Joint action with artificial agents: human-likeness in behaviour and morphology affects sensorimotor signaling and social inclusion

Francesca Ciardo[1], Davide De Tommaso[1], & Agnieszka Wykowska[1]*

[1]Social Cognition in Human-Robot Interaction, Italian Institute of Technology, Genoa, 16152, Italy

*Corresponding Author:

Agnieszka Wykowska

Via Enrico Melen, 83

16152, Genoa, Italy.

Email: Agnieszka.Wykowska@iit.it

Competing Interest Statement: All the authors declare no competing interests.

Keywords: Human-Artificial agents interaction, sensorimotor communication, joint action, prosocial behaviour.

Abstract

Sensorimotor signaling is a key mechanism underlying coordination in humans. The increasing presence of artificial agents, including robots, in everyday contexts, will make joint action with them as common as a joint action with other humans. The present study investigates under which conditions sensorimotor signaling emerges when interacting with them. Human participants were asked to play a musical duet either with a humanoid robot or with an algorithm run on a computer. The artificial agent was programmed to commit errors. Those were either human-like (simulating a memory error) or machine-like (a repetitive loop of back-and-forth taps). At the end of the task, we tested the social inclusion toward the artificial partner by using a ball-tossing game. Our results showed that when interacting with the robot, participants showed lower variability in their performance when the error was human-like, relative to a mechanical failure. When the partner was an algorithm, the pattern was reversed. Social inclusion was affected by human-likeness only when the partner was a robot. Taken together, our findings showed that coordination with artificial agents, as well as social inclusion, are influenced by how human-like the agent appears, both in terms of morphological traits and in terms of behaviour.

Introduction

We coordinate activities with others on a daily basis – from paddling together in a kayak to playing in music ensembles. Successful coordination is achieved thanks to a complex plethora of cognitive mechanisms that allow us to continuously exchange implicit (nonverbal) signals with our partners. For instance, if you are kayaking in double with a friend that is less trained than you and s/he is not able to keep up with your tempo, you will notice it, even if s/he does not tell you. To ensure the achievement of your goal, namely, keeping the kayak moving, you will probably change your tempo to adapt to theirs. In doing so, your friend might interpret your change in behaviour as intention to help them (McEllin et al., 2018; Vesper et al., 2011).

Nonverbal signaling in joint action task

When performing actions with others, in the so-called joint-action scenario, we continuously use non-verbal signals, such as gaze direction, bodily posture, movement kinematics, reduction of behavioural variability, or prosody (see Vesper, et al., 2017 for a review). Such signaling strategy has been recently defined as action-based communication (Sebanz & Knoblich, 2021) or sensorimotor communication (Vesper & Sevdalis, 2020) and it mainly relies on body movements (e.g., Sartori, et al. 2009; Pezzulo, et al., 2013; Remland, et al. 1995). Sensorimotor communication serves three functions in social interactions: (a) informing the co-agent about our action intentions, (b) facilitating real-time coordination; (c) eliciting joint emotional and aesthetic experiences (see Vesper & Sevdalis, 2020 for a review). An example of using sensorimotor communication to highlight own action intentions is the exaggeration of action kinematics to reduce perceptual ambiguity and to facilitate action-recognition for the infant, i.e., motionese (Brand et al., 2002; Koterba & Iverson, 2009).

Reduction of behavioural variability as a form of nonverbal signaling

When the aim is to achieve real-time coordination, partners tend to reduce their behavioural variability. For instance, using a finger-tapping task, Vesper et al. (2011) showed that when pairs of participants were asked to synchronize discrete button presses in a reaction time task (Vesper et al., 2011), the variability in their responses was reduced compared to when they produced a simple reaction time task alone. Also, McEllin et al., (2018) showed that in a joint scenario in which pairs of participants were asked to play a virtual xylophone, participants modulated velocity parameters depending on whether their partner knew, or did not know, the action sequence to be performed. This suggests that when in a joint-action scenario, humans reduce response variability, with the intention to help the partner to “stay in the loop”, without committing errors in coordination (e.g., Vesper et al., 2011; McEllin et al., 2018; Sacheli et al., 2013). This is true for social interactions characterized by cooperative intent and goal-interdependency (Deutsch, 2011). On the contrary, when competing in a joint action task, individuals tend to intentionally modify their movements and behaviour (i.e., by being more variable or misleading) (e.g., Tomeo et al.; 2021) to be less informative with respect to predictions made by the competitor.

Adaptation to partner’s errors in the context of Dyadic Motor Plan

It is important to note that even if humans have the means to achieve coordination through the above-mentioned mechanisms, and even if our coordinated activities are trained over many hours, we are still prone to making mistakes. Only a few studies so far have investigated the impact of errors on joint action dynamics, mainly focusing on self- and-other error and action monitoring (e.g., Loehr et al., 2013; Moreau, 2021). A recent study by Sacheli and colleagues (Sacheli et al., 2021) showed that errors in joint action have an impact on sensorimotor signaling. Using a musical turn-taking task, the authors showed that violation of expectations driven by the partner’s error triggers an implicit tendency to correct the error, sacrificing individual efficiency in favor of sensorimotor signaling.

Here, we propose to consider adaptation to partner’s errors in joint action tasks within the embodied cognition framework (Loehr et al., 2013) and the Dyadic Motor Plan model (Sacheli et al., 2013; 2018). According to the embodied cognition framework, individuals activate their sensorimotor representation of an action when observing actions of others (Blakemore & Frith;

2005; Wilson & Knoblich, 2005; Grafton, 2009; Schubert & Semin; 2009). This action simulation or motor resonance has been indicated as the process that supports interpersonal coordination, as it allows to integrate self- and other-internal models within a Dyadic Motor Plan (DMP) (Prinz, 1990; Jeannerod, 2006; Vesper et al., 2010; Herwig, 2015). According to the DMP account, during interactive tasks, we represent our own and others' actions in terms of their contribution to the achievement of the ***joint*** goal (i.e., paddling so that the kayak goes straight). This allows us to represent and predict the effects of our own and our partner’s actions jointly. Thus, when partner’s action effects do not contribute to the achievement of the joint goal as predicted (for example, they are paddling at a different tempo or incorrectly), we select an appropriate response based on the effect this error produces with respect to the overarching joint goal (Sacheli et al., 2013; 2018). In effect, sensorimotor signaling emerges when a DMP is established. When a mismatch in the predicted behaviour of the partner and the observed one is detected, partners are ready to adapt consequently to achieve the joint goal.

Sensorimotor signaling as the basis for social inclusion of artificial agents

By ensuring coordination and achievement of joint goals, sensorimotor signaling indirectly impacts prosocial attitudes (Michael et al., 2020). For instance, it has been shown that subsequent to coordination tasks, individuals show higher cooperation and helping behaviour towards their partners (Kokal et al., 2011). Similarly, Hove and Risen (2009) found that the degree of synchrony between participants in a finger-tapping task correlates with subsequent affiliation ratings (Hove & Risen, 2009). It has been proposed that successful joint action increases social bonding and group membership by increasing the perceived similarity between co-agents. Thus, the increased trustworthiness and pro-sociality reported following interactive tasks that rely on coordination may be the result of the group-membership effect (Michael et al., 2020; Tajfel, 1970) which raises the expectation that in a future interaction in-group members will act toward in-group interest (Michael et al., 2016). Taken together, sensorimotor signaling seems to be a key mechanism of social cognition ensuring effective social interactions both directly and indirectly.

Joint action with artificial agents

At present, we interact not only with other humans but also with artificial agents that may (or not) be embodied. Soon, the application of artificial agents (including robots) within everyday contexts, such as workplaces, homes (Horvitz, 2016), and clinical settings (Ciardo & Wykowska, 2020) will make joint action with them as common as a joint action with other humans. For instance, robots will likely be involved in rescue operations during emergencies or could be taking the role of a partner in training our motoric skills in sports, maybe also training paddling skills. The interesting aspect of robots is that despite their artificial nature, they can induce in humans similar social cognitive mechanisms as those elicited by other humans during social interactions (Wiese et al., 2017; Wykowska, 2020; Ciardo et al., 2020; Hinz et al., 2021; Abubshait et al., 2020). This is particularly true for those robots that are designed to resemble humans in appearance, i.e., humanoids. A critical aspect during joint action with artificial agents is their morphology. Indeed, thanks to their embodied nature and their ability to move and act (potentially autonomously) within our environment, robots are artificial agents that can resemble humans not only in their physical appearance but also in their motor repertoire. Martini and colleagues investigated the relationship between human-likeness in morphology and robots' capability of inducing gaze following in humans. The authors showed that the degree to which humans follow the gaze direction of a robot does not linearly decrease with human-likeness in the morphology. Rather, the relationship between morphology and gaze following is best described by an inverted u-shaped pattern (Martini, Buzzell, and Wiese, 2015). In a recent study, Abubshait et al. (2020) tested the interplay between physical and behavioural human-likeness on joint attention. The authors showed that while physical appearance modulated joint attention only for lifelike interactions with the robot, behavioural features, such as the reliability of gaze signals, modulated joint attention across different types of interactions (lifelike vs. lab-based). Similarly, Ghiglino and colleagues (2020) showed that less attentional engagement is needed to process and interpret artificial agents' behaviour when it closely resembles one of a human being. Thus, the discussion about artificial agents' human-likeness should not be limited to the physical appearance per se, but it should be extended to their behaviour as well (Metta et al., 2008).

Aim of study

In the present study, we aimed at evaluating conditions under which an artificial agent (a humanoid robot or a computer program run on a standard PC) induces human sensorimotor signaling (***nonverbal***) during joint action. We focused on the ***reduction of variability*** as a measure of sensorimotor signaling in an interaction context, allowing for forming a joint goal, and thereby, for a ***dyadic motor plan*** to be established. Specifically, we were interested in cases in which the human should adapt to an ***erring behaviour of the artificial agent*** (thereby violating predictions formed through DMP), especially if the errors resemble human-like behaviour, in comparison to mechanical failures. In line with the above argumentation, we reasoned that sensorimotor signaling (and thus coordination) should also affect the ***social inclusion*** *of* the agent.

In short, we were interested in whether (i) humans would exhibit sensorimotor signaling towards an erring artificial agent in a joint action task, (ii) the signaling behaviour would be dependent on the human-likeness of the agent's erring behaviour, and (iii) the signaling behaviour would be related to a higher tendency to socially include the agent. In addition, we aimed at testing whether all these effects would depend on human-like appearance and motor repertoire of the artificial agent.

To address these questions, we designed an experimental paradigm in which human participants were asked to play a musical duet either with the iCub (Metta et al., 2008) humanoid robot (Experiment 1) or a computer algorithm (Experiment 2) that were programmed to make errors during their performance. We manipulated the human-likeness of the erring behaviour between-subjects in a way to reproduce a typical human mistake, i.e., a memory error, or a machine-like failure, i.e., entering a repetitive behavioural loop. After the joint task, we tested – by using a ball-tossing game inspired by the Cyberball paradigm (Williams & Jarvis, 2006; Ciardo et al., 2020) – whether our manipulation affected also the willingness to be socially inclusive towards the artificial agent. The Cyberball task has been extensively used in social psychology research to evaluate ostracism and social acceptance. In the original version of the study (Williams & Jarvis, 2006), participants are asked to virtually toss a ball towards two players, in a three-player game. One of the two players usually ostracizes the other one. The two players usually resemble in-group and out-group individuals. Group membership can be defined by any relevant characteristic shared (or not) between the participant and one of the players (e.g. race, sex, or social status). The typical result is that participants tend to toss the ball more often toward the ostracized player if s/he belongs to the ingroup instead of if s/he is an outgroup member.

We reasoned that the human-likeness of erring behaviour displayed by the artificial agent should affect sensorimotor signaling. Specifically, we predicted that when the erring behaviour belongs to a human repertoire (swapping by mistake an element of the sequence) and is also displayed by an embodied humanoid robot, the error itself does not compromise the DMP, as participants are still able to represent it in terms of its effects on the joint goal. As a consequence, their performance should display characteristics of sensorimotor signaling as a strategy to recover coordination and reach the joint goal. When the error is mechanical, i.e., the agent moves in an endless loop, the DMP should be compromised since participants are not able to represent the effects of the agent's action on DMP. In consequence, they cannot adapt their performance as much as in the case of human-like erring behaviour. Specifically, we predicted the effects of human-likeness on accuracy, performance in the task, and variability. We also predicted that these effects should be observed for the embodied robot, due to the possibility of representing its actions at the sensorimotor level. The effects were expected to be attenuated for the algorithm run on a standard PC, due to its non-human-like motor repertoire.

Regarding the impact of sensorimotor signaling during a joint task on willingness to include the agent as an in-group member we predicted that participants should prefer to socially include the agent that showed a human-like erring behaviour instead of a mechanical one. Specifically, we expected that the probability to toss the ball toward the iCub or the computer should be higher for those participants who experienced a human-like error instead of a mechanical one. Such a result would support the idea that the human-like erring behaviour increased the social inclusion and willingness to interact again with the agent. Also in this case, we predicted stronger differential effects across the erring conditions for the humanoid robot relative to the computer algorithm, due to the robot's human-likeness in appearance.

Materials and Methods

Participants

Seventy-three right-handed adults (27 males; mean age = 23.7 ± 3.8 years) took part in the study. Participants were recruited through the "Join the Science" mailing list (http://www.great-campus.it/join-the-science/). The experimenter sent an e-mail with brief information about the study, the expected duration, and compensation. Inclusion criteria were: (i) age between18 and 35 years and (ii) right-handedness. Exclusion criteria were self-reported neurological or motor disorders. All participants had a normal or corrected-to-normal vision and were not informed about the purpose of the study. All participants gave their informed written consent. The studies were conducted under the ethical standards laid down in the 1964 Declaration of Helsinki and were approved by the Local Ethical Committee (Comitato Etico Regione Liguria). According to the Ethical approval, we collected only demographic information about age, sex, and handedness. After having signed the consent form, participants filled in a series of questionnaires to address general attitudes towards robots[1] (see SM 1.1). Participants received 15 Euros for their participation.

In total, the data of 14 participants were excluded from data analysis (see data analysis section for further details). Therefore, the final sample size was N= 59, see Table 1 for demographics of each experimental condition.

[1] Please note that the questionnaires that were administered to have a qualitative description of our sample to be able to check if our experimental groups did not differ regarding a priori biases towards robots. Thus, average scores of the questionnaires are not considered a dependent measure of interest for our study and are reported in the Supplementary Materials.

Table 1. Demographic informations of the final sample across experimental conditions.

	iCub partner condition	*Computer partner condition*
Human-like erring condition	N= 14 6 males mean age = 21.8 ± 2.9 years	N= 13 4 males mean age = 22.3 ± 2.8 years
Mechanical erring condition	N= 16 6 males mean age = 22.8 ± 2.5 years	N= 16 7 males mean age = 25.8 ± 4.1 years

Experimental setup and stimuli

For a visual representation of the setup and the task, see Video 1 (https://osf.io/38yg6)

Our study was implemented in a multi-modal human-robot interactive scenario, where the participant and the robot interacted with a vertical touch screen by producing a set of periodic audio sequences. The experimental setup included a PC controlling the stimuli and responses, an iCub humanoid robot (Metta et al., 2008), and a vertically positioned multi-touch screen (1099.4 x 634.0 x 36.8 mm, 60Hz). iCub's pointing gesture was pre-defined as in Ciardo et al., 2019. The default trajectory time of iCub's arm was defined by design to be 350 ms. Participants performed the task standing next to iCub facing the touch screen (see Fig. 1). They were presented with a black screen divided into two equal portions by a white midline. In each hemifield of the screen, a music "pad" was presented. The pad consisted of an array of six coloured dots (Ø: 5 cm) positioned on the vertices of a hexagon, all equidistant from the center (see Fig. 1). Each dot corresponded to a specific tone. The pad was centered with respect to iCub's right arm. Tone duration was 450 ms and it was estimated empirically, based on iCub's minimum period achievable in the audio sequence.

Procedure

The task consisted of three phases: *training, teaching*, and *duet* where the duet was the actual experimental task. The first two phases were comparable across conditions

Training phase. Participants were instructed to invent a melody by tapping a sequence (sequence invented at participants' own will) of 6 different colored dots on the vertical touchpad. Participants were asked to repeat the sequence four times. Thus, each melody was composed of twenty-four taps. Participants could choose to start their sequence with any of the six dots and follow the order they preferred. However, each tone could be played only once within each sequence. Participants were instructed to try to keep a constant tempo. During the training, only participants' music pad was presented on the screen while the robot was standing next to them in a resting position (i.e., with its arm along the body). The training ended when participants were able to execute their melody (i.e., 24 taps) correctly ten times.

Teaching phase. Participants were instructed to "teach" the melody to iCub. Thus, while they were playing the sequence, iCub performed the same task as a follower. It reacted merely by repeating the participant's actions. Once a dot selection was detected, the task controller sent a request to the robot for tapping the same dot. To induce the belief that iCub was learning and improving during the teaching phase, we sequentially decreased the average delay of iCub's tap with respect to the participant's tap (delay condition: 650, 550, and 450ms). In this phase of the experiment, we decided not to add self-generated mistakes to the performance of the robot. See SM 2 for a detailed description of the teaching phase.

Duet phase. In this phase, participants were instructed as follows: *"iCub has now learned how to play your melody correctly with the right tempo, and now it can play on its own. Your task now is to play a duet with iCub, trying to maintain synchrony and not making errors".*

iCub's music pad was programmed to play following the average period estimated from the last four trials of the teaching phase. Specifically, the inter-tap interval between two consecutive taps was equal to the average time differences between two consecutive taps collected from the human participants in the last four trials of the teaching phase. In this way, we ensured that the robot's music pad was playing according to a tempo tailored to each participant.

Across participants, we manipulated the context of the interaction by programming the robot and the music pad to produce an error in 60% of the trials. For half of the participants, in the erroneous trials, the iCub switched one element of the melody by pressing the wrong key (Human-like error), while for the other half of participants it interrupted the melody and moved back and forth between two keys in an “endless’ loop (Mechanical error), see Video 1 (https://osf.io/38yg6). The duet phase included 24 trials in total, the number of correct and erroneous trials was 10 and 14, respectively. Correct and erroneous trials were fully randomized. Participants were not informed about how iCub and its music pad were programmed.

In all three experimental phases the trial procedure was as follows: At the beginning of each trial, the pad was presented as inactive (i.e., empty circles, see Video 1) with a central white circle. To begin the trial, participants had to press the white circle until it turned yellow, and the pad became active (i.e., all circle outlines turned into filled circles). Participants performed the tapping sequence always with their right arm.

In Experiment 2, the apparatus, stimuli, and procedure were the same as in the iCub experiment, with the only exception that instead of the iCub robot, the participants performed the task standing next to a computer controlling the task, see Video 2 (https://osf.io/38yg6/). To help participants to detect where the algorithm will point next, a white dot resembling the mouse cursor was presented with the music pad, giving the participants the impression that the algorithm was “tapping”.

Social inclusion

The willingness to include the agent as an in-group social partner was evaluated after the interactive task by using a ball-tossing game inspired by the Cyberball paradigm (Williams & Jarvis, 2006; Ciardo et al., 2020). Stimuli were pictures of a female human partner and a picture of either the iCub or the computer. The act of throwing the ball was simulated by presenting a 1sec video of a schematic ball moving. Participants were asked to choose to pass the ball to whoever they wanted. Each trial started with the presentation of pictures of the human player and the artificial co-agent, on the left and right sides of the screen, respectively[2], while the name of the participant was presented at the bottom. Upon receiving the ball, the participant had 500 ms to press either the “Z” or “M” key on a standard QUERTY-keyboard, to pass the ball to the human player (“Z”) or the artificial agent

[2] Please note that the choice of presenting the artificial agent always on the right side of the screen was motivated by the fact that during the interactive task, the partner was always located on the right side of the setup, see Fig.1.

player (“M”). Timeout was highlighted with a 500 ms feedback display. The task included 240 trials plus trials to replace timeouts. A short pause was given to participants after 120 trials. In the computer experiment, the picture of the iCub was replaced with a drawing of a computer.

Data analysis

In both experiments, we focused the analysis on the duet phase (see SM for data analysis of the teaching phase and questionnaires). We excluded data of participants that performed below 50% of the accuracy of correct trials (i.e., the trials in which the robot or the computer performed correctly). In total, the data of 5 participants were excluded from the iCub condition and the dataset of 9 participants from the Computer condition.

All the analyses were conducted using the lme4 package (Bates et al., 2014) in R. Parameter estimates (β) and their associated t-tests (t, p), calculated using the Satterthwaite approximation for degrees of freedom (Kuznetsova et al., 2015) are presented to show the magnitude of the effects, with bootstrapped 95% confidence intervals (Efron & Tibshirani, 1994).

Sensorimotor signaling

To address sensorimotor signaling we focused only on correct trials, namely, trials in which both the artificial agent and the participant performed the sequence correctly[3]. The dependent variables were the average inter-tap interval (ITI) asynchrony and its variability. First, we estimated the ITI as the time interval between two consecutive taps of the same trial, namely the difference between the timestamps of a tap and its previous one. The first tap of each trial was excluded. Then, we estimated the difference between the ITI of the artificial agent (iCub or computer) and the ITI of the human, i.e., “ITI asynchrony”. The standard deviation of the ITI asynchrony is the “variability in ITI asynchrony”, see SM Fig 2.

[3] Please note that the analysis was focused on correct trials only because in erring trials, i.e., when the robot made errors, the motor behaviours of the robot and the motor behaviour of the participant differered in their spatial componets (i.e. distance between two consecutive taps). Also, given the random nature of errors, such differences in the spatial aspects of the robot’s movement were not comparable neither across trials nor participants.

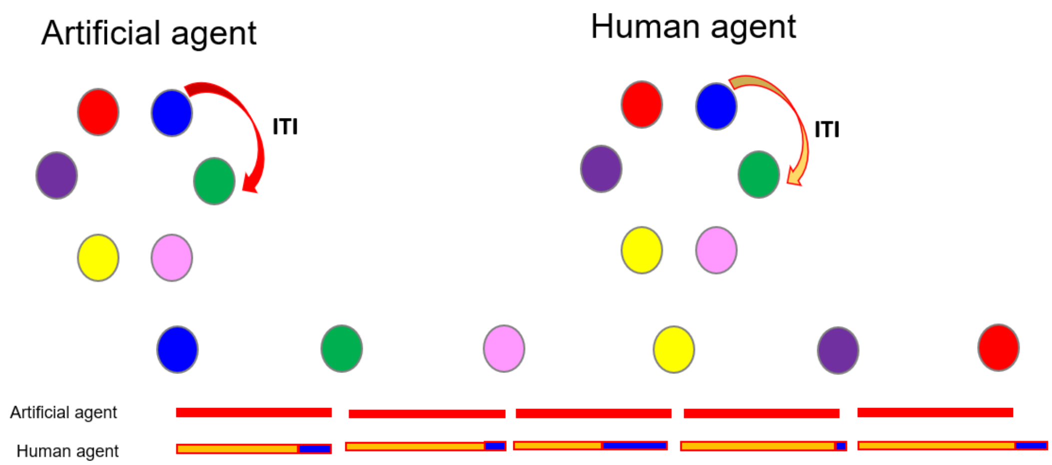

Figure 1. The figure depicts how Intertap Interval was defined for each agent within each trial (24 taps). Please note that the picture is a mere graphical representation of inter tap intervals (ITI) and that the length of the depicted segment is arbitrary. For each agent, the average tapping period was estimated as the time interval between two consecutive taps of the same trial, namely the difference between the timestamps of a tap and its previous one, also defined as ITI, after the exclusion of the first tap of each trial. Thus, the average period of the human agent is the average of all the yellow segments. The average period for the artificial agent (iCub or computer) corresponded to the average of all the red segments. ITI asynchrony is the difference between the ITI of the artificial agent (iCub or computer) and the ITI of the human. ITI asynchrony is depicted as the blue segments. Thus, mean ITI asynchrony and mean variability in ITI asynchrony correspond to the mean and standard deviation of the blue segments, respectively

Mean ITI asynchrony and variability in ITI asynchrony were estimated for each trial as indexes of signaling between the human and the artificial partner. Both measures were then compared separately across the Erring conditions (Human-like or Mechanical). Although mean ITI asynchrony and variability in ITI asynchrony are related measures, they reflect two different aspects of performance. The former indicates on average how well participants were able to predict the tapping time of the partner and to perform their action accordingly within a single trial. The second one reflects how well they were able able to maintain such precise predictions within a given trial. The first one is related to motor control capabilities, i.e. the ability to predict (based on the action-perception link) when an event will occur and when the required action is to be performed. On the other hand, variability in performance indicates how much participants adapt their performance to the asynchronies they detected between their own and others' actions. It is identified as an index of coordination strategies.

Task Performance

To address how Human-like or Mechanical erring behaviour affects task performance when coordinating our actions with artificial agents, we focused on error rate and tapping period. Errors were defined as trials in which the participant pressed incorrectly the buttons of the music pad or did not complete the sequence, i.e. less than 24 keypresses were performed. The tapping period of each trial was estimated as the mean of all the time differences between two consecutive taps within the same 24 –tap sequence, see Fig.2.

Arcsine-transformed error rate and Average tapping period were modeled as a function of Error Occurrence, i.e., if the trial included an error from the artificial partner (Yes, No), and Erring condition (Human-like or Mechanical), plus their interactions, as fixed effects, and participants as a random effect.

Social inclusion of the artificial agent

The frequencies of choosing to toss the ball toward the iCub robot or the computer were submitted to a logistic mixed model with Erring condition as a fixed effect and participant as a random effect.

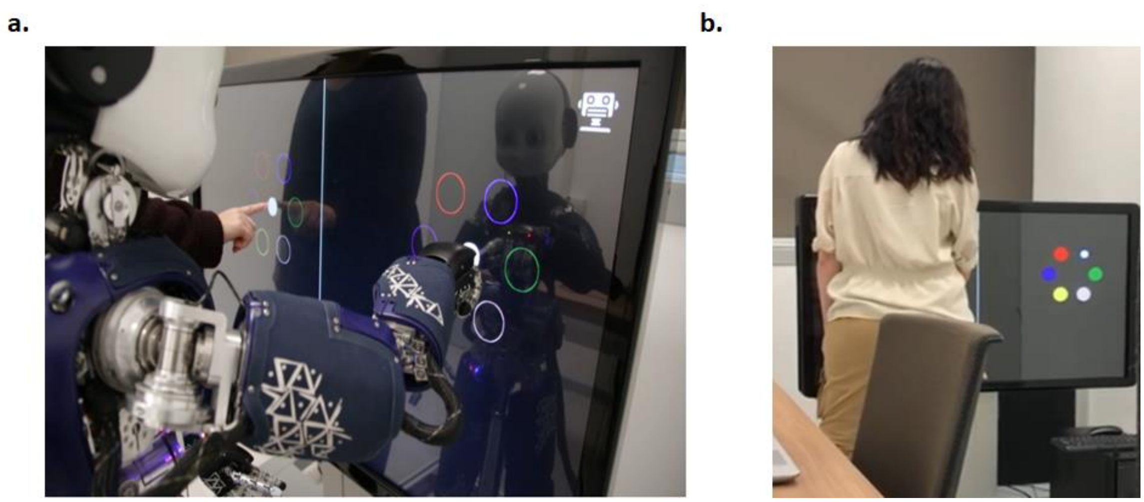

Figure 1. Panel a: iCub experiment setup. Panel b: Computer setup experiment

Results

Sensorimotor signaling

Experiment 1: iCub partner

Results showed that although the mean ITI asynchrony did not differ across conditions [Wilcoxon-test = 0.356, p-value = 0.706], variability in ITI asynchrony was affected by the Erring condition [Wilcoxon-test = 0.502, p-value = 0.001]. Specifically, the variability in ITI asynchrony was lower for those participants who interacted with the iCub committing human-like errors compared to those who interacted with the mechanically-erring robot, see Fig.3.

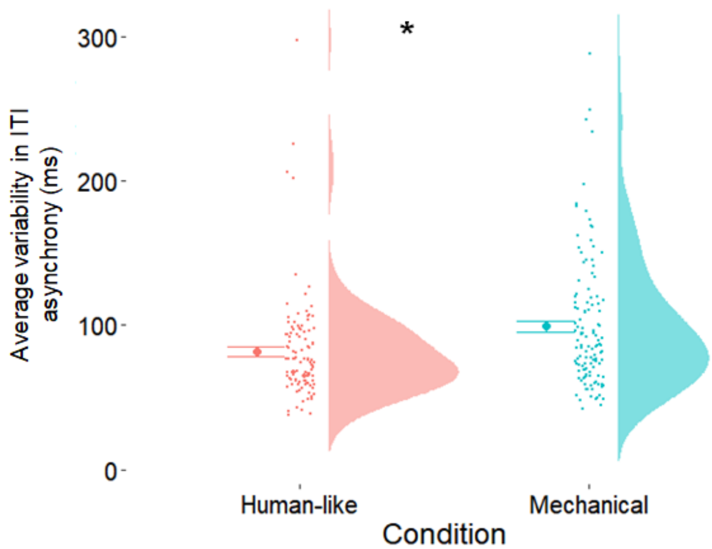

Figure 3. Average variability in ITI asynchrony plotted as a function of Erring condition (Human-like or Mechanical) in Experiment 1 (iCub partner). Asterisk denotes a significant difference.

Experiment 2: Computer partner

Results showed that the mean ITI asynchrony did not differ across conditions [Wilcoxon-test = 0.326, p-value = 0.121]. However, the variability in ITI asynchrony was modulated by the Erring condition [Wilcoxon-test = 0.757, p-value = 0. 001]. Specifically, the variability in ITI asynchrony was lower for participants who experienced the computer erring in a mechanical way compared to those who were exposed to the human-like error context, see Fig.4.

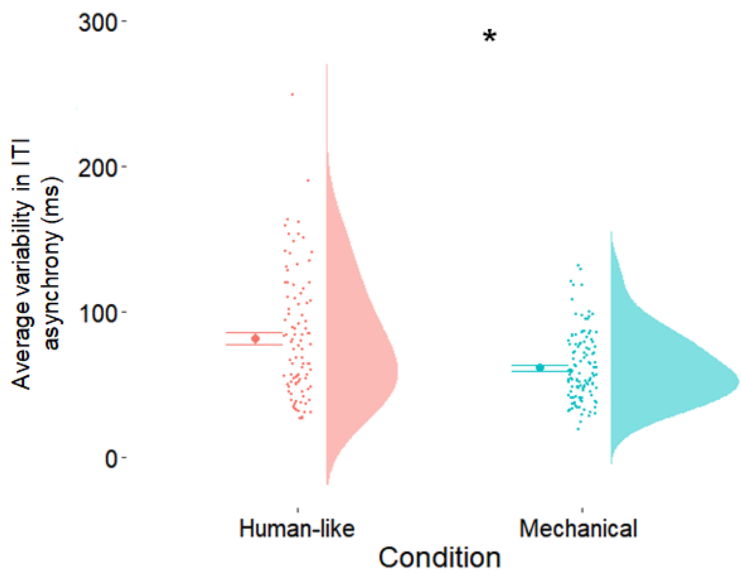

Figure 4. Average variability in ITI asynchrony *plotted as a function of Erring condition (Human-like or Mechanical) in Experiment 2 (Computer partner). Asterisk denotes a significant difference.*

Task Performance

Experiment 1: iCub partner

Results showed a marginal interaction between Error Occurrence and Erring condition on error rates [β = 0.196, t_{28} = 1.951, p = 0.061, 95% CI = (0.018, 0.377)], see Fig.5a. No main effect or interaction was found for average tapping period [all p > 0.250].

Experiment 2: Computer partner

Results showed a main effect of Error Occurrence on error rates [β = 0.181, t_{27} = 2.945, p = 0.007, 95% CI = (0.061, 0.301)] with higher error rate for trials in which the computer made an error (0.34±0.18%) compared to when it performed the melody correctly (0.20±0.15%), see Fig. 5b. No main effect or interaction was found for average tapping period [all p > 0.511].

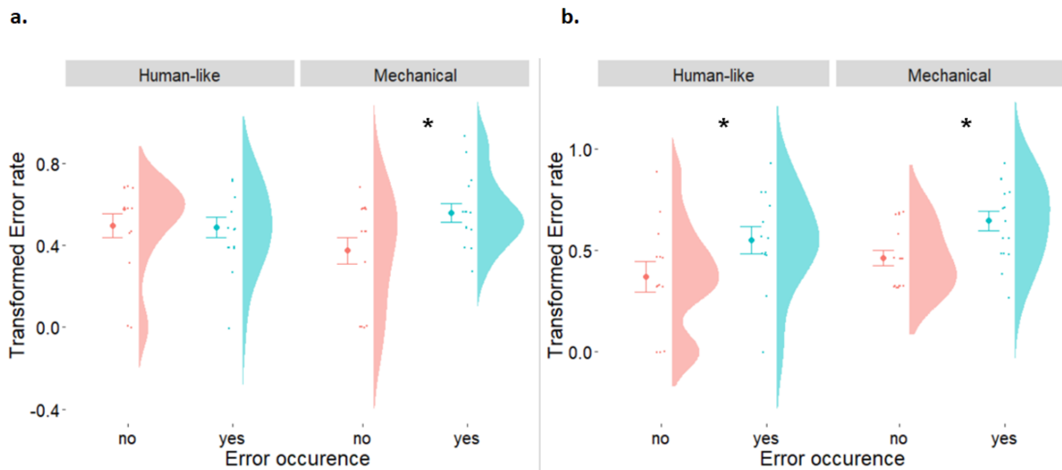

Figure 5. Arcsin-transformed error rate as a function of Error Occurrence, i.e., if the trial included or not an error from iCub (Yes, No), and Erring condition (Human-like or Mechanical) for the iCub(panel a) and the computer (panel b) experiment. Asterisks denote significant differences.

Social inclusion of the artificial agent

Experiment 1: iCub partner

The analysis revealed that the probability to choose the robot instead of the human player was significantly increased for those participants who interacted with a human-like erring robot [β = .45, z = 6.25, p < 0.001, CI = (0.31; 0.59)], see Table 2 and Fig. 6a. Specifically, the increase in the probability of choosing iCub was 1.57 times higher.

Experiment 2: Computer partner

The analysis revealed that the probability to choose the computer did not differ between the two Erring conditions [β = -.15, z = -1.22, p = 0.223, CI = (-0.40; 0.10)], see Fig.6b.

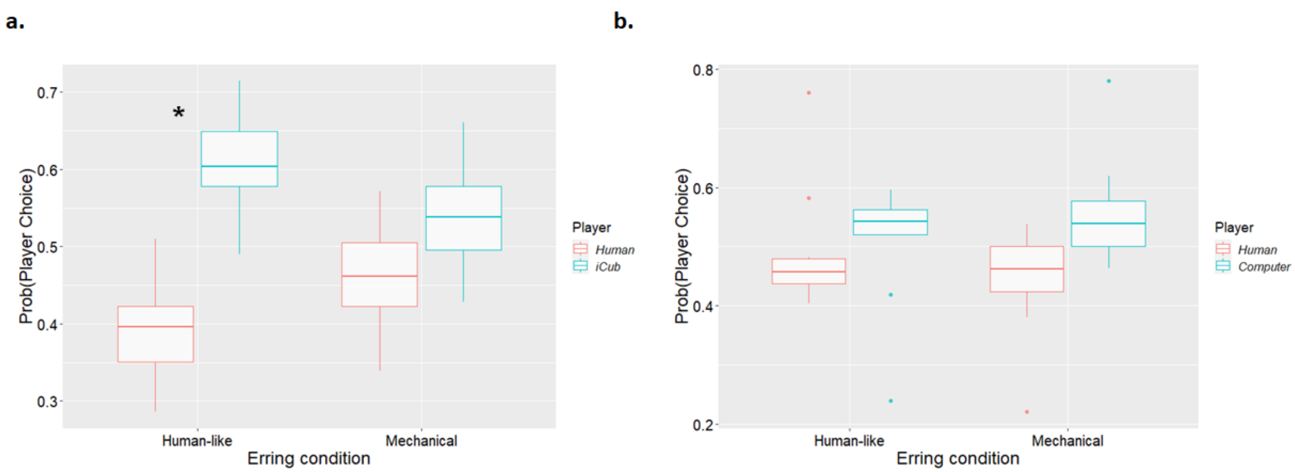

Figure 6. The frequencies of choosing to toss the ball toward the human or the artificial player are plotted as a function of the Erring Condition (Human-like or Mechanical) for the iCub (panel a) and the computer (panel b) experiment. Asterisk denotes a significant difference.

Table 2. The average percentage of choosing the artificial agent in the ball-tossing game as a function of the erring behaviour condition and experiment.

Experiment	Erring behaviour condition	
iCub experiment	Human-Like	61%
	Mechanical	54%
Computer experiment	Human-Like	52%
	Mechanical	55%

Discussion

The present study aimed to evaluate whether (i) humans would exhibit sensorimotor signaling towards an erring artificial agent (robot or computer) in a joint action task, (ii) the signaling behaviour would be dependent on the human-likeness of the agent's erring behaviour, and (iii) if the signaling behaviour would result in a higher tendency to socially include the artificial agent. To address these questions, we designed an experimental paradigm in which human participants were asked to play a musical duet either with the iCub humanoid robot or with an algorithm on a computer that were programmed to make human-like or mechanical errors during their performance. After the joint task, we tested – with an adapted Cyberball (Williams & Jarvis, 2006; Ciardo et al., 2020) game – whether our manipulation affected also the willingness to include the agent as an in-group social partner.

Joint action with artificial agents: Sensorimotor signaling and adaptation to partner's errors

We predicted that the human-likeness of erring behaviour displayed by the agent should affect sensorimotor signaling and adaptation to partner's errors, especially when the agent has a human-like shape and motor repertoire. Specifically, we reasoned that when the erring behaviour belongs to a human repertoire, the error itself does not compromise the DMP, as participants are still able to represent it in terms of its effect on the joint goal. On the contrary, mechanical errors presumably compromise the DMP, thereby impairing adaptation. Thus, we expected higher accuracy, faster performance, and lower variability when the embodied agent committed a human-like error compared to when it failed in a mechanical-way. Our results showed that indeed, when interacting with iCub, participants showed lower variability in their performance when the error was a mismatch of the sequence, resembling a human error, compared to when the robot showed a mechanical error. Reduction of behavioural variability during joint action has been considered as a form of nonverbal signaling (Sebanz & Knoblich, 2021; Vesper & Sevdalis, 2020) which aims to make oneself more predictable to help the partner maintain and recover coordination when a joint goal is established (e.g., Vesper et al., 2011; McEllin et al., 2018; Sacheli et al., 2013). This is in line with Sacheli and colleagues' study (2021), showing that in human-human interactions, the violation of expectations driven by the partner's error triggers an implicit tendency to correct the error, sacrificing individual efficiency in favor of sensorimotor signaling. In a similar vein, in our study, when the robot error occurred, participants reduced their variability to facilitate coordination with iCub. Interestingly, this occurred only when the erring behaviour resembled a human-like error. In such condition, participants were still able to represent the errors in terms of their effectiveness in reaching (or not) the joint goal, namely, playing the melody in synchrony. Thus, although the iCub was unreliable as a partner, participants were still able to establish a DMP allowing them to evaluate and predict how the robot was (or not) contributing to reach the joint goal and they adapted their performance consequently. When the iCub error was a mechanical failure, the impossibility to represent iCub's behaviour in terms of its effects presumably prevented the establishment of a DMP. As a consequence, participants might have not been able to adapt their performance and ended out of the loop, by committing significantly more errors in the trials in which also iCub committed errors, relative to correct trials.

On the other hand, when the partner was a computer algorithm run on a standard PC, participants showed a reversed pattern. Specifically, lower variability in performance was found for the condition in which the computer displayed a mechanical failure compared to the human-like error. Such result was unexpected, as when the partner is an agent with a non-human like motor repertoire, the lack of a motoric component of the action should have prevented any form of action simulation or motor resonance (Blakemore & Frith; 2005; Wilson & Knoblich, 2005; Grafton, 2009; Schubert & Semin; 2009). As a consequence, the non-motoric "action" of a computer should not have been represented in terms of their contribution to the DMP, irrespectively of the human-likeness manipulation. It might be that in the case of the computer algorithm partner, participants decreased their variability in the condition in which the agent displayed behaviour that was better fitting to their representation of that agent. A standard computer might have been represented by participants as a mechanical device and thus, its mechanical erring behaviour might have fit participants' expectations. This might have elicited higher degree of implicit mechanisms of cooperative signaling.

This would, however, suggest an alternative explanation for the effect in Experiment 1 (with iCub). Also, in this case, the effect might not have been driven so much by the human-likeness per se, but rather how much a behaviour fits to the representation of the agent. A humanoid robot resembles a human more than a disembodied computer algorithm and thus a human-like error is what participants might have expected from the robot. This might have induced cooperative signaling.[4]

Apart from behavioural variability, participants' performance accuracy was also affected by the erring behaviour differently across the two experiments. When interacting with the iCub robot, the failures of the co-agent only marginally affected accuracy. However, when the partner was the computer, participants' performance was influenced by the erring behaviour. Specifically, the error rate was higher for those trials in which the computer failed. Such a result indicates that participants were not able to complete the task when the computer failed. Such a phenomenon is well known in human-factors literature as the Out-of-the-Loop problem (OOTL), i.e., the impairment in human performance in performing a task when a failure in an automatized system occurs. Humans that are OOTL usually take longer or are unable to decide if, and how, they should intervene (Berberian, et al., 2017; Norman, 1991). The OOTL phenomenon has been listed as one of the major causes of incidents in a highly automatized work environment, such as air traffic control in civilian aviation (Norman, 1991). The fact that errors of the iCub only marginally affected performance accuracy suggests that the human-like motor repertoire of the robot prevented participants to end up OOTL. Interestingly, the marginal interaction effect on error rate when the partner was the iCub seemed to be driven by the mechanical erring condition (cf. Fig. SM 4a), speaking in favor of the idea that inability to represent an error in the context of DMP results in the OOTL phenomenon.

Taken together, our results suggest that during joint action with artificial agents physical appearance of the partner and behavioural human-likeness may interact. Specifically, it might be that the human-like appearance and motor repertoire of the iCub might have triggered a different representation, expectation, and prediction about its behaviour than a standard computer. Indeed, several studies showed that despite their artificial nature, humanoid robots can trigger in humans attribution of intentionality (for a review see Perez-Osorio & Wykowska, 2020). According to Daniel

[4] It is important to highlight that the differences in the sensorimotor signaling across agents cannot be explained by participants' lack of error perception. When asked directly, participants reported a failure in the partner behaviours equally across conditions (see SM 3).

Dennett (1971; 1987), when interacting with an agent, humans adopt different types of stances, i.e., Intentional, Design/ Mechanical, or Physical. When the partner is another human, we adopt the intentional stance, namely, we explain and predict their behaviour (and errors) as resulting from mental operations. When facing artificial and mechanical systems, like computers, their behaviour (and failures) are explained and predicted referring to the way they were designed or programmed to act, i.e., we adopt the design/mechanical stance. A recent series of studies showed that humans can explain behaviours of the iCub robot using both intentional and mechanical stances (Marchesi et al., 2019).

The adoption of an Intentional or Mechanical stance might be crucial for the type of internal model we build about artificial agents, affecting, in consequence, also the DMP. Indeed, within a DMP, we represent our own and others' actions in terms of their contribution to the achievement of the joint goal. That is, the different types of internal models that we have regarding our partner will not only result in different expectations about how s/he can contribute to the joint goal, but also in different representations of how *we* need to contribute to it. This happens, for example, when we interact with partners of different physical characteristics or expertise in a task. For example, when kayaking, the DMP will be different, depending on whether our partner is a child or an instructor. In the former case, the DMP relies on the representation of a partner that is not as strong as we are, which brings us to expect that s/he contributes less to the paddling. In contrast, when the partner is an instructor, the DMP relies on the representation of a partner that has more expertise than we do, resulting in the expectation that s/he would contribute to paddling substantially. In our study, when the iCub's error was human-like, the violation of expectations related to the error was still plausible within the internal representation of the robot as an intentional agent. As a consequence, participants could adapt their performance as they would with another intentional agent. On the contrary, when the robot failed mechanically, the error was not plausible within their representation of the "intentional" robot, thus participants were not able to explain the error and ended "out of the loop". In a similar vein, when the partner was the computer, participants were able to interpret and adapt, only when the failure was plausible within the representation of the computer agent, namely, in the mechanical erring condition, which fit the "computer" representation and expectations.

Sensorimotor signaling as the basis for social inclusion of artificial agents

The final aim of the study was to evaluate the impact of sensorimotor signaling during a joint task on the social inclusion of artificial agents (i.e., willingness to include the agents as in-group social partners). Thus, after the joint action tasks, participants performed a ball-tossing game inspired by the Cyberball paradigm (Williams & Jarvis, 2006; Ciardo et al., 2020). We predicted that after the interactive task, participants should prefer to interact again with the agent after it showed a human-like erring behaviour rather than a mechanical one. This should be particularly pronounced for the human-like robot agent, due to its more social presence. Results showed that indeed the probability of choosing iCub as the receiver of the ball, instead of the human avatar, was higher for those participants who interacted with the human-like erring robot, relative to those who interacted with the robot which was erring in a mechanical way. Interestingly, this was not the case in the standard computer experiment. Indeed, after the interaction with the computer partner, the probability to toss the ball toward the artificial partner was equal across erring conditions, suggesting that the effect on social inclusion is not driven by the violation of expectation per se.

These results suggest a transfer effect between the interactive task and the willingness to include iCub as an in-group social partner. Specifically, the possibility to maintain a dyadic motor plan during the joint task might have led participants to perceive the interaction as smoother and the iCub as a trustworthy partner, despite the errors. Also, it is possible that the human-like error increased the perceived similarity between participants and the iCub, resulting in a group membership effect. Previous evidence showed that the perceived similarity between self and partner is crucial in affecting social cognition mechanisms (Ciardo et al., 2021). Ciardo and colleagues showed that joint attention is influenced by both implicit and explicit cues of similarity elicited by age (Ciardo et al., 2014; 2021) or the attitude of the partner during the interaction (Ciardo et al., 2015). Similarly in our study, participants might have perceived the human-like behaviour as a more cooperative attitude. Indeed, although in both conditions iCub made an error in 60% of the trials, in the human-like condition, following the error, it continued to play, although incorrectly. On the contrary, in the mechanical erring condition, the robot continued moving back and forth between keys interrupting playing altogether. In the former case, participants might have perceived the behaviour of iCub as an attempt to recover from its error. Thereby, they might have perceived the robot as cooperative or more committed to reaching the joint goal. Notably, this effect was not due only to the behaviour, as subsequent to the interaction with the computer failing in a human-like way, participants did not show a preference to interact with it again. Thus, the social inclusion of artificial agents is influenced by a joint effect of human-likeness of appearance and of behaviour.

Limitations and future directions

The study has some limitations that need to be addressed in future research. Firstly, we used a computer as a partner in Experiment 2. Such a choice did not allow us to directly compare participants' performance between the two agent conditions. Indeed, although the actions of the robot and the computer were comparable in terms of their effects in reaching (or not) the joint goal, they differed in the amount of information associated with them. iCub's actions were characterized not only by the visual and auditory effects they were producing on the music pad but also by motoric information that was lacking when the partner was a computer program running on a standard PC.

Another point that can be examined in future research is the manipulation of the reliability of the robot. Indeed, in our study, the iCub (and the computer as well) were committing an error in the majority of the trials (60%), thus they were unreliable partners. It remains to be answered whether the frequency at which the artificial agent violates our expectations can affect sensorimotor signaling and social inclusion. Finally, in the mechanical erring condition, participants might have interpreted the whole system as being faulty, instead of a failure of the agent in completing the task. Future studies should include a specific question about this possibility in the manipulation check interview.

Conclusions

In the present study, we examined conditions under which artificial partners elicit sensorimotor signaling in a joint musical task, and what are the consequences of erring behaviour on social inclusion of artificial agents. Our results showed that when interacting with artificial agents, human-likeness both in physical appearance and in the behaviour of an artificial agent have an interactive impact on coordination and social inclusion in joint tasks with artificial agents.

Acknowledgments

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.

FC was in part supported by H2020–Marie Skłodowska-Curie grant agreement no. 893960.

The content of this paper is the sole responsibility of the authors. The content of this paper 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.

Credit Authors Statement:

FC conceived, designed, and performed the study; collected and analyzed the data, discussed and interpreted the results; wrote the manuscript.

DDT integrated and programmed the technical components of the experimental task.

AW conceived and designed the study; discussed and interpreted the results; wrote the manuscript.

All authors reviewed the manuscript.

References

Abubshait, A., Weis, P. P., & Wiese, E. (2021). Does context matter? Effects of robot appearance and reliability on social attention differs based on lifelikeness of gaze task. *International Journal of Social Robotics, 13*(5), 863-876.

Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol., 59*, 617-645.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823.

Berberian, B., Somon, B., Sahaï, A., & Gouraud, J. (2017). The out-of-the-loop Brain: a neuroergonomic approach of the human automation interaction. *Annual Reviews in Control, 44*, 303-315.

Blakemore, S. J., & Frith, U. (2005). The learning brain: Lessons for education. Blackwell publishing.

Brand, R. J., Baldwin, D. A., & Ashburn, L. A. (2002). Evidence for 'motionese': modifications in mothers' infant-directed action. *Developmental science, 5*(1), 72-83.

Ciardo, F., & Wykowska, A. (2020). La robotica assistenziale sociale come strumento per promuovere lo sviluppo socio-cognitivo: vantaggi, limiti e prospettive future. *Sistemi intelligenti, 32*(1), 9-25.

Ciardo, F., Marino, B. F., Actis-Grosso, R., Rossetti, A., & Ricciardelli, P. (2014). Face age modulates gaze following in young adults. Scientific reports, 4(1), 1-7.

Ciardo, F., De Tommaso, D., & Wykowska, A. (2019, March). Humans socially attune to their "follower" robot. In 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI) (pp. 538-539). IEEE.

Ciardo, F., Beyer, F., De Tommaso, D., & Wykowska, A. (2020). Attribution of intentional agency towards robots reduces one's own sense of agency. *Cognition, 194*, 104109.

Ciardo, F., Ricciardelli, P., Lugli, L., Rubichi, S., & Iani, C. (2015). Eyes keep watch over you! Competition enhances joint attention in females. *Acta Psychologica, 160*, 170-177.

Ciardo, F., De Angelis, J., Marino, B. F., Actis-Grosso, R., & Ricciardelli, P. (2021). Social categorization and joint attention: Interacting effects of age, sex, and social status. *Acta Psychologica, 212*, 103223.

Ciardo, F., Ghiglino, D., Roselli, C., & Wykowska, A. (2020, November). The effect of individual differences and repetitive interactions on explicit and implicit attitudes towards robots. In International Conference on Social Robotics (pp. 466-477). Springer, Cham.

Dennet, D. (1987). True Believers. Intentional Stance, 13, 36.

Dennett, D. C. (1971). Intentional systems. *The Journal of Philosophy, 68*(4), 87-106.

Deutsch, M. (2011). Cooperation and competition. In Conflict, interdependence, and justice (pp. 23-40). Springer, New York, NY.

Efron, B., & Tibshirani, R. J. (1994). An introduction to the bootstrap. CRC press.

Ghiglino, D., Willemse, C., De Tommaso, D., & Wykowska, A. (2021). Mind the eyes: artificial agents' eye movements modulate attentional engagement and anthropomorphic attribution. *Frontiers in Robotics and AI, 8*.

Grafton, S. T. (2009). Embodied cognition and the simulation of action to understand others. *Annals of the New York Academy of Sciences, 1156*(1), 97-117.

Herwig, A. (2015). Linking perception and action by structure or process? Toward an integrative perspective. *Neuroscience & Biobehavioral Reviews, 52*, 105-116.

Hinz, N. A., Ciardo, F., & Wykowska, A. (2021). ERP markers of action planning and outcome monitoring in human–robot interaction. *Acta Psychologica*, *212*, 103216.

Horvitz, E. (2016). One hundred year study on artificial intelligence.

Hove, M. J., & Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. *Social cognition, 27*(6), 949-960.

Jeannerod, M. (2006). Motor cognition: What actions tell the self (No. 42). Oxford University Press.

Kokal, I., Engel, A., Kirschner, S., & Keysers, C. (2011). Synchronized drumming enhances activity in the caudate and facilitates prosocial commitment-if the rhythm comes easily. *PloS one, 6*(11), e27272.

Koterba, E. A., & Iverson, J. M. (2009). Investigating motionese: The effect of infant-directed action on infants' attention and object exploration. *Infant Behavior and Development, 32*(4), 437-444.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). Package 'lmertest'. R package version, 2(0), 734.

Loehr, J. D., Kourtis, D., Vesper, C., Sebanz, N., & Knoblich, G. (2013). Monitoring individual and joint action outcomes in duet music performance. *Journal of cognitive neuroscience, 25*(7), 1049-1061.

Marchesi, S., Ghiglino, D., Ciardo, F., Perez-Osorio, J., Baykara, E., & Wykowska, A. (2019). Do we adopt the intentional stance toward humanoid robots?. *Frontiers in psychology, 10*, 450.

Martini, M. C., Buzzell, G. A., & Wiese, E. (2015, October). Agent appearance modulates mind attribution and social attention in human-robot interaction. In International Conference on Social Robotics (pp. 431-439). Springer, Cham.

McEllin, L., Sebanz, N., & Knoblich, G. (2018). Identifying others' informative intentions from movement kinematics. *Cognition, 180*, 246-258.

Metta, G., Sandini, G., Vernon, D., Natale, L., & Nori, F. (2008, August). The iCub humanoid robot: an open platform for research in embodied cognition. In *Proceedings of the 8th workshop on performance metrics for intelligent systems* (pp. 50-56).

Michael, J., Sebanz, N., & Knoblich, G. (2016). Observing joint action: Coordination creates commitment. *Cognition, 157*, 106-113.

Michael, J., McEllin, L., & Felber, A. (2020). Prosocial effects of coordination–What, how and why?. *Acta Psychologica, 207*, 103083.

Moreau, Q., Tieri, G., Era, V., Aglioti, S. M., & Candidi, M. (2021). The monitoring system is attuned to others' actions during dyadic motor interactions. *bioRxiv*.

Norman, D. A. (1991). Cognitive artifacts. *Designing interaction: Psychology at the human-computer interface, 1*(1), 17-38.

Perez-Osorio, J., & Wykowska, A. (2020). Adopting the intentional stance toward natural and artificial agents. *Philosophical Psychology, 33*(3), 369-395.

Pezzulo, G., Donnarumma, F., & Dindo, H. (2013). Human sensorimotor communication: A theory of signaling in online social interactions. *PloS one, 8*(11), e79876.

Prinz, W. A common coding approach to perception and action. In *Relationships between perception and action* pp. 167-201 (Springer, Berlin, Heidelberg, 1990).

Remland, M. S., Jones, T. S., & Brinkman, H. (1995). Interpersonal distance, body orientation, and touch: Effects of culture, gender, and age. *The Journal of social psychology, 135*(3), 281-297.

Sacheli, L. M., Musco, M. A., Zazzera, E., & Paulesu, E. (2021). Mechanisms for mutual support in motor interactions. *Scientific reports, 11*(1), 1-16.

Sacheli, L. M., Tidoni, E., Pavone, E. F., Aglioti, S. M., & Candidi, M. (2013). Kinematics fingerprints of leader and follower role-taking during cooperative joint actions. *Experimental brain research, 226*(4), 473-486.

Sacheli, L. M., Arcangeli, E., & Paulesu, E. (2018). Evidence for a dyadic motor plan in joint action. *Scientific reports, 8*(1), 1-12.

Sartori, L., Becchio, C., Bara, B. G., & Castiello, U. (2009). Does the intention to communicate affect action kinematics?. *Consciousness and cognition, 18*(3), 766-772.

Schubert, T. W., & Semin, G. R. (2009). Embodiment as a unifying perspective for psychology. *European Journal of Social Psychology, 39*(7), 1135-1141.

Sebanz, N., & Knoblich, G. (2021). Progress in Joint-Action Research. Current Directions in *Psychological Science, 30*(2), 138-143.

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific american, 223*(5), 96-103.

Tomeo, E., Cesari, P., Aglioti, S. M., & Urgesi, C. (2013). Fooling the kickers but not the goalkeepers: behavioral and neurophysiological correlates of fake action detection in soccer. *Cerebral Cortex, 23*(11), 2765-2778.

Vesper, C., & Sevdalis, V. (2020). Informing, coordinating, and performing: A perspective on functions of sensorimotor communication. *Frontiers in Human Neuroscience, 14*, 168.

Vesper, C., Butterfill, S., Knoblich, G., & Sebanz, N. (2010). A minimal architecture for joint action. *Neural Networks, 23*(8-9), 998-1003.

Vesper, C., Van Der Wel, R. P., Knoblich, G., & Sebanz, N. (2011). Making oneself predictable: Reduced temporal variability facilitates joint action coordination. *Experimental brain research, 211*(3-4), 517-530.

Vesper, C., et al. (2017). Joint action: mental representations, shared information and general mechanisms for coordinating with others. *Frontiers in psychology, 7*, 2039.

Wiese, E., Metta, G., & Wykowska, A. (2017). Robots as intentional agents: using neuroscientific methods to make robots appear more social. *Frontiers in psychology, 8,* 1663.

Williams, K. D., & Jarvis, B. (2006). Cyberball: A program for use in research on interpersonal ostracism and acceptance. *Behavior research methods, 38*(1), 174-180.

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological bulletin, 131*(3), 460.

Wykowska, A. (2020). Social robots to test flexibility of human social cognition. *International Journal of Social Robotics, 12*(6), 1203-1211.

Supplemental Material:

Joint action with artificial agents: human-likeness in behaviour and morphology affects sensorimotor signaling and social inclusion

SM 1 Questionnaires addressing attitudes towards robots.

After arriving in the lab participants filled out the following questionnaires:

•The Frankenstein Syndrome Questionnaire (FSQ [1]): self-report scales that investigate the anxiety perceived towards robots in contexts of interaction.
•The Negative Attitudes Towards Robots Scale (NARS [2]): self-report scales that investigate negative attitudes towards robots
•The Robotic Social Attitude Scale (RoSAS [3]): a self-report questionnaire that investigates the attribution of anthropomorphic characteristics to robots
Self-report questionnaires presentation and data collection were controlled by OpenSesame software. Average scores for each subscale are presented in Table 1 for iCub and Computer partner separately (Experiment 1 and 2, respectively).
The analyses on questionnaire responses between the two experiments show that there were no differences in participants' general attitudes towards robots across the two experiments (cf. Table 1)

Table1. Average scores and standard deviations for the subscales of the NARS, FSQ, and RoSAS.

Questionnaire	Subscale	iCub partner M	iCub partner SD	Computer partner M	Computer partner SD	Mann-Whitney-Wilcoxon Test
FSQ	General anxiety toward humanoid robots	37.38	11.93	34.21	11.08	W = 353 p-value = 0.297
	Apprehension toward social risks of humanoid robots	24.38	5.31	22.21	4.76	W = 331.5 p-value = 0.167
	Trustworthiness of developers of humanoid robots	23.48	5.55	21.07	3.81	W = 296, p-value = 0.053
	Expectations for humanoid robots in daily life	27.14	7.98	27.83	5.93	W = 516 p-value = 0.138
NARS	Negative attitudes toward situations and interactions with robots	11.76	4.20	9.69	3.02	W = 332.5 p-value = 0.170
	Negative attitudes toward the social influence of robots	13.28	4.11	12.55	3.81	W = 376.5, p-value = 0.497
	Negative attitudes toward emotions in interaction with robots	8.03	8.03	7.14	7.14	W = 332.5, p-value = 0.170
RoSAS	Competence	7.02	1.27	7.45	1.19	W = 525.5, p-value = 0.104
	Discomfort	2.69	1.08	2.94	1.28	W = 458.5, p-value = 0.559
	Warmth	3.10	1.51	3.84	1.73	W = 534, p-value = 0.0785

SM 2 Detailed Teaching phase procedure.

In the teaching and duet phases, participants were presented with their music pad and an identical music pad in front of the robot. The robot's music pad played in a pre-programmed way, and the robot was moving its hand and finger in line with each successive "tap" to give the impression that it was causing the tone to play. It was located at the minimum possible distance to avoid damages to both the robot's arm and the screen. Participants were told that iCub's tapping was executed in a touchless manner by means of an infrared system embedded into the touchscreen's frame. Before beginning these phases of the task, we showed and asked participants to experience the touchless tapping modality using the infrared frame of the touchscreen.

In the teaching phase, participants were instructed to teach the melody to iCub. Thus, while they were playing the sequence, iCub performed the same task as a follower. It reacted merely by repeating the participant's actions. Once a dot selection was detected, the task controller sent a request to the robot for tapping the same dot. To induce the belief that iCub was learning and improving during the teaching phase, we manipulated the average delay of iCub's tap (delay condition). In this phase of the experiment, we decided not to add self-generated mistakes to the performance of the robot. The delay introduced represented the iCub's response time for tapping a single dot in relation to the human's tap, that is, the time between the detection of the participant's tap and the iCub's tap. The delay conditions were: 650, 550, and 450 ms. The values have been selected empirically taking into account iCub's arm movement trajectory time (350 ms) ± estimated variability of the position controller (100ms) [4]. The teaching phase comprised 22 trials, in which participants had to perform their 24-dots sequence keeping the tempo as constant as possible. In the first 6 trials, the iCub performed with a delay of 650 ms. Then, in the 6 subsequent trials, iCub's performance was delayed by 550 ms. In the last 10 trials, iCub tapped on each dot with a delay of 450 ms, giving the impression that it has improved its performance and learned the melody. Participants' performance in the last four trials was used to model iCub' s behaviour in the duet phase. If participants made an error in executing their trial, the trial was aborted and restarted.

SM 3 Data analysis and results of the Teaching phase.

In the teaching phase, participants were instructed to teach the melody to iCub. Thus, while they were playing the sequence, iCub performed the task as a follower. Across the teaching phase, the robot reduced its delays on tapping with respect to participants' tap: Delay conditions were: 650, 550, and 450 ms. Participants' average periods were modeled as a function of Delay condition as a fixed effect and participants as a random effect. Analyses were conducted using the lme4 package [5] in R. Parameter estimates (β) and their associated t-tests (t, p), calculated using the Satterthwaite approximation for degrees of freedom [6] are presented to show the magnitude of the effects, with bootstrapped 95% confidence intervals. The analysis was run separately for the iCub and Computer experiments.

Experiment 1: iCub partner

Results showed that compared to the 650 ms delay condition participants performed faster only when the iCub performed with the shortest delay (450ms) [$\beta = 135.87$, $t_{29.81} = 6.761$, $p < 0.001$, 95% CI = (96.482, 175.250)] (822.97 vs 958.84 ms). See Fig1a.

Experiment 2: Computer partner

Results showed that, compared to the 650 ms delay condition, participants performed faster both when the computer played with an intermediate delay (550ms) [$\beta = 76.38$, $t_{29.61} = 3.625$, $p < 0.001$, 95% CI = (35.090, 117.674)] (953.87vs 1060.73 ms), and when the delay was of 450 ms [$\beta = 183.24$, $t_{29.61} = 8.697$, $p < 0.001$, 95% CI = (141.95, 224.53)] (877.49 vs 1060.73 ms). See Fig1b.

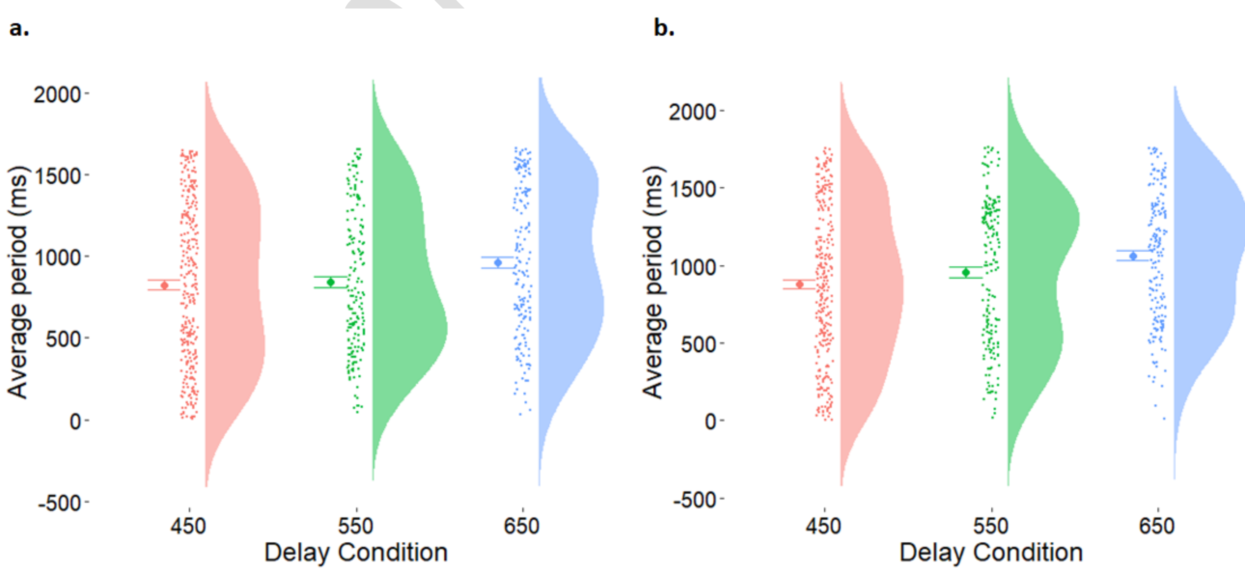

Figure 2 Average period in performing the melody across AI agents delays condition during the teaching phase, for the iCub (a) and Computer (b) partner.

SM 4 Manipulation Check

At the end of the experiment participants were asked to answer verbally two questions:

Q1: Did you notice the erring behaviour?

Q2: How would you describe the errors of the robot/computer?

The experimenter took notes of participants' replies and two independent raters categorized the answers as a function of the following dimensions:

Q1: Yes, No;

Q2: Intentional, Mechanical, n/a

Only replies on which both evaluators agreed were considered for the analysis. Of 59 respondents, 83.1% reported the occurrence of an error in the partner's performance (N = 49) while the remaining 16.9% did not notice any error in the performance of the partner (N = 10). The chi-square test indicated no difference in error detection between the type of agent (iCub vs Computer) or between the Erring behaviour condition (Human-like vs. Mechanical), [$\chi_2 = 0.679$, df = 3, p= 0.871], see Fig.2.

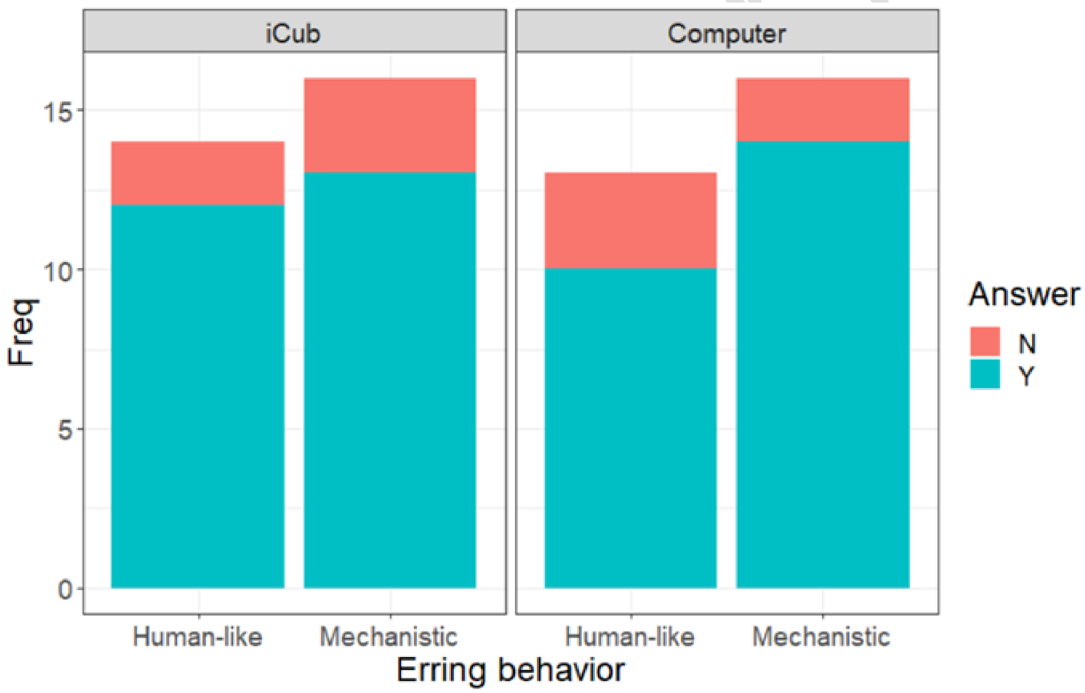

Figure *2*: Frequencies of responses to Q1 question plotted as a function of the Erring behaviour condition (Human-like vs. Mechanical) and the type of partner participants interacted with (iCub vs. Computer).

Out of the 49 participants who recognized the partner's errors, 32.7% described the errors referring to intentionality (N = 16), 40.8% described the error using mechanical or physical words (N = 20), and 26.5% were not able to describe the type of error (N=13). The chi-square test indicated no difference in how participants described the error between the type of agent (iCub vs. Computer) or Erring behaviour condition (Human-like vs. Mechanical), [$\chi_2 = 11.625$, df = 3, p= 0.071], see Fig.3. It is perhaps worthwhile to note that when the agent was the iCub robot, participants never explained the human-like error as "mechanical". Although this is only a qualitative indication, it might be an interesting point for future studies.

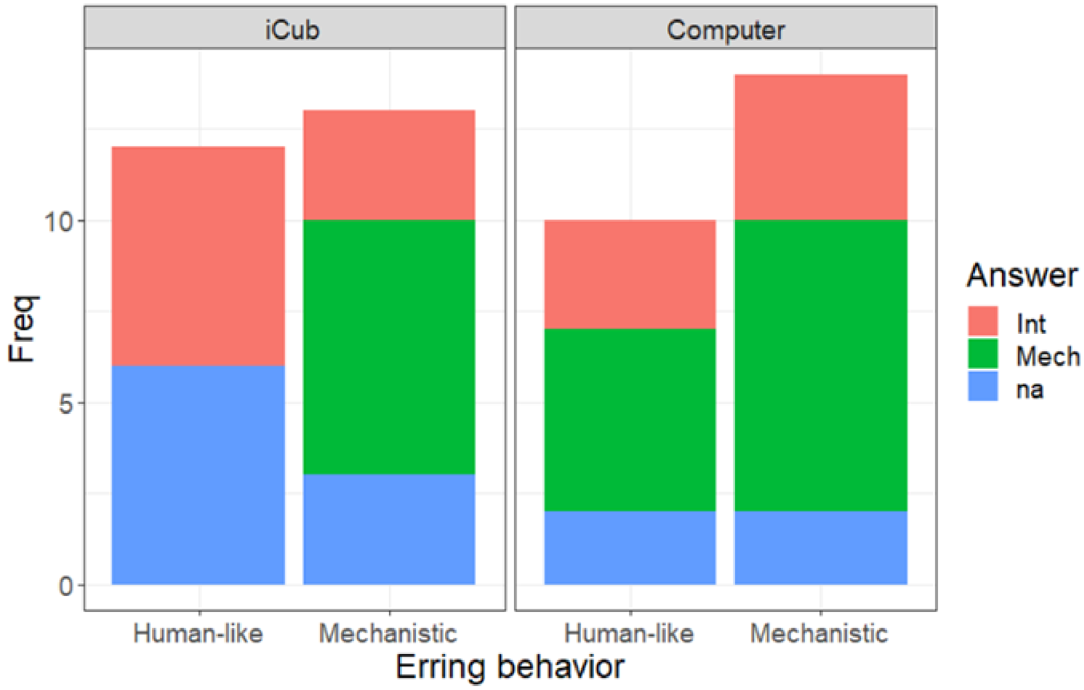

Figure 3: Frequencies of responses to Q2 question plotted as a function of the Erring behaviour condition (Human-like vs. Mechanical) and the type of partner participants interacted with (iCub vs. Computer).

References

[1] D.S. Syrdal, T. Nomura, & K. Dautenhahn. The Frankenstein Syndrome Questionnaire–Results from a quantitative cross-cultural survey. In: Proceedings of the International Conference on Social Robotics, 270–279, (2013).

[2] D.S. Syrdal., K. Dautenhahn, K.L. Kony, & M.L. Walters. The Negative Attitudes towards Robots Scale and Reactions to Robot Behaviour in a Live Human-Robot Interaction Study. Adaptive and Emergent Behaviour and Complex Systems (2009)

[3] C.M. Carpinella, A.B. Wyman, M.A. Perez, & S.J. Stroessner. S.J.: The Robotic Social Attributes Scale (RoSAS). In Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction, 254–262, (2017)

[4] Ciardo, F. et al. Effects of erring behaviour in a human-robot joint musical task on adopting Intentional Stance toward the iCub robot. *In 30th IEEE International Conference on Robot and Human Interactive Communication* (in press)

[5] D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823.(2014)

[6] A. Kuznetsova, P.B. Brockhoff, R.H.B. Christensen, Package ‘lmertest’. R package version, 2(0), (2015)