] as an in-group member, as a function of familiarity with its functionality. Participants' familiarity with robot's inner workings was operationalized as a short familiarization training phase, during which participants were provided with technical knowledge regarding iCub's inner workings and functionality. Interestingly, our results showed that a short familiarization with the robot's functionality does not significantly affect the frequency of its social inclusion; in other words, providing participants with technical knowledge regarding the robot's workings and functionality does not modulate their tendency to socially include the robot. Therefore, we concluded that humans naturally tend to socially include robots, independently of the knowledge they have about its

inner functioning. In our view, these results might extend current knowledge regarding the relationship between familiarity with robots and social inclusion, at the same time being beneficial for designing social robots able to account for humans' propensity and attitudes towards them.

That said, in the following section we are going to introduce prior evidence that is relevant for the purpose of the present study. Specifically, we are going to describe the Cyberball paradigm [15] in detail, since our experimental method also relies on that, and how it has been widely

employed to investigate social inclusion/exclusion. Furthermore, as our study aims to explore the relationship between familiarity with robots and social inclusion of robots, we are going to present a brief review of literature, highlighting why familiarity is a relevant factor for the perception of robots as in-group members.

## 2 Related Work

This section introduces prior work showing how the Cyberball paradigm has been designed and validated by social psychologists to address social inclusion/exclusion, and how it has been adapted from human-human interactions to social contexts in which humans engaged in an interaction with a robot.

### 2.1 Cyberball Paradigm: from Human-Human to Human-Robot Interaction

In the original version of the Cyberball paradigm, participants were recruited to log on to an online experiment, in which they played a virtual ball-tossing game with two other avatar players, depicted as animated icons on the screen [15]. Crucially, the program could vary the number of tosses received by the participant. Ostracized participants received the ball only twice at the beginning of the game, and thus obtained fewer ball tosses as compared to the other players. Conversely, included participants repeatedly received the ball. Although the experimenters presented the task to participants as merely a way to train their mental visualization skills, results showed that participants cared about the extent to which the other players included them in the game. Specifically, if participants were over-included (i.e., they received the ball in half of the total tosses), or included (i.e., they received the ball in one-third of the tosses), they reported more positive feelings- in terms of meaningful existence, perceived control, and self-esteem- than ostracized participants (i.e., the ones who received the ball only one-sixth of the total tosses) [15]. Since the publication of the original study, the Cyberball paradigm has been extensively used in many different experimental contexts, by varying either structural aspects of the task (i.e., number of tosses, number of players), or research questions behind- for example, how participants' gender modulates social inclusion or the impact of cultural differences. For instance, Bernstein and colleagues manipulated the ethnicity of confederates so that Caucasian American participants took part in the Cyberball game with either same-ethnicity (i.e., Caucasian American) or other-ethnicity confederates (i.e., African American). Results showed that participants rated the experience of being included more positively, as

well as the experience of being ostracized more painful, when carried out by in-group members- namely, same-ethnicity confederates [17].

Notably, some versions of the Cyberball have been employed also in the Human-Robot Interaction (HRI) field, to examine factors potentially modulating social inclusion of robots; again, cultural differences seem to play a role. For example, a recent study [18] tested participants of two different nationalities, namely Chinese and UK participants who were representative of a more collectivistic and a more individualistic country, respectively. Results showed that, when asked to perform the Cyberball game with the humanoid robot iCub, the more participants displayed a collectivistic stance, the more they tended to include the robot in the game. More recently, Rosenthal von der Pütten and Bock [19] explored the effect of ostracism on humans when they were excluded by robots in a verbal interaction. The authors reported that humans are oversensitive to ostracism, even when they are excluded by robotic agents.

Taken together, this evidence demonstrates that Cyberball is a robust paradigm, that can be successfully used to investigate social inclusion/exclusion- both in human-human and in human-robot interactions. However, it remains unclear what factors contribute to social inclusion of robots; in particular, individuals' degree of familiarity and technical knowledge regarding robots has not been systematically studied. Therefore, in the next section we will minutely discuss this concept from different points of view, showing how it is relevant to social inclusion of robots.

## 2.2 Familiarity with Robots

Past research in social psychology demonstrated that familiarity, in terms of repeated exposure to a stimulus, breeds liking, as it reduces prejudice towards others and enhances the probability of treating them as social peers [20–22]. It might be because familiarity reduces, over time, people's apprehension towards novel and potentially threatening stimuli, such as other (unknown) humans [23, 24]. By repeatedly interacting with novel stimuli, individuals become more familiar with them, since they gain more knowledge. This, in turn, helps individuals to understand that these stimuli are not inherently threatening; therefore, over time, people start to like them more [25, 26].

There is evidence showing that the same seems to occur when the novel stimuli are robots, as people report liking them more and being more well-disposed towards them after repeated interactions [27, 28]. However, in the context of social interactions with robots, the role of familiarity for social inclusion seems to be not as straightforward as for humans. For example, a recent study involving a group of clinicians who conducted a robot-assisted intervention was

the first study showing that the degree of previous experience seems to be critical for the social inclusion of robots [27]. However, results indicated that the more clinicians were experienced with robots, the less they were social including the robot [27]. These results highlighted a relationship between familiarity and social inclusion in the opposite direction as compared to what happens with humans, for whom familiarity leads to increase liking (and thus promoting social inclusion) [20–22]. More recent evidence suggested that social inclusion of robots mainly depends on the way a robot is presented, and to a lesser degree on previous experience with them [29]. Specifically, the authors asked two samples of participants varying in their degree of technical education to perform the Cyberball game twice, i.e., before and after the presentation of a short video in which the humanoid robot iCub was either depicted as a mechanical artifact or as an intentional agent. Results showed that participants tended to socially include the robot in the game only after being exposed to iCub depicted as a mechanical artifact, regardless of participants' type of technical, formal education.

In contrast, further evidence showed that social inclusion of a humanoid robot, measured by means of the Cyberball game, was modulated by its degree of human-likeness. Ciardo and colleagues asked participants to play a musical duet with a humanoid robot programmed to commit errors, either in a human-like or in a machine-like way. Results showed that the probability of tossing the ball to the robot (social inclusion) was higher for participants who played the duet with the human-like erring robot, as compared to those who interacted with the robot erring in a mechanical way [30].

Taken together, this evidence shows that individuals' familiarity with robots definitively plays a role in social inclusion of robots. At the same time, these results conflict with each other, leaving open the question of whether exposure and information about robots (and specifically their inner workings and functionality) increase or decrease the likelihood of including the robot in one's social sphere- which represents the main point of interest of this study.

In this context, two factors might potentially contribute: the physical presence of an embodied robot, and the attribution of intentional agency to the robot. Given the importance of both factors to the phenomenon of interest (social inclusion of robots), in the following sections they will be addressed separately.

*1) Physical presence of an embodied robot.* In the study of Ciardo and colleagues [27], the Cyberball game was designed with the use of static pictures of a robot avatar. However, it is reasonable to think that social and physical presence of another agent has an impact on our social cognition, and on the way we process information about

that agent [31–33]. Regarding the importance of physical presence of robots, recent evidence showed that people responded differently, in terms of both individual attitudes and task performance, when the robot was “co-present”, i.e., both physically embodied and physically present in the user’s environment, as compared to when it was only “tele-present”, i.e., physically embodied by displayed on a screen. Indeed, robot’s physical embodiment positively affected people’s responses towards robots only when that embodiment allowed the robot to share the same physical space with the human [34]. Therefore, perhaps the question of whether technical knowledge about the mechanical functionality of a robot increases or decreases its social inclusion should be answered by means of protocols in which the Cyberball interaction happens with a physically embodied robot rather than only displayed on a screen in the shape of an avatar. To the best of our knowledge, only one study recently implemented a “physical” version of a ball-tossing game based on the Cyberball paradigm [35]. In this study, the authors manipulated the robot’s tosses in such a way that participants could feel excluded by the game, included, or over-included by the robot. Results showed that being excluded by the robot led participants to an ostracism experience which involved feeling rejected, with an impact on psychological needs such as control, belonging, and meaningful existence. Conversely, being included or over-included led participants to an experience associated with social inclusion [35]. Although this study’s results showed that social inclusion/exclusion by robots impacts individuals’ psychological mechanisms, it did not investigate factors influencing this effect, which is the purpose of the present study.

In the next section, we will focus on the potential contribution of the second issue to the social inclusion of robots: the attribution of intentional agency to robots.

2) *Adoption of the Intentional Stance towards robots.* The second potential factor that might be at play in social inclusion is related to the concept of the Intentional Stance [36]. According to Daniel Dennett, we adopt the Intentional Stance towards others when we treat them as intentional agents; namely, when we explain and predict their behaviour with reference to their (true or assumed) mental states, such as beliefs or desires. In contrast, with mechanical artifacts such as machines, we most likely adopt the Design Stance, namely, we are more prone to explain and predict their behaviour with reference to how these artifacts have been designed to behave. However, previous research in the field of HRI has demonstrated that we adopt the Intentional Stance not only towards humans, but sometimes also towards robots- although the Design Stance would be more appropriate, given that robots are mechanical artifacts [37–39]. To allow studying this philosophical

concept experimentally, Marchesi and colleagues [37] created the InStance Test (IST) (see Fig. 1 below). The IST differentiates between people’s tendency to explain and predict the behaviour of a humanoid robot with reference to mental states (Intentional Stance) and their tendency to explain and predict the robot’s behaviour with reference to its mechanical functionality (Design Stance). The IST comprises 34 visual scenarios, each of them containing three images that portray the humanoid robot iCub [16, 40]. Each scenario is linked with two descriptions: one explains the robot’s actions using mechanistic terminology (mechanistic description), whereas the other characterizes the robot’s behaviour in terms of mental states, such as “iCub likes”, “iCub desires” (mentalistic description). In Marchesi et al. study [37], participants were instructed to move a cursor along a slider, positioning it towards the description that best corresponded to their interpretation of the presented scenario. Results indicated a slight overall propensity towards the mechanistic description at the group level. However, there was also some tendency to adopt the Intentional Stance. This implies that, according to the New Ontological Category (NOC) hypothesis [41], participants were not firmly committed to adopting the Design Stance, thus showing uncertainty regarding the optimal stance to explain the robot’s actions. In other words, some participants were more likely to adopt the Intentional Stance, whereas some others were more likely to adopt the Design Stance.

Further evidence demonstrated individuals’ spontaneous propensity to adopt the Intentional Stance towards robots, at both behavioral and electrophysiological levels [42]. Notably, the IST resulted to be sensitive to detecting changes in the likelihood of adopting the Intentional Stance as a function of experience that one might have with robots. In three experiments, participants were exposed to the iCub robot exhibiting either human-like behaviors or entirely “mechanistic” reactions to the surroundings. Overall, results showed that individuals’ likelihood of adopting the Intentional Stance increased in the human-like condition, while remaining steady in the mechanical condition. The authors concluded that even brief interactions with humanoid robots can modulate the likelihood of adopting the Intentional Stance, with factors such as the robot’s behaviour and social engagement playing a role in that modulation. Most importantly, the results of this study showed also that the IST is a sensitive tool to detect changes in the propensity of adopting the Intentional Stance. The sensitivity of the IST was further confirmed by other studies which used different types of interactions with the robot [43–46].

Importantly for this study, apart from robot behaviour and social engagement, another factor is also crucial for adopting the Intentional Stance, namely familiarity with the inner workings of the robot. Similarly to affecting social

inclusion, familiarity also affects the likelihood of adopting the Intentional Stance towards the robot [47]. In this study, the authors recruited two groups of participants: one with a background in robotics, while the other one with a background in psychotherapy. The assumption was that roboticists should be most familiar with the mechanical explanations of (human) behaviour and are not familiar with robots. The results showed that psychotherapists indeed displayed a significantly larger likelihood of adopting the Intentional Stance towards robots, relative to roboticists. This was also paralleled in differential neural activity measured with electroencephalography (EEG).

Therefore, this paper combines the three main themes described above: social inclusion, familiarity with robots, and Intentional Stance. In detail, the paper addresses the question of whether exposing participants to knowledge about the mechanical functionality of the robot, and thus triggering a higher likelihood of adopting the Design Stance towards the robot (rather than the Intentional Stance) affects social inclusion.

### 3 Aims

In the present study, we aimed to address the following research questions:

RQ1: would participants' acquired technical knowledge about inner functioning of the humanoid robot iCub affect their tendency to socially include it in the Cyberball game?

RQ2: would participants' acquired technical knowledge about inner functioning of the humanoid robot iCub affect their likelihood of adopting the Intentional Stance towards the robot?

In order to address these questions, we set out to design an embodied version of the Cyberball task, in which participants would be interacting with the physical humanoid robot iCub. We exposed participants to a familiarization phase where explanations were given regarding the way iCub functions mechanically. Finally, we also administered the IST test pre- and post- familiarization to address the question of whether familiarization affects the propensity to adopt the Intentional Stance. In the next section, we will describe the design in detail, together with corresponding hypotheses.

#### 3.1 Social Inclusion: Our Embodied Adaptation of the Cyberball Game

To operationalize individuals' tendency to socially include robots, we developed an "embodied" version of the Cyberball ball-tossing game [15], namely, a version of the game involving an embodied humanoid robot sharing the same

physical and social space with participants and another human player, with all the three players performing the game together in real-time.

It is important to highlight that, in our version of the Cyberball, the term "embodiment" refers to the nature of the agents involved in the game, namely, the human confederate and the humanoid robot iCub, and not to the nature of the task itself. Indeed, the ball that the three players exchanged during the game was a virtual object displayed on the screen, and not a physical ball. Although it would have been extremely interesting to employ a physical object, in such a way as to make the Cyberball game entirely "embodied", we decided against this option for two main reasons.

First, the physical constraints of the iCub robot do not allow performing a complex movement, such as grasping and handing over a ball (or another soft object), for many repetitions without the risk of breaking. Second, even though the robot would not run the risk of breaking, its movement would have been extremely different from the human's movement, for example in terms of preparation and execution time. This might affect the frequencies of passing the ball to iCub rather than the human confederate. In contrast, having a virtual object (i.e., the ball), whose presentation and timings can be rigorously controlled in the same manner across the three players (the participant, the other human confederate, and iCub) would limit the occurrence of potential confounds given by the nature of the movements performed by the different agents.

In our version of the task, participants played the game with two other players, namely the humanoid robot iCub and another human player (a confederate, unbeknownst to the participant). Participants were free to choose which player to toss the ball to. The iCub robot was programmed to equally alternate the ball between the two players, whereas the other player—the confederate—was instructed to exclude iCub from the game. Specifically, the confederate was instructed (and trained) to toss the ball to iCub twice at the beginning of the game (first two tosses), and then only eight more times distributed along the whole game, which amounted to 10% of total tosses performed by the confederate. This procedure was implemented to give participants the impression that the robot was ostracized by the other human, thus allowing us to examine whether participants would tend to re-include iCub in the game. Re-inclusion was operationalized as the number of tosses that a participant would perform towards iCub during the game.

In the next sections, we will describe separately the hypothesized link between individuals' degree of familiarization with robots and the two phenomena of interest (social inclusion of robots, and adoption of the Intentional Stance).



*RQ1. Familiarization with robot's mechanical functionality.*

To assess the potential effect of participants' knowledge about robots (which we believed would affect the likelihood of adopting the Intentional Stance), participants were asked to perform the Cyberball in two separate— but identical— sessions, namely before and after a familiarization phase where the experimenter would explain participants some of the functionalities of the iCub robot. Specifically, after the first session of the Cyberball (Pre-Session), participants were invited to leave the cabin where the game took place and take a seat outside, in a position where the robot would be visible. Then, the experimenter started giving participants some technical details about the iCub's functionalities. In detail, the experimenter would tell, as a cover story, that she had to turn off the robot to avoid overheating. As part of the cover story, the experimenter would start explaining step-by-step which modules need to be deactivated, and the corresponding function— e.g., the ones controlling facial expressions (*emotioninterface*), or iCub's gaze (*iKinGazectrl*; [48]). Subsequently, the experimenter turned on the robot again, and showed participants some functionalities of the Motor Graphical User Interface (GUI) of iCub. Specifically, participants were told that each joint of the robot could be controlled in terms of pitch, roll, and yaw to perform a specific movement. After the explanation, the experimenter showed it in real-time to participants, by changing the parameters controlling the shoulder joints of the robot and inviting them to observe the robot while performing the corresponding movement (i.e., raising the left arm).

In total, this familiarization phase lasted approximately ten minutes, and it was meant to provide participants with some technical knowledge about robot's functionalities. Then, participants performed again the Cyberball game ("Post" session).

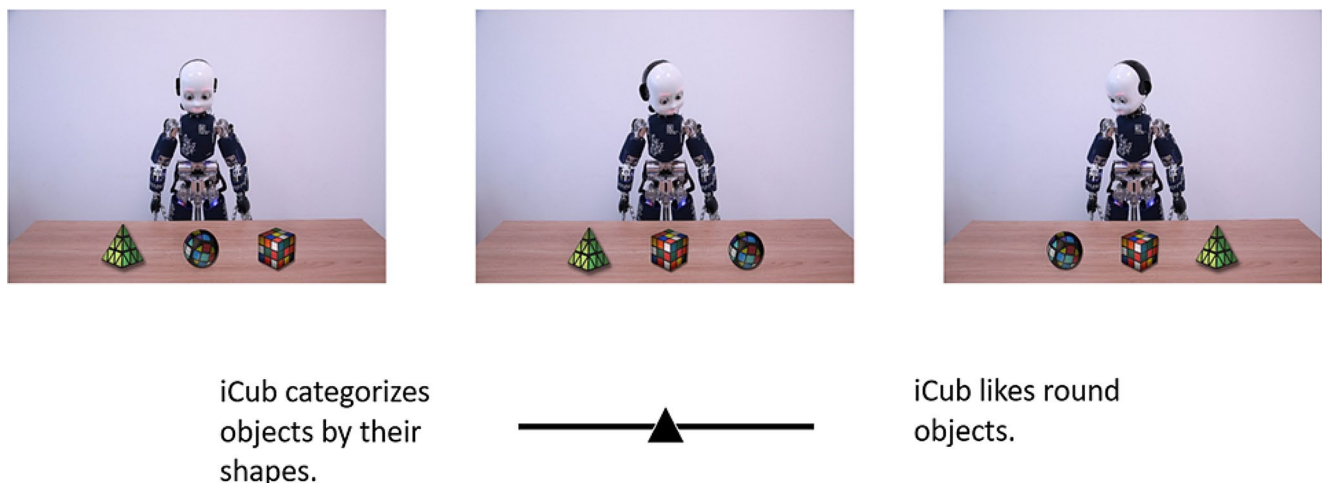
H1: We hypothesized that, if acquiring some technical knowledge in terms of robot's functionality affects the social inclusion of robots, then we should observe a difference in social inclusion, that is, the number of tosses towards the robots across the two sessions (i.e., Pre- vs. Post- sessions). Due to conflicting previous results, we did not have a hypothesis regarding which direction this modulation would occur.

*RQ2: The likelihood of adopting the Intentional Stance towards robots: the InStance Test.*

As a measure of participants' likelihood of adopting the Intentional Stance towards robots, we employed the Instance Test (IST) [37]. As described above, this test has been created to assess which strategy people tend to use when they explain robot's behaviors. It consists of 34 fictional scenarios depicting the iCub robot while performing various daily activities. Each scenario comprises three different pictures illustrating a sequence of events, with a scale (ranging from 0 to 100) providing a mechanistic description of the scenario on one extreme, and an intentional description on the other. By moving the cursor on a slider's scale towards one of the two extremes, in each scenario participants rate whether they think that iCub's behavior has a mentalistic explanation, such as desire or preference, or a mechanistic explanation, such as malfunctioning or calibration (see Fig. 1 for an example scenario).

We asked participants to fill out the IST in two separate sessions, namely before and after the familiarization phase about iCub's functionalities.

H2: we hypothesized that, if the degree of familiarity with robots affected the likelihood of adopting the Intentional Stance towards robots, the presentation of a robot as a mechanical device should reduce the likelihood of treating it as an intentional agent, which would result in reducing the IST Post-Session scores, relative to IST Pre-Session scores.



**Fig. 1** Screenshot of an example scenario taken for the Instance Test (IST) [37], with mechanistic explanation on the left, and mentalistic explanation on the right. Note that, in this study, the IST has been administered in Italian

## 4 Materials and Methods

### 4.1 Participants

Forty-six participants were recruited to participate in the study (Age range: 18–45 years old,  $M_{\text{Age}} = 25$ ,  $SD_{\text{Age}} = 6$ ; 21 males). All participants were right-handed, except for 5 left-handed and 1 ambidextrous. All participants had normal or corrected-to-normal vision, and no previous neurological disorders. The study was approved by the local ethical committee (Comitato Etico Regione Liguria) and conducted in accordance with the ethical standards laid down in the 2013 Declaration of Helsinki. Before the experiment, all participants gave written informed consent; at the end of the experiment, they were all debriefed about the purpose of the study. They all received an honorarium of 15 € for their participation. All participants were naïve to the purpose of the experiment and had no previous experience or training regarding the technicalities of the iCub robot.

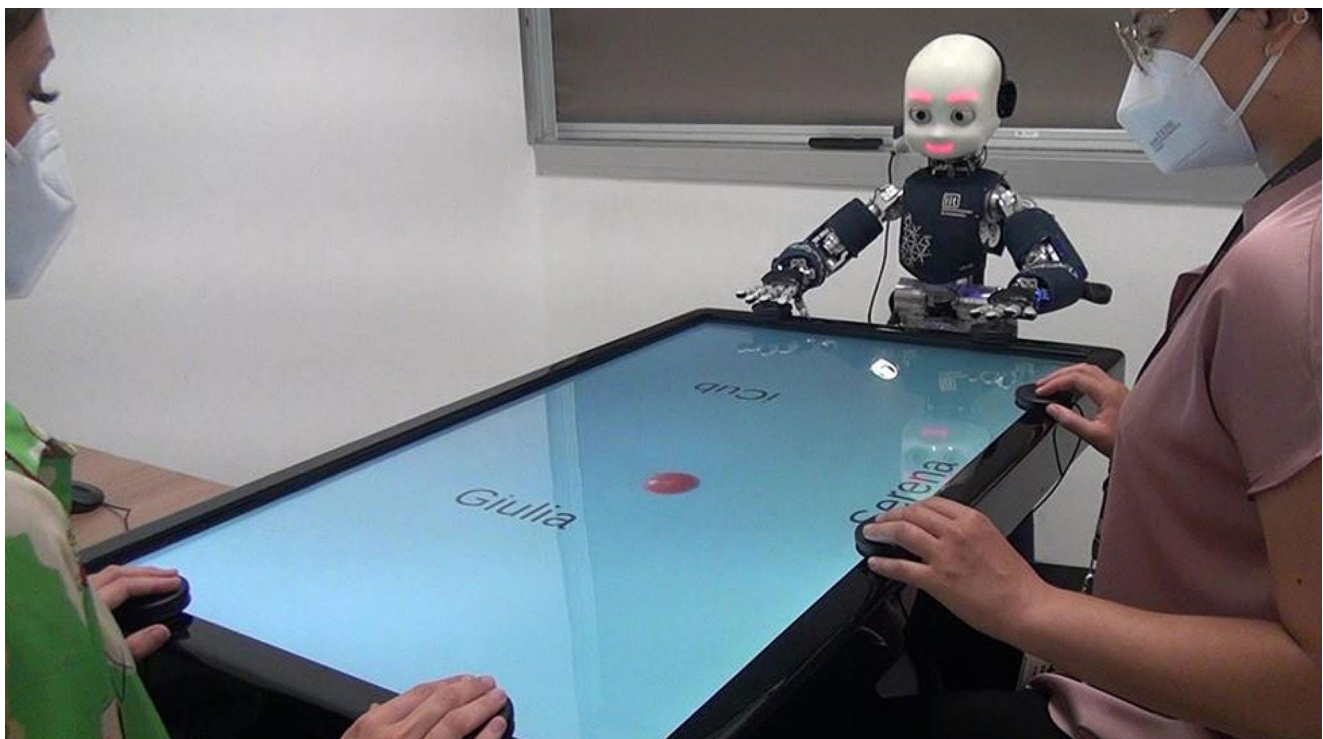
### 4.2 Apparatus and Stimuli

The experimental setup comprised the humanoid robot iCub, a workstation with a 21' inches screen to launch and control the Cyberball game, a big screen placed in a horizontal position to display the Cyberball (dimension =  $68 \times 155$  cm; resolution:  $1920 \times 1080$ ), three pairs of Logitech buttons, two

stools (one for participants, and one for the confederate), a set of speakers, and two laptops to present the InStance Test (IST) (see Fig. 2). Stimuli presentation and response collection of the Cyberball game were controlled using Psychopy v. 2021.2.3 [49]. The InStance Test was programmed using Psychopy v2020.1.3 [49].

### 4.3 The Robot Platform

The iCub is a humanoid robotic platform with 53 degrees of freedom (DoF). In our experiment, we designed the robot to verbally interact with participants in a short conversation, greeting them and inviting them to play the Cyberball game. During the game, the robot would be programmed to press buttons to toss the ball towards the participants or the confederate. The robot button pressing behaviors were implemented via the middleware Yet Another Robot Platform (YARP; [50]) using the position controller following a minimum jerk profile for head and hand movements to realize the keypresses. The greeting sentences at the beginning and the end of the experiment were played by the experimenter via a Wizard-of-Oz manipulation (WoOz; [51, 52]). The WoOz manipulation consists of an experimenter completely (or partially) remotely controlling a robot's actions (for a review, see [51]). In addition, we programmed the robot to look in the direction of participants' and confederate's eyes, since mutual gaze in human-robot interaction



**Fig. 2** Experimental setup of the Cyberball game. Before starting, the experimenter asked both participants and the human confederate for their names, which were then displayed on the screen in front of them for the entire duration of the game

has been shown to be a pivotal mechanism that influences human social cognition [9–11]. The gaze behaviour was implemented using the 6-DoF iKinGazeCtrl [48], using inverse kinematics to produce eye and neck trajectories. Facial expressions on the robot were controlled via the YARP emotion interface module.

#### 4.4 Experimental Procedure

At the beginning of the experiment, participants were asked to fill out the first half of the InStance Test (IST), and the confederate pretended to do the same to give participants the idea that they were both naïve participants. Subsequently, they were invited to enter the cabin where both the robot and the screen were placed, to perform the first session of the Cyberball (Pre-Session). Before the beginning of the game, the robot was programmed to greet participants and introduce itself (“Hello, I am iCub. Do we start playing?”). Then, the robot gazed at both players (i.e., the participant and the confederate), before lowering its head and gazing at the screen, being ready to play.

Each trial started with the presentation of both participants’ and the confederate’s names on the screen. The toss of the ball was represented by a one-second animation of a red ball, moving from one player to another for the entire duration of the game. As previously mentioned, iCub was programmed to equally alternate the ball between participants and the confederate. Conversely, the confederate was instructed to exclude the robot in the game, in such a way that she tossed the ball to iCub twice at the beginning of the game and then tossed the ball to the robot only eight times in total. These additional eight tosses were randomly distributed during the game, for a total of 10 ( $\pm 2$ ) tosses performed by the confederate across the whole duration of the game. This amounted to 10% of tosses overall.

Participants were always free to choose which player (i.e., iCub or the confederate) to toss the ball to. When participants (or the confederate) received the ball, before tossing it, they were instructed to wait until their name turned from black to red. After their name turned red, they had 500 ms to decide which player to toss the ball to, being as fast as possible. To choose the player on their right side, participants had to press the right button, whereas they had to press the left button to choose the player on their left side. To make sure that participants’ dominant hand did not bias their responses, we counterbalanced iCub’s position across participants, relative to the other player. Specifically, for the first half of participants, iCub was placed on participants’ right side, whereas for the other half, iCub was standing on their left side. If participants took more than 500 ms to press a button to choose the recipient of their toss, a red “TIME-OUT” statement appeared on the screen, and the trial was

considered invalid. The task comprised 100 trials in which participants received the ball, namely, participants had to choose to toss the ball to either of the players 100 times.

After performing the Pre-Session Cyberball, participants were invited to leave the cabin, and sit outside to see both iCub and the workstation controlling the robot. Then, participants were exposed to the familiarization phase about the robot’s functionality.

At the end of the familiarization phase, both participants and the confederate were asked to fill out the second half of the IST. Then, they were invited to re-enter the cabin, and perform again the Cyberball game (“Post” session). At the end of the experimental session, participants were asked whether they realized that the other human player was a confederate, or whether they believed that she was another naïve participant. Notably, all participants (except for one) reported that they did not understand that the other human was a confederate (see Fig. 3 for an overview of the various steps of the experimental session).

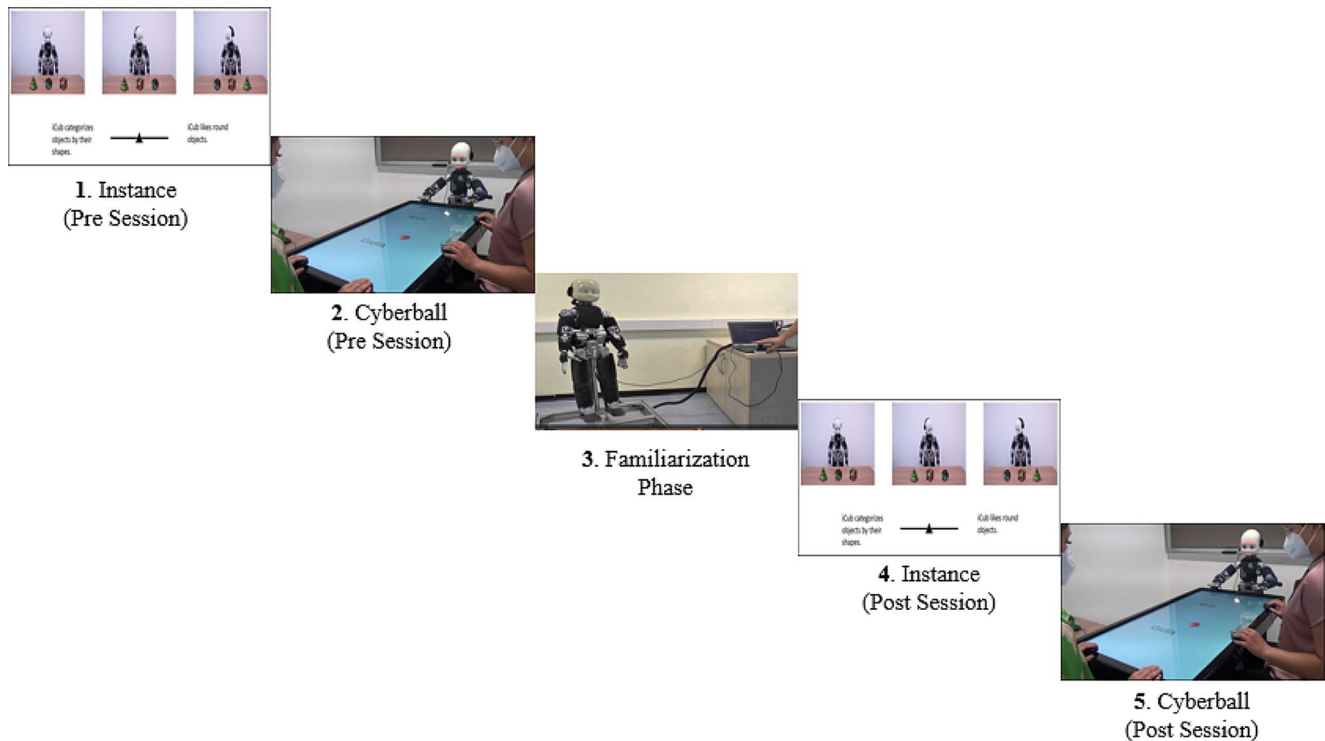
## 5 Statistical Analysis and Results

### 5.1 Data Preprocessing

Data from one participant were excluded because the confederate tossed the ball to the iCub robot more times than planned, namely, more than 10% of tosses overall (two tosses at the beginning of the game, plus other eight tosses randomly distributed during the rest of the game). It is important to note that participants’ removal upon confederate’s number of tosses is justified by the experimental design, according to which the confederate was instructed to pass the ball to iCub only 10% of tosses overall. Therefore, we decided to remove participants when the confederate tossed the ball more times than planned, as it potentially introduces a confound- i.e., participants might not perceive the robot as being ostracized by the confederate.

Moreover, data from another participant were excluded because the person realized that the other human player was a confederate instructed by the experimenters. This resulted in a final sample of  $N=44$ . Afterward, all trials deviating more than  $\pm 2.5$  SD from participants’ mean Reaction Times (RTs) were excluded from the subsequent analyses (2.52% of the total number of trials, Mean<sub>Pre\_Excluded</sub> = 86.95 ms, SD<sub>Pre\_Excluded</sub> = 104.73 ms; Mean<sub>Post\_Excluded</sub> = 88.73 ms, SD<sub>Post\_Excluded</sub> = 100.71 ms). We decided to apply an outlier exclusion method based on standard deviations because it introduces only small biases [53]. Specifically, the method based on the sample mean plus/minus a coefficient (usually between 2 and 3 times the standard deviation) is a common practice to detect and exclude outliers [54]. Following the





**Fig. 3** Experimental procedure. Participants started the experimental session by completing the first part of the IST (1). Then, they played the first session of the Cyberball game with iCub and the confederate (2), before being exposed to the familiarization phase about the robot's

functionality (3). Afterward, they were asked to fill out the second part of the IST (4), and they performed a second session of the Cyberball game (5)

same line of reasoning, participants with a low number of valid trials were excluded from further analyses, as keeping them would have introduced a larger bias in the distribution [53]. Therefore, data from other six participants were excluded due to a low number of remaining trials [i.e.,  $< 10$  in one or more combinations of Session (Pre- vs. Post-Session) and Recipient (Human confederate vs. Robot)] after outliers' removal, resulting in a final sample size of  $N=38$ . Analyses were conducted in JASP v.0.14.1 (2020), and in R studio v. 4.0.5 [55] using the *lme4* package [56]. Parameters estimated ( $\beta$ ) and their associated t-tests ( $t$ ,  $p$ -value) were calculated using the Satterthwaite approximation method for degrees of freedom; they were reported with the corresponding bootstrapped 95% confidence intervals [57].

## 5.2 Degree of Familiarity and Social Inclusion

To assess participants' tendency to re-include the robot in the Cyberball game, and how it was modulated by their acquired knowledge about robot's functionalities, we ran a linear mixed-effects model. Before running the model, we checked whether the assumptions of the model (linear relationship between the explanatory variables and the response variable, homogeneity of variances, and normality

of residuals) were met. Since all the assumptions were met, we built the following model (see Supplementary Materials, point SM.1.1, for more details regarding the assumption checks).

The frequency of participants' tosses was considered as the dependent variable and modelled as a function of two fixed factors: Session (Cyberball Pre vs. Post- familiarization phase), and Recipient, namely, which player participants decided to toss the ball to (the Human confederate vs. iCub), plus their interactions. Participants were considered as the random effect.

Results showed a significant main effect of Recipient [ $\beta=19.57$ ,  $SE=2.07$ ,  $t=9.45$ ,  $p<0.0001$ , 95% CI = (15.53, 23.62)], with an overall higher frequency of participants' tosses towards iCub compared to the human player [Robot: 3811 tosses (61%); Human: 2436 tosses (39%)] (see Fig. 4). No other main effect or interaction resulted to be significant (all  $ps > 0.31$ ).

## 5.3 Participants' Performance in Re-Including the Robot between Pre- and Post- Sessions

Additionally, we explored participants' performance (operationalized as their Reaction Times, RTs) when reincluding the robot among the two sessions. To do so, we focused only

on trials in which participants re-included the robot, and we built a linear mixed-effects model. Before performing the model, we checked whether the assumptions of the model (linear relationship between the explanatory variables and the response variable, homogeneity of variances, and normality of residuals) were met. Since all the assumptions were met, we built the following model (see Supplementary Materials, point SM.1.2, for more details regarding the assumption checks).

Specifically, participants' RTs were considered as the dependent variable and modelled as a function of Session (Pre- vs. Post familiarization session) as fixed effect. Participants were considered as the random effect. Results showed a significant main effect of Session [ $\beta = -20.23$ ,  $SE = 1.95$ ,  $t = -10.39$ ,  $P = < 0.0001$ , 95% CI = (-24.04, -16.31)]. Contrasting the marginal means showed that participants were significantly faster in reincluding the robot during the Post-familiarization session, as compared to the Pre-session [ $b = 20.2$ ,  $z = 10.38$ ,  $p = < 0.0001$ ;  $M_{\text{Pre-Session}} = 166$  ms,  $M_{\text{Post-Session}} = 143$  ms] (see Fig. 5).

To be thorough, we performed the same model as above, but considering only trials in which participants did not re-include the robot (and thus the Recipient was the human confederate). Assumption checks of the model were first controlled, and they were met (see Supplementary Materials, point SM.1.3, for more details). Therefore, participants' RTs were still considered as the dependent variable and modelled as a function of Session (Pre- vs. Post- familiarization session) as fixed effects; participants were the random effect. Results showed a significant main effect of Session [ $\beta = -21.99$ ,  $SE = 2.4$ ,  $t = -9.15$ ,  $p = < 0.0001$ , 95% CI = (-26.71, -17.29)]. Contrasting the marginal means showed that participants were significantly faster

in the Post-session also when tossing the ball to the human confederate [ $b = 22$ ,  $t = 9.16$ ,  $p = < 0.0001$ ;  $M_{\text{Pre-Session}} = 166$  ms,  $M_{\text{Post-Session}} = 140.5$  ms].

## 5.4 Degree of Familiarity and Adoption of Intentional Stance towards Robots

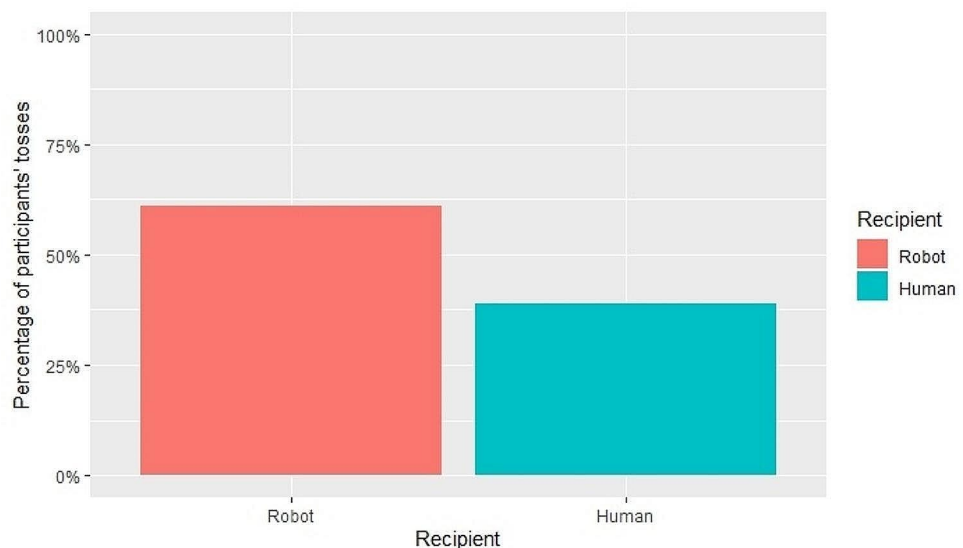
To assess whether participants' likelihood of adopting the Intentional Stance, operationalized as scores in the Instance Test (IST), was modulated by their acquired knowledge about robot's functionality and design, we first performed a Shapiro-Wilk test to assess whether data were normally distributed. Since it was not the case ( $p = 0.03$ ), we performed a non-parametric Wilcoxon signed-rank test to compare IST means across sessions (Pre- vs. Post-Session). Results did not show a significant difference in IST scores between the two sessions [ $W = 416.5$ ,  $r_b = 0.12$ ,  $p = 0.51$ , 95% CI  $r_b = (-0.24; 0.45)$ ;  $\text{Mean}_{\text{pre}} = 41.76$ ,  $\text{Mean}_{\text{post}} = 40.44$ ] (see Fig. 6).

## 5.5 Relationship between Individual Tendency towards Social Inclusion and Adoption of the Intentional Stance

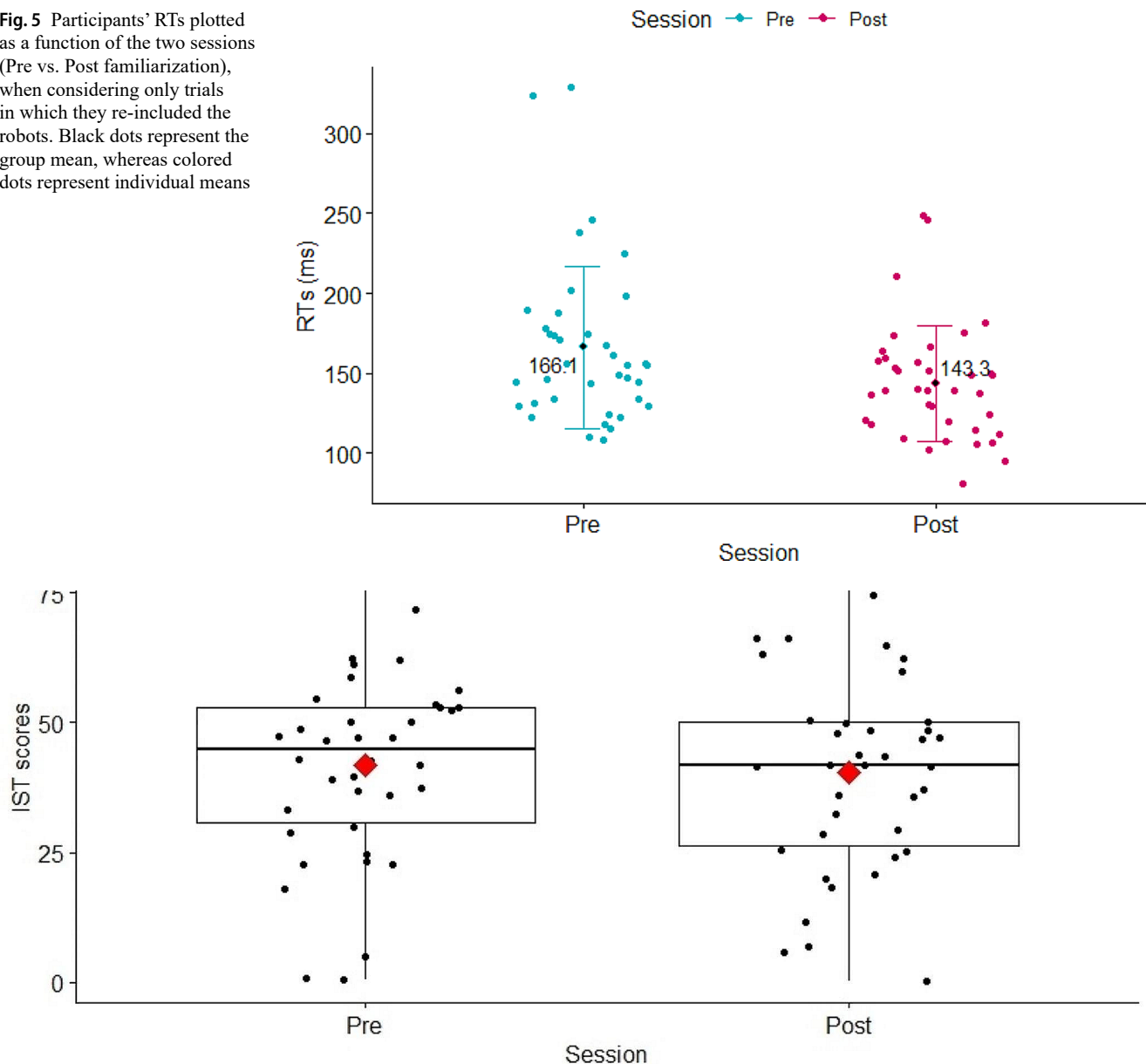
To assess whether there was a relationship between participants' individual tendency to socially re-include the robot in the Cyberball game, and their individual likelihood of adopting the Intentional Stance towards robots, we ran a correlation between the rate of robot choice and the Instance scores in the Pre-Session. Results did not show a significant correlation between the two variables ( $r_{36} = -0.08$ ,  $p = 0.7$ ).

In addition, we ran a correlation between the rate of choosing to toss the ball to the robot and the IST scores in the Post session, however, we did not find any significant

**Fig. 4** Frequency of participants' tosses, reported in the form of relative percentage over the total number of tosses, and plotted as a function of the Recipient (i.e., which player participants decided to toss the ball to: "Human" stands for the human confederate, whereas "Robot" stands for iCub)



**Fig. 5** Participants' RTs plotted as a function of the two sessions (Pre vs. Post familiarization), when considering only trials in which they re-included the robots. Black dots represent the group mean, whereas colored dots represent individual means



**Fig. 6** Participants' scores at the IST, plotted separately for each session ("Pre" stands for the IST session before the familiarization phase, whereas "Post" stands for the IST session after the familiarization

phase). The red diamonds represent the mean, whereas the black dots represent participants' individual means

result ( $r_{36} = -0.2$ ,  $p = 0.37$ ). The same occurred when running a correlation between the rate of choosing the robot in the Pre-Session and the delta, namely the difference between IST scores in the Post- and in the Pre-Session ( $r_{36} = 0.06$ ,  $p = 0.73$ ).

## 6 Discussion

In the present study, our aim was twofold. First, we wanted to investigate whether providing participants with technical knowledge about the inner functioning of the humanoid iCub robot modulates their tendency to perceive robots as social entities and thus treat them as in-group members. Second, we aimed to assess whether familiarity with robots modulates participants' likelihood of ascribing intentional traits to robots, i.e., adopting the Intentional Stance.

To measure participants' tendency to socially include the robot, we adapted a classical social psychology paradigm, i.e., the Cyberball game [15], to a setup involving a physically embodied robot. In our version of the Cyberball game, a human participant was instructed to toss a virtual ball (by means of two button presses) towards a fellow human player- i.e., the confederate- or a humanoid robot- i.e., the iCub robot. Using a physically embodied robot, which shared the same physical space with the other players, rather than only its picture avatar on the screen, we increased the realism and ecological validity of this paradigm. The game was divided into two sessions, each lasting until the participant reached 100 tosses.

To address RQ1, thus examining the potential impact of technical knowledge about robots on social inclusion, participants were exposed to a familiarization phase between the two Cyberball sessions. During the familiarization phase, the experimenter explained the role of each module controlling iCub, and how both the software and the mechanical components would interact to create the behaviors of the robot. This was done to give participants some technical knowledge about the robot as a mechanical system being controlled by hardware and software components, despite iCub being interactive with them, in a seemingly human-like way, during the Cyberball sessions. Our results showed a main effect of Recipient, meaning that our participants tossed more often the ball to the robot (i.e., including it more in the game) as compared to the human confederate, independent of familiarization with the robot's mechanical functions. This result is in contrast with previous evidence showing that the more people are experienced with robots, the less they were socially including the robot in the Cyberball game, and the more they were negative towards social interaction with robots [27]. Thus, our findings can be interpreted considering the behaviour of the human confederate, who adopted an exclusion strategy towards the robot in both sessions. Once participants realized that the fellow human was intentionally excluding the robot, they started re-including the robot in the game. One possible explanation may be found in the link between social inclusion and self-esteem. Since social inclusion is one of humans' primary needs, akin to other primary needs such as sustenance and shelter [44], providing social inclusion (and thus, behaving in a prosocial manner) towards others might increase one's own self-esteem and perception of reputation [58]. In the light of this, it is plausible that participants tended to socially include iCub, especially after seeing that the other human- i.e., the confederate- was excluding it. This would promote their own self-esteem, and through it, their well-being, and a positive representation of themselves. Under this interpretation, the robot was arguably seen as a social

agent calling for being socially excluded, despite the experimenter's explanation of the robot as a mere human-made artifact.

Regarding the fact that the familiarization phase did not alter the re-inclusion behavior of participants towards the robot, it might suggest that explaining to participants the robot's functionalities (from both software and mechanical point of view) was not enough to acquire a deep knowledge representation that would modulate their disposition to include the robot in as in-group member. Indeed, past research has shown that experts, namely people with an understanding derived from the accumulation of a large body of knowledge, represent this knowledge by applying deeper and more abstract principles of domain, as compared to novices who tend to have a shallower representation [59–61].

Interestingly, when we focused on the trials where the participants were re-including the robot (or not), the familiarization phase did affect the implicit behavior of our participants, showing that they were faster in the Post Session. This shows that it was easier for participants to re-include the robot after they have become familiar with its functionality. This might have been the result of practice during the task, meaning that in general our participants became faster in tossing the ball during the second Session, regardless of the agent they intended to toss the ball to.

To address RQ2, thus examining the potential impact of technical knowledge about robots on participants' likelihood of adopting the Intentional Stance, we administered the Instance Test (IST) in two sessions, namely, at the beginning of the experiment (before the first Cyberball session), i.e., Pre-Session, and right after the end of the familiarization phase, i.e., Post-Session. Our results showed no differences in IST scores in the Post-familiarization phase, relative to Pre-familiarization phase, thus suggesting that the short exposure to technological knowledge did not modulate the stance participants would take towards the robots. This was in contrast with our initial hypothesis, according to which the presentation of a robot as a mechanical device should reduce the likelihood of treating it as an intentional agent- which, in turn, would result in lower IST scores in the Post-Session relative to Pre-Session. It is also in contrast with previous evidence showing that the more participants were educated, in terms of accumulated knowledge regarding robots, the less they were willing to perceive robots as intentional agents [47, 62]. Also in this case, it might be that the familiarization phase was too short and too superficial, thus not providing a 'real robot experience' being able to modulate the initial stance that people took towards the iCub robot.

Beyond our main research questions, we decided to examine whether there was a relationship between participants'



tendency to socially include the robot, and their likelihood of adopting the Intentional Stance towards the robot. We correlated the number of tosses towards iCub in the first Cyberball session (Pre-Session, before the familiarization phase) at the individual subject level, with the individual scores at IST. Our results did not show any significant correlation between the two, suggesting that in our case, these two constructs are not associated with each other. Indeed, social inclusion and adoption of the Intentional Stance might be grounded in different socio-cognitive mechanisms. Future research should further explore the relationship between social inclusion and the adoption of the Intentional Stance.

In sum, the results of the present study show that when participants play a social inclusion game with a physically embodied robot, they tend to socially include the robot. This might reflect their need to increase self-esteem through a behaviour that is considered socially appropriate. Interestingly, this attitude is not modulated by acquired knowledge about the robot's mechanistic inner workings and functionality. This might suggest that individuals' initial attitudes towards robots are difficult to modify through short-duration instructions about the robot's functionalities.

## 7 Limitations and Future Directions

The current study presents two main limitations. The first limitation regards the familiarization phase, which indeed might have been too simple to modulate the inclusion/exclusion behaviour in the Session of the Cyberball administered after the familiarization. Similarly, the session might have been too short to modulate the adoption of the Intentional Stance. Future studies should address this point by creating different levels of complexity of the familiarization phase (i.e., simple introduction to the robot vs. the one presented here vs. a more technical hands-on training) and comparing them to directly investigate the effect of complexity of the familiarization on the inclusion/exclusion behaviour in the subsequent task. A second point that needs further investigation is participants' knowledge about technology. Future research might provide individuals with more extended knowledge about technology, to deepen their knowledge representations about mechanical functioning of social robots and thus modify their initial representation and social behaviors. Finally, in future studies it might be an interesting question to test whether the level of honorarium that participants receive (or are being told they would receive) influences their re-inclusion behaviour.

Another option might be to move the Cyberball paradigm to a “fully embodied” version, namely, involving not only an embodied robot physically sharing the same space as participants but also a physical object (e.g., a soft ball

instead of the ball projected on the screen). Once the technical challenges mentioned above regarding the robot itself are addressed (i.e., a better control algorithm), the option for a fully embodied Cyberball game should be systematically explored. Literature showed that touching/holding a soft object is effective for reducing negative emotions [63], for example social exclusion. Following this line of reasoning, it might be that using a physical, soft object during the Cyberball game might strengthen the probability of socially re-including the player that is ostracized, i.e., the iCub robot, in the game. To the best of our knowledge, no previous studies directly investigated the relationship between the use of a physical object- e.g., a soft toy- and the social inclusion in a Cyberball task. Only one study recently investigated the relationship between holding a soft object and the social pain induced by social exclusion, by means of the Cyberball game [64]. Surprisingly, the authors found that holding a soft object while performing the Cyberball task increased participants' subjective ratings of social inclusion. Finally, given that this setup would be more ecologically valid, since humans usually manipulate physical objects with embodied co-agents, it might potentially make easier to generalize further results to real-world scenarios.

## 8 Conclusions

In conclusion, this study aimed to explore the impact of exposure to technical knowledge about robots on (i) social inclusion of robots and (ii) the likelihood of adoption of the Intentional Stance towards such artificial agents. Our results showed that short familiarization with robots does not affect either social inclusion or the Intentional Stance. Importantly, however, the results showed that participants were likely to re-include the robot in the game after exclusion by the other participant.

This suggests that humans have the propensity to socially include humanoid robots as in-group members. Thus, including robots in human social environments might not be as difficult as it seems. Future research needs to examine whether results obtained in the lab generalize to real-world contexts.

**Acknowledgements** The authors thank Alessia Frulli for acting as a confederate during the experiment, and Federico Rospo for his help in data collection.

**Author contribution** CR designed the study; collected and analyzed the data; discussed and interpreted the results; and wrote the manuscript. SM designed the study; collected and analyzed the data; discussed and interpreted the results; and wrote the manuscript. NSR and DDT programmed the experiment; integrated the iCub robot into the experimental protocol; and wrote the manuscript. AW designed the study; discussed and interpreted the results; and wrote the manuscript. All authors revised the manuscript.

**Funding** Open access funding provided by Istituto Italiano di Tecnologia within the CRUI-CARE Agreement.

This work has received support from the European Research Council under the European Union's Horizon 2020 research and innovation program, ERC Starting Grant, G.A. number: ERC-2016-StG-715058, awarded to AW. The content of this article is the sole responsibility of the authors. The European Commission or its services cannot be held responsible for any use that may be made of the information it contains.

**Data Availability** The dataset analyzed during the current study is available, together with videos of the experiment, at the following link: <https://osf.io/etk5m/>.

## Declarations

**Conflict of interest** The authors declare that this study has been conducted in the absence of any commercial or financial relationship that could be construed as a potential conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Kerr NL, Levine JM (2008) The detection of social exclusion: evolution and beyond. *Group Dyn Theory Res Pract* 12(1):39–52. <https://doi.org/10.1037/1089-2699.12.1.39>
- Wesselmann ED, Wirth JH, Pryor JB, Reeder GD, Williams KD (2013) When do we ostracize? *Soc. Psychol. Personal. Sci.* 4(1):108–115. <https://doi.org/10.1177/1948550612443386>
- Baumeister RF, Leary MR (1995) The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychol Bull* 117:497–529. <https://doi.org/10.1037/0033-2909.117.3.497>
- Pickett CL, Gardner WL (2005) The social monitoring system: enhanced sensitivity to social cues as an adaptive response to social exclusion, in *The social outcast: Ostracism, social exclusion, rejection, and bullying*, in K. D. Williams, J. P. Forgas, & W. von Hippel (Eds.), pp. 213–226
- Buss DM (1991) Do women have evolved mate preferences for men with resources? A reply to Smuts. *Ethol Sociobiol* 12(5):401–408. [https://doi.org/10.1016/0162-3095\(91\)90034-N](https://doi.org/10.1016/0162-3095(91)90034-N)
- Duncan LA, Park JH, Faulkner J, Schaller M, Neuberg SL, Kenrick DT (2007) Adaptive allocation of attention: effects of sex and sociosexuality on visual attention to attractive opposite-sex faces. *Evol. Hum. Behav.* 28(5):359–364, Sep. <https://doi.org/10.1016/j.evolhumbehav.2007.05.001>
- Castelli L, Tomelleri S, Zogmaister C (Jun. 2008) Implicit ingroup metafavoritism: subtle preference for ingroup members displaying ingroup bias. *Pers Soc Psychol Bull* 34(6):807–818. <https://doi.org/10.1177/0146167208315210>
- Eyssel F, Kuchenbrandt D (Dec. 2012) Social categorization of social robots: Anthropomorphism as a function of robot group membership. *Br J Soc Psychol* 51(4):724–731. <https://doi.org/10.1111/j.2044-8309.2011.02082.x>
- Kompatsiari K, Perez-Osorio J, Davide DT, Metta G, Wykowska A (2018) Neuroscientifically-Grounded Research for Improved Human-Robot Interaction, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Madrid, Spain, pp. 3403–3408, 2018. <https://doi.org/10.1109/IROS.2018.8594441>
- Kompatsiari K, Ciarlo F, Tikhonoff V, Metta G, Wykowska A (2018) On the role of eye contact in gaze cueing. *Sci Rep* 8(1). <https://doi.org/10.1038/s41598-018-36136-2>
- Kompatsiari K, Ciarlo F, Wykowska A (2022) To follow or not to follow your gaze: the interplay between strategic control and the eye contact effect on gaze-induced attention orienting. *J Exp Psychol Gen* 151(1):121–136. <https://doi.org/10.1037/xge0001074>
- Oberman LM, McCleery JP, Ramachandran VS, Pineda JA (2007) EEG evidence for mirror neuron activity during the observation of human and robot actions: toward an analysis of the human qualities of interactive robots. *Neurocomputing Int J* 70:13–15. <https://doi.org/10.1016/j.neucom.2006.02.024>
- Gazzola V, Rizzolatti G, Wicker B, Keysers C (2007) The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. *NeuroImage* 35(4):1674–1684. <https://doi.org/10.1016/j.neuroimage.2007.02.003>
- Wykowska A, Chellali R, Al-Amin MM, Müller HJ (2014) Implications of robot actions for human perception. How do we represent actions of the observed robots? *Int J Soc Robot* 6(3):357–366. <https://doi.org/10.1007/s12369-014-0239-x>
- Williams KD, Jarvis B (2006) Cyberball: A program for use in research on interpersonal ostracism and acceptance. *Behav. Res. Methods*, 38(1):174–180. <https://doi.org/10.3758/BF03192765>
- Metta G et al (2010) The iCub humanoid robot: an open-systems platform for research in cognitive development. *Neural Netw* 23:8–9. <https://doi.org/10.1016/j.neunet.2010.08.010>
- Bernstein MJ, Sacco DF, Young SG, Hugenberg K, Cook E (Aug. 2010) Being 'In' with the In-Crowd: the effects of social exclusion and inclusion are enhanced by the Perceived Essentialism of Ingroups and outgroups. *Pers Soc Psychol Bull* 36(8):999–1009. <https://doi.org/10.1177/0146167210376059>
- Marchesi S, Roselli C, Wykowska A (2021) Cultural values, but not nationality, predict social inclusion of Robots. In: Li H, Ge SS, Wu Y, Wykowska A, He H, Liu X, Li D, Perez-Osorio J (eds) *Social Robotics. Lecture Notes in Computer Science*, vol 13086. Springer International Publishing, 13086, Cham, pp 48–57. [https://doi.org/10.1007/978-3-030-90525-5\\_5](https://doi.org/10.1007/978-3-030-90525-5_5)
- Rosenthal-von Der A, Pütten, Bock N (2023) Seriously, what did one robot say to the other? Being left out from communication by robots causes feelings of social exclusion, *Hum.-Mach. Commun*, 6, pp. 117–134. <https://doi.org/10.30658/hmc.6.7>
- Mrkva K, Van Boven L (2020) Salience theory of mere exposure: relative exposure increases liking, extremity, and emotional intensity. *J Pers Soc Psychol* 118(6):1118–1145. <https://doi.org/10.1037/pspa0000184>
- Zebrowitz LA, White B, Wieneke K (2008) Mere exposure and racial prejudice: exposure to other-race faces increases liking for strangers of that race. *Soc Cogn* 26(3):259–275. <https://doi.org/10.1521/soco.2008.26.3.259>
- Brewer M, Miller N (1988) Contact and cooperation. In: Katz PA, Taylor DA (eds) *Eliminating racism. Perspectives in social psychology (a Series of texts and monographs)*. Springer, Boston, MA. [https://doi.org/10.1007/978-1-4899-0818-6\\_16](https://doi.org/10.1007/978-1-4899-0818-6_16)
- Bornstein RF (1989) Exposure and affect: overview and meta-analysis of research, 1968–1987. *Psychol Bull* 106(2):265–289. <https://doi.org/10.1037/0033-2909.106.2.265>

24. Harrison AA (1997) Mere exposure. In advances in experimental social psychology. *Adv Exp Soc Psychol* 10:39–83. [https://doi.org/10.1016/S0065-2601\(08\)60354-8](https://doi.org/10.1016/S0065-2601(08)60354-8)
25. Zajonc RB (1968) Attitudinal effects of mere exposure. *J. Pers. Soc. Psychol.* 9(2):pt.2, pp. 1–27. <https://doi.org/10.1037/h0025848>
26. Montoya RM, Horton RS, Vevea JL, Citkowitz M, Lauber EA (2017) A re-examination of the mere exposure effect: the influence of repeated exposure on recognition, familiarity, and liking. *Psychol Bull* 143(5):459–498. <https://doi.org/10.1037/bul0000085>
27. Ciardo F, Ghiglini D, Roselli C, Wykowska A (2020) The effect of individual differences and repetitive interactions on explicit and implicit measures towards robots. In: Wagner AR et al (eds) *Social robotics. ICSR 2020: lecture notes in Computer Science*; 2020 Nov 14–18.; Golden, Colorado. Springer, Cham, pp 466–477. [https://doi.org/10.1007/978-3-030-62056-1\\_39](https://doi.org/10.1007/978-3-030-62056-1_39).
28. Heerink M (2011) Exploring the influence of age, gender, education and computer experience on robot acceptance by older adults, *Proceedings of the 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*; Mar 6–9 Lausanne, Switzerland, IEEE; 2011. <https://doi.org/10.1145/1957656.1957704>
29. Roselli C, Ciardo F, Wykowska A (2022) Social inclusion of robots depends on the way a robot is presented to observers. *Paladyn J Behav Robot* 13(1):56–66. <https://doi.org/10.1515/pjbr-2022-0003>
30. Ciardo F, De Tommaso D, Wykowska A (2022) Joint action with artificial agents: human-likeness in behaviour and morphology affects sensorimotor signaling and social inclusion. *Comput Hum Behav* 132:107237. <https://doi.org/10.1016/j.chb.2022.107237>
31. Schilbach L et al (2013) Toward a second-person neuroscience. *Behav Brain Sci* 36(4):393–414. <https://doi.org/10.1017/S0140525X12000660>
32. Redcay E, Schilbach L (2019) Using second-person neuroscience to elucidate the mechanisms of social interaction. *Nat Rev Neurosci* 20(8):495–505. <https://doi.org/10.1038/s41583-019-0179-4>
33. Bolis D, Schilbach L (2020) ‘I interact therefore I am’: The self as a historical product of dialectical attunement, *Topoi*, 39(3):521–534. <https://doi.org/10.1007/s11245-018-9574-0>
34. Li J (2015) The benefit of being physically present: a survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *Int J Hum Comput Stud* 77:23–37. <https://doi.org/10.1016/j.ijhcs.2015.01.001>
35. Erel H et al (2021) Excluded by robots: can robot-robot-human interaction lead to ostracism? in *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*, Boulder CO USA: ACM, pp. 312–321, 2021. <https://doi.org/10.1145/3434073.3444648>
36. Dennett DC (1971) Intentional systems. *J Philos* 68(4):87–106. <https://doi.org/10.2307/2025382>
37. Marchesi S, Ghiglini D, Ciardo F, Perez-Osorio J, Baykara E, Wykowska A (2019) Do we adopt the intentional stance toward humanoid robots? *Front Psychol* 10:450. <https://doi.org/10.3389/fpsyg.2019.00450>
38. Marchesi S, De Tommaso D, Perez-Osorio J, Wykowska A (2022) Belief in sharing the same phenomenological experience increases the likelihood of adopting the intentional stance toward a humanoid robot. 11. <https://doi.org/10.1037/tmb0000072>
39. Thellman S, de Graaf M, Ziemke T (2022) Mental state attribution to robots: A systematic review of conceptions, methods, and findings, *ACM Trans. Hum.-Robot Interact*, p. 3526112. <https://doi.org/10.1145/3526112>
40. Natale L, Bartolozzi C, Pucci D, Wykowska A, Metta G (2017) iCub: the not-yet-finished story of building a robot child. *Sci Robot* 2(13):eaq1026. <https://doi.org/10.1126/scirobotics.aq1026>
41. Kahn PH, Shen S (2017) Who’s there? A New Ontological Category (NOC) for Social Robots. In: Budwig N, Turiel E, Zelazo PD (eds) in *New perspectives on Human Development*, 1st edn. Cambridge University Press, pp 106–122. <https://doi.org/10.1017/CBO9781316282755.008>.
42. Bossi F, Willemse C, Cavazza J, Marchesi S, Murino V, Wykowska A (2020) The human brain reveals resting state activity patterns that are predictive of biases in attitudes toward robots. *Sci Robot* 5(46):eabb6652. <https://doi.org/10.1126/scirobotics.abb6652>
43. Ciardo F, De Tommaso D, Wykowska A (2021) Effects of erring behavior in a human-robot joint musical task on adopting Intentional Stance toward the iCub robot, in 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN), Vancouver, BC, Canada: IEEE, pp. 698–703, 2021. <https://doi.org/10.1109/RO-MAN50785.2021.9515434>
44. Marchesi S, Perez-Osorio J, Tommaso DD, Wykowska A (2020) Don’t overthink: fast decision making combined with behavior, in 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), pp. 54–59, 2020. <https://doi.org/10.1109/RO-MAN47096.2020.9223522>
45. Abubshait A, Wykowska A (Nov. 2020) Repetitive robot behavior impacts perception of intentionality and gaze-related attentional orienting. *Front Robot AI* 7:565825. <https://doi.org/10.3389/frobt.2020.565825>
46. Navare UP, Kompatsiari K, Ciardo F, Wykowska A (2022) Task sharing with the humanoid robot iCub increases the likelihood of adopting the intentional stance, in 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), Napoli, Italy: IEEE, pp. 135–140, 2022. <https://doi.org/10.1109/RO-MAN53752.2022.9900746>
47. Roselli C, Navare UP, Ciardo F, Wykowska A (2023) Type of education affects individuals’ adoption of intentional stance towards robots: an EEG study. *Int J Soc Rob* 16:185–196. <https://doi.org/10.1007/s12369-023-01073-2>
48. Roncone A, Pattacini U, Metta G, Natale L (2016) A cartesian 6-DoF gaze controller for humanoid robots, in *Robotics: Science and Systems XII, Robotics: Science and Systems Foundation*, <https://doi.org/10.15607/RSS.2016.XII.022>
49. Peirce J et al (2019) PsychoPy2: experiments in behavior made easy. *Behav Res Methods* 51(1):195–203. <https://doi.org/10.3758/s13428-018-01193-y>
50. Metta G, Fitzpatrick P, Natale L (2006) YARP: Yet Another Robot Platform. *Int. J. Adv. Robot. Syst*, 3(1):8. <https://doi.org/10.5772/5761>
51. Riek L (2012) Wizard of Oz studies in HRI: a systematic review and New Reporting guidelines. *J Hum Robot Interact* 119–136. <https://doi.org/10.5898/JHRI.1.1.Riek>
52. Kelley JF (1984) An iterative design methodology for user-friendly natural language office information applications. *ACM Trans Inf Syst* 2(1):26–41. <https://doi.org/10.1145/357417.357420>
53. Berger A, Kiefer M (2021) Comparison of different response time outlier exclusion methods: a simulation study. *Front Psychol* 12:675558. <https://doi.org/10.3389/fpsyg.2021.675558>
54. Leys C, Ley C, Klein O, Bernard P, Licata L (2013) Detecting outliers; do not use standard deviation around the mean, use absolute deviation around the median. *J Exp Soc Psychol* 49(4):764–766. <https://doi.org/10.1016/j.jesp.2013.03.013>
55. RStudio T, RStudio, Boston (2020) MA. <http://www.rstudio.com/>
56. Kuznetsova A, Brockhoff PB, Christensen RHB (2017) lmerTest Package: tests in Linear mixed effects models. *J Stat Softw* 82(13). <https://doi.org/10.18637/jss.v082.i13>
57. Efron B, Tibshirani RJ (1994) An introduction to the bootstrap. CRC
58. Berman JZ, Silver I (2022) Prosocial behavior and reputation: when does doing good lead to looking good? *Curr Opin Psychol* 43:102–107. <https://doi.org/10.1016/j.copsyc.2021.06.021>

59. Hmelo-Silver CE, Pfeffer MG (2004) Comparing expert and novice understanding of a complex system from the perspective of structures, behaviors, and functions. *Cogn Sci* 28(1):127–138. [https://doi.org/10.1207/s15516709cog2801\\_7](https://doi.org/10.1207/s15516709cog2801_7)
60. Means ML, Voss JF (1985) Star wars: a developmental study of expert and novice knowledge structures. *J Mem Lang* 24(6):746–757. [https://doi.org/10.1016/0749-596X\(85\)90057-9](https://doi.org/10.1016/0749-596X(85)90057-9)
61. Chi MTH, Feltovich PJ, Glaser R (1981) Categorization and representation of physics problems by experts and novices. *Cogn Sci* 5(2):121–152. [https://doi.org/10.1207/s15516709cog0502\\_2](https://doi.org/10.1207/s15516709cog0502_2)
62. Heerink M (2011) Exploring the influence of age, gender, education and computer experience on robot acceptance by older adults, in *Proceedings of the 6th international conference on Human-robot interaction*, Lausanne Switzerland: ACM, pp. 147–148. <https://doi.org/10.1145/1957656.1957704>
63. Tjewe-A-Sin M, Tops M, Heslenfeld DJ, Koole SL (2016) Effects of simulated interpersonal touch and trait intrinsic motivation on the error-related negativity. *Neurosci Lett* 617:134–138. <https://doi.org/10.1016/j.neulet.2016.01.044>
64. Ikeda T, Takeda Y (2019) Holding soft objects increases expectation and disappointment in the Cyberball task. *PLoS ONE* 14(4):e0215772. <https://doi.org/10.1371/journal.pone.0215772>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Cecilia Roselli** is currently a postdoctoral researcher in the Social Cognition in Human-Robot Interaction (S4HRI) unit, at the Italian Institute of Technology (Genoa, Italy). She received her Bachelor's (2015) and Master's (2017) degrees in Psychology at the University of Turin; then, she received her Ph.D. in Bioengineering and Robotics at the Università degli Studi di Genova in 2022. Her research interests primarily focus on social cognition. Her current research investigates the vicarious Sense of Agency phenomenon in a shared social context with various kinds of artificial agents, and how familiarity with robots impacts the likelihood of attributing intentionality to them.

**Serena Marchesi** is a postdoctoral researcher at the Social Cognition in Human-Robot Interaction (S4HRI) group, coordinated by Dr. Agnieszka Wykowska. Serena obtained her Ph.D. from the University of Manchester, in collaboration with IIT under the supervision of Prof. Angelo Cangelosi and Dr. Agnieszka Wykowska. Moreover, Serena won the A\*STAR Research Attachment Programme (ARAP) research award, which allowed her to spend more than a year at the Agency for Science, Technology and, Research (A\*STAR) in Singapore, under the supervision of Dr. Wu Yan. Previously, Serena obtained a second level Post Graduate Master in Clinical Neuroscience at the Catholic University of the Sacred Heart in Milan, under the supervision of Prof. Michela Balconi and a Master Degree in Cognitive Science and Decision-Making Processes at the Università Statale in Milan. Serena's research interests focus on the social and moral cognition processes involved in human-human and human-robot interaction, and how individual and cultural differences can affect them.

**Nicola Severino Russi** received his Master's degree in Industrial Engineering and Innovation Science from Università Politecnica delle Marche in 2019. Afterwards, he worked as an automation engineer in Biesse Group, before joining the Social Cognition in Human Robot Interaction (S4HRI) unit at the Italian Institute of Technology (Genoa, Italy) as a software engineer. He is currently working as an industrial engineer at the MISTER Smart Innovation in Bologna, Italy.

**Davide De Tommaso** works as research engineer in the Social Cognition in Human-Robot Interaction research group. He received his Master's degree in computer engineering from the University of Catania in 2010. He began his research activity at the Italian Institute of Technology in 2011 where he joined the Learning & Interaction group as a PhD student. In 2014 he obtained his PhD in Robotics, Cognition and Interactive Technologies. Today he is responsible for providing integrated solutions to support research activities in the design and development of experiments in human-robot interaction, in laboratory and clinical environments. In recent years he mainly deals with humanoid robots, EEG, eye-tracking and motion tracking systems, virtual and augmented reality.

**Agnieszka Wykowska** is the head of the unit "Social Cognition in Human-Robot Interaction" at the Italian Institute of Technology (IIT), in Genoa, Italy. She is also the Coordinator of the Center for Human Technologies, at the IIT. Her background is Cognitive Neuroscience (2006), Ludwig Maximilian University Munich (LMU) and philosophy (2001), Jagiellonian University in Krakow. She obtained a Ph.D. in psychology (2008) from the LMU. In 2016 she was awarded the ERC Starting grant "InStance: Intentional Stance for Social Attunement". She is Editor-in-Chief of *International Journal of Social Robotics*. Since July 2022 she serves in the role of President of the European Society for Cognitive and Affective Neuroscience (ESCAN). She is also board member of Association of ERC Grantees and a delegate to the European Research Area (ERA) Forum— an EU expert group shaping EU science policies. Her research foci are interdisciplinary, bridging psychology, cognitive neuroscience, robotics and healthcare. She combines cognitive neuroscience methods with human-robot interaction to understand the human brain mechanisms in interaction with other humans and with technology.