Check for updates

RESEARCH ARTICLE

Investigating the effect of cardio-visual synchrony on prosocial behavior towards a social robot

[version 1; peer review: 2 approved with reservations]

Katie A. Riddoch[1], Emily S. Cross[1-3]

[1]Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, UK
[2]Department of Cognitive Science, Macquarie University, Sydney, Australia
[3]MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Sydney, Australia

v1 **First published:** 21 Feb 2023, **3**:37
https://doi.org/10.12688/openreseurope.15003.1
Latest published: 21 Feb 2023, **3**:37
https://doi.org/10.12688/openreseurope.15003.1

Abstract

Background: Robots are being designed to alleviate the burden of social isolation and loneliness, particularly among older adults for whom these issues are more widespread. While good intentions underpin these developments, the reality is that many of these robots are abandoned within a short period of time. To encourage the longer-term use and utility of such robots, researchers are exploring ways to increase robot likeability and facilitate attachment. Results from experimental psychology suggest that interpersonal synchrony (the overlap of movement/sensation between two agents) increases the extent to which people like one another. **Methods:** To investigate the possibility that synchrony could facilitate people's liking towards a robot, we undertook a between-subjects experiment in which participants interacted with a robot programmed to illuminate at the same rate, or 20% slower, than their heart rate. To quantify the impact of cardio-visual synchrony on prosocial attitudes and behaviors toward this robot, participants completed self-report questionnaires, a gaze-cueing task, and were asked to strike the robot with a mallet.
Results: Contrary to pre-registered hypotheses, results revealed no differences in self-reported liking of the robot, gaze cueing effects, or the extent to which participants hesitated to hit the robot between the synchronous and asynchronous groups.
Conclusions: The quantitative data described above, as well as qualitative data collected in semi-structured interviews, provided rich insights into people's behaviours and thoughts when socially engaging with a humanoid social robot, and call into question the use of the broad "Likeability" measurement, and the appropriateness of the 'hesitance to hit' paradigm as a measure of attachment to a robotic system.

Keywords

social robotics, human–robot interaction, heart rate synchrony, prosocial behavior

Open Peer Review

Approval Status ? ?

	1	2
version 1 21 Feb 2023	? view	? view

1. **Alejandro Galvez-Pol**, University of the Balearic Islands, Palma de Mallorca, Spain

2. **Francesco Bossi**, University of Pisa, Pisa, Italy

Any reports and responses or comments on the article can be found at the end of the article.



This article is included in the Horizon 2020 gateway.



This article is included in the European Research Council (ERC) gateway.



This article is included in the Artificial Intelligence and the Social Sciences and Humanities collection.



This article is included in the Experimental Psychology and Cognitive Neuroscience collection.

Corresponding author: Emily S. Cross (emily.cross@glasgow.ac.uk)

Author roles: Riddoch KA: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Visualization, Writing – Original Draft Preparation; **Cross ES**: Conceptualization, Funding Acquisition, Methodology, Project Administration, Resources, Supervision, Validation, Visualization, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This research was financially supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 677270). The research was also supported by the Economic and Social Research Council (grant reference 1945868) and the Leverhulme Trust (PLP-2018-152).
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2023 Riddoch KA and Cross ES. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Riddoch KA and Cross ES. **Investigating the effect of cardio-visual synchrony on prosocial behavior towards a social robot [version 1; peer review: 2 approved with reservations]** Open Research Europe 2023, **3**:37 https://doi.org/10.12688/openreseurope.15003.1

First published: 21 Feb 2023, **3**:37 https://doi.org/10.12688/openreseurope.15003.1

Introduction

Rapidly ageing populations around the world, paired with an insufficient number of carers, has led to developments in "gerontechnologies" - devices designed to improve the health and wellbeing of older adults (Broekens *et al.*, 2009; Fozard *et al.*, 2000). Examples include location monitors, medication reminders, and fall detection systems (Bharucha *et al.*, 2009). It is proposed that gerontechnologies will allow people to retain their independence and remain in their own homes as they age, as opposed to entering into a care home facility (Benefield & Holtzclaw, 2014; Piau *et al.*, 2014). Although so-called "aging in place" is a desire reported by many people (Rantz *et al.*, 2005), living alone is also associated with social isolation (lack of contact with others) and loneliness (distress that results from the discrepancy between one's desired and actual social relationships; Pinquart & Sörenson, 2003). Longitudinal studies on older adults have demonstrated that feelings of loneliness are linked to more symptoms of depression (Cacioppo *et al.*, 2006), reduced physical activity (Hawkley *et al.*, 2009), and impaired cognition (O'Luanaigh *et al.*, 2012). A study of 1604 older adults also found that those who reported being lonely on average experienced greater difficulty with walking, stair climbing, and completing activities of everyday living (e.g. bathing, dressing, etc.) compared to those who did not report being lonely (Perissinotto *et al.* (2012).

A role for robots in reducing loneliness

In an attempt to reduce loneliness and the myriad of associated health problems, companies are in the process of developing 'companion robots' - machines designed to be engaging, comforting, and respond to the user in an intuitive manner (Broekens *et al.*, 2009; Young *et al.*, 2009). Rather than having the capabilities of assistive robots designed to carry out physical tasks, such as carrying food or fetching medication, companion robots are designed to connect with users in a socio-emotional way. The need for such companion robots to be developed has been further reinforced by social distancing measures introduced as part of the recent COVID-19 pandemic, with prominent roboticists championing a role for social robots as ideal tools for providing care and companionship when contact with other people brings increased infection risk (Yang *et al.*, 2020).

One such companion robot is "Paro" – a robotic seal developed within the Japan National Institute of Advanced Industrial Science and Technology. Equipped with microphones and tactile sensors, the Paro robot can move and vocalise in response to a user's voice or touch. Despite its simplicity, studies demonstrate that individuals (specifically, older adults with dementia) enjoy interacting with Paro, and that when the robot is present (compared to not), care home residents engage in more conversation with staff and other residents (Kelly *et al.*, 2021; Takayanagi *et al.*, 2014). The latter finding suggests that a social robot could act as a social facilitator (encouraging interactions between humans), opposed to something which leads to increased isolation (e.g., if the person engages with a robot in lieu of other humans).

A number of research groups have suggested that after novelty effects wear off, social robots will be neglected and users will fail to reap long-term benefit (e.g., Leite *et al.*, 2013; Woo *et al.*, 2021). It is pertinent to note however, that such suggestions are predominantly based on research with children - often within classroom settings. With regards to social robots for older adults, current research findings are much more promising. For example, Bradwell and colleagues (2020) conducted a 6-month diary study in a supported living facility where residents had the opportunity to interact with a robotic cat, and found that older adults engaged with the robotic cat increasingly often - shifting from short, structured sessions, to more frequent requests. By the end of the study, staff in the facility reported that the robot was "continually present" and that an estimated "80% of clients loved the cat". Such findings are promising in terms of robot acceptance in support living facilities, however, it remains unclear the extent to which these findings might apply to older adults living independently, or other types of robots (e.g., those inspired by humans, opposed to pet animals). It could be the case that like a third of assistive technologies, social robots as well are at risk of being abandoned within the first three months of use (Gurley & Norcio, 2009).

For robots to have a long-term positive impact, it is vital to conduct further research with independent older adults, and to develop a clearer understanding of the features or behaviors that might facilitate bonding and attachment to robotic systems. As robots are often ascribed intentions and treated like social entities by people of all ages in a variety of laboratory and naturalistic experiments (Hortensius & Cross, 2018; Hortensius *et al.*, 2018; Wykowska *et al.*, 2016), research from psychology and the cognitive sciences has the potential to play a significant role in characterising which factors and attributes of robots will lead to the long-term acceptance and enjoyment by human users. One of the factors currently receiving considerable research attention for its potential to facilitate stronger bonds between humans and robots is interpersonal synchrony (Henschel & Cross, 2020; Lehmann *et al.*, 2015; Mörtl *et al.*, 2014).

Interpersonal synchrony

Interpersonal synchrony refers to the overlap of movement and/or sensation in time or form (for example, when we tap in sync, compared to out of sync, with another person) (Hove & Risen, 2009). Studies have demonstrated that experimentally-induced movement synchrony can have significant positive effects on prosocial behavior towards 1) other people (increased donating in public goods games and improved rapport) and 2) social robots (increased liking and perceived intelligence; Hove & Risen, 2009; Lehmann *et al.*, 2015; Mörtl *et al.*, 2014; Mogan *et al.*, 2017; Rennung & Göritz, 2016). These studies raise the intriguing possibility that movement synchrony could be used to facilitate increased liking of social robots. In the present study, however, we shift our focus to another kind of interpersonal synchrony that has the potential to be introduced in a more subtle and effortless manner: namely, cardio-visual synchrony.

Cardio-visual synchrony refers to an overlap between a visual stimulus and an observer's heart rate – for example, a light bulb flashing at the same speed as one's heart rate. Lab-based studies have identified that cardio-visual synchrony can impact the extent to which a person perceives an object as being part of their own body, or as part of the self. Specifically, when viewing body parts (e.g., an image their own face, or a rubber hand in place of their own), participants feel greater self-identification to those which are illuminating in a synchronous manner relative to their heart rate (Sel *et al.*, 2017; Suzuki *et al.*, 2013). Demonstrating that cardio-visual synchrony can lead to an increased self-other overlap towards an object could be a significant finding in terms of facilitating liking and prosocial behaviors towards a robot, as Hove & Risen (2009) proposes that positive effects of interpersonal synchrony (such as prosocial behaviors and increased liking) are the result of a perceived blurring of the "self" and "other" (Hove & Risen, 2009).

If found to be effective, a cardio-visual synchrony intervention could be relatively easy to implement in robotic systems used in social contexts with human users as: 1) it involves minimal programming (unlike movement synchrony), 2) it could be facilitated by an inexpensive commercially available heart rate monitor, and 3) it could operate independently of the other behaviors of the person, and importantly, the robot - that is, the movement, speech, and other existing functions of the robot are unaffected by the presence of the flashing lights.

Current study

To investigate the effect of cardio-visual synchrony on the perception of a robotic agent, we monitored the heart rate of the participant as they interacted with a humanoid robot (using a wrist-based heart monitor), and relayed this information to the robotic system in real time. As a result, the robot's shoulder lights illuminated in a manner synchronous (at the same rate) or asynchronous (20% slower) relative to the participants heart rate. To determine whether cardio-visual synchrony leads participants to perceive a Pepper robot as more likable and behave in a more prosocial manner towards it, we used a between-subjects design and several qualitative and behavioral measures designed to probe participants' awareness of and response to our experimental manipulation.

The current study was designed to address three primary pre-registered predictions. Given the literature suggesting that interpersonal synchrony increases likeability (see 'Interpersonal synchrony'), we predict that the synchronous group, compared to the asynchronous group, will perceive the robot as more likable. We hypothesise that this difference will be reflected in the ratings on the validated 'Liking' scale of the Godspeed questionnaire - with the synchronous group scoring higher than the asynchronous group.

We propose that the increased liking will also be reflected in how long participants hesitate after being asked to hit the robot. Specifically, we predict that the synchronous group, compared to the asynchronous group, will hesitate for longer after being asked to hit the robot with a mallet. We also predict that more individuals in the synchronous group, compared to the asynchronous group, will refuse to hit the robot.

To investigate whether perceptions of the robot as a social agent differ between the two groups, we also explore gaze cueing data. On the basis of experiments suggesting that mind-perception modulates gaze cueing effects (Morgan *et al.*, 2018; Teufel *et al.*, 2010; Teufel *et al.*, 2009; Wiese *et al.*, 2012), we anticipate that participants will be slower to respond to human faces, compared to the arrows, replicating the basic gaze cueing effect. We also expect that participants in the synchronous group will exhibit slower reaction times to the robot stimuli, compared to the asynchronous group. This is on the basis of research suggesting that a synchronous agent is perceived as more "like me" and as a social being opposed to an object as a result (Sel *et al.*, 2017; Suzuki *et al.*, 2013).

We wish to note that, during the course of running the present study, it became clear that the debriefing procedures we had planned were yielding incredibly rich data and insights about participants' subjective experience of socially engaging with (and being asked to damage) the robot. In response to the richness of these insights and our wish to delve into more qualitative aspects of participants' experiences in this study, we have made the qualitative data collected as part of this project the focus of an additional, exploratory paper, which was recently published (Riddoch & Cross, 2021). We wish to make clear that the same participants and data sets form the foundation of the present manuscript and the qualitative paper (Riddoch & Cross, 2021), with our focus in the present manuscript on the pre-registered quantitative questions, analyses and implications.

Methods

Preregistration and data

Prior to data collection, all manipulations, measures, and the sample size justification and main hypotheses were pre-registered on the Open Science Framework (Riddoch & Cross, 2020). Consistent with recent proposals (Simmons *et al.*, 2011; Simmons *et al.*, 2012), we report all manipulations and all measures in the study. In addition, following open science initiatives (Munafò *et al.*, 2017), the data, stimuli, and analysis code associated with this study are freely available on the Open Science Framework. By making the data available, we enable others to pursue tests of alternative hypotheses, as well as more exploratory analyses.

Ethics and consent

All study procedures were approved by the College of Science and Engineering Ethics Committee (University of Glasgow, Scotland) on 19th June 2019 – approval number 300180265. All participants provided written informed consent prior to taking part in the study.

Sample size justification

Due to the time and resources associated with recruiting a sample including older individuals, the decision was made to use Bayesian Sequential Hypothesis Testing as outlined

by Best, Barsalou and Papies (2018). To determine our minimum and maximum sample size, we undertook two power analyses in G*Power3.1. Both were undertaken on the basis of an independent T-Test, comparing scores on the Likeability scale of the Godspeed questionnaire (Bartneck *et al.*, 2009a).

To find a large effect (Cohen's d = 0.8), G*Power3.1 indicated that using a power of 0.8 and an alpha level of 0.05 (5%), we would need a minimum of 42 participants (21/group). To find a medium effect (Cohen's d = 0.5) G*Power3.1 indicated that using a power of 0.8 and an alpha level of 0.05 (5%), we would need to test 102 participants (51/group). Our maximum sample size was based on a medium effect size, as a small effect size might not be of as much interest commercially (at least initially). That is, if we found a small effect, we would argue that the synchrony setup might not be a compelling area of development for those designing and producing robots.

As outlined in our preregistration, lodged on the Open Science Framework (Riddoch & Cross, 2020), we initially tested 42 participants (the minimum sample size), then calculated the updated Bayes Factor (BF) after every 4 participants. As the BF was below 6 (considered "strong" evidence that the alternate hypothesis is true, opposed to null; Schönbrodt & Wagenmakers, 2018) we continued recruiting and testing until n=89. Testing was halted at 89 participants opposed to 102 (the maximum sample size we pre-registered), due to time constraints and difficulties recruiting individuals over the age of 60.

Participants

89 individuals took part in the experiment, however the data from 12 individuals were excluded as they encountered problems which affected their experience with Pepper (error lights within Pepper, loss of Bluetooth/WIFI connection, and hearing problems). As a result, the final sample included 77 individuals aged 18-83 (mean age = 43.36, SD = 21.38), with 31 individuals over the age of 60 ("older adults"). Participants were individuals residing in Glasgow (Scotland, UK) and were initially recruited by word of mouth (in person, such as friends, family members, and colleagues of the researchers; via email, such as sharing via approved university mailing lists for participant recruitment; and through social media advertisements on Twitter) followed by snowball sampling. All individuals had normal or corrected-to-normal vision and hearing, and no previous experience interacting with the robot used in the study. All individuals were compensated £10 for their participation.

Participants were assigned to the synchronous (n=40) or asynchronous (n = 37) group prior to arrival, using a random group generator online. Such generators allow the experimenter to specify how many participants will be tested (e.g., n=77), and how many groups the individuals should be split into (in this case, 2 groups). The generator then splits the individuals randomly, into the specified number of groups. Ideally, to avoid inadvertently biasing the results, the experimenter would have been blind to the condition of the participant. However, this was not possible in the present study due to the logistical complexity paired with lack of resourcing. In the future, if repeating similar work, we would recommend that researchers consider how the experimenter can be blind to the condition of the participant, without adding the social pressure associated with an extra researcher in the testing space.

To check that the groups were sufficiently matched, in terms of their demographics, we performed a series of statistical tests. Results of the independent samples t-tests revealed no significant differences between the groups regarding age, nor on questionnaires assessing negative attitudes towards robots (Nomura *et al.*, 2008). Such tests also revealed no significant difference between the groups in terms of anthropomorphic tendencies (Waytz *et al.*, 2010), and general empathetic concern (Batchelder *et al.*, 2017). See Figure 1 for a visualisation of between-group similarities, in terms of their questionnaire scores.

Both synchronous and asynchronous groups contained more women than men (57.50% and 62.16%, respectively). In the synchronous group, one individual identified as "Agender", and in the asynchronous group, one person identified as "Non-Binary".

Experiment design

Participants assigned to the synchronous group interacted with a humanoid robot (see 'Apparatus' for more details) whose shoulder lights were programmed to illuminate at the same rate as their heartbeat. This was achieved by sampling the participant's heart rate using a wearable heart rate monitor and relaying this information to the robot in real time. In the asynchronous group, the heart rate data were sampled in the same way, however the lights of the robot were programmed to flash at a rate 20% slower than the individual's heart rate (thus producing an asynchronous cardio-visual experience).

Regardless of group allocation, participants completed the same tasks with the robot, and the same measures. To account for between-group differences, baseline measures were taken from both groups before they interacted with the robot.

Apparatus

Robotic platform. The robot used in the experiment was the Pepper Robotic system - a commercially available humanoid robot from SoftBank Robotics (Tokyo, Japan). See Figure 2 for an image of the Pepper robotic system. Pepper is 120cm tall and features 2 in-built cameras, as well as microphones and tactile sensors, which allow it to detect objects and movement in the environment. Pepper is already being introduced to social spheres, and is already being trialled in hospital and service industry contexts (Foster *et al.*, 2016; Niemelä *et al.*, 2017; Tanioka, 2019).

Pepper also has expressive movement and speech capabilities that can run autonomously, but for the purpose of experimental control we used a Wizard of Oz set up wherein we controlled

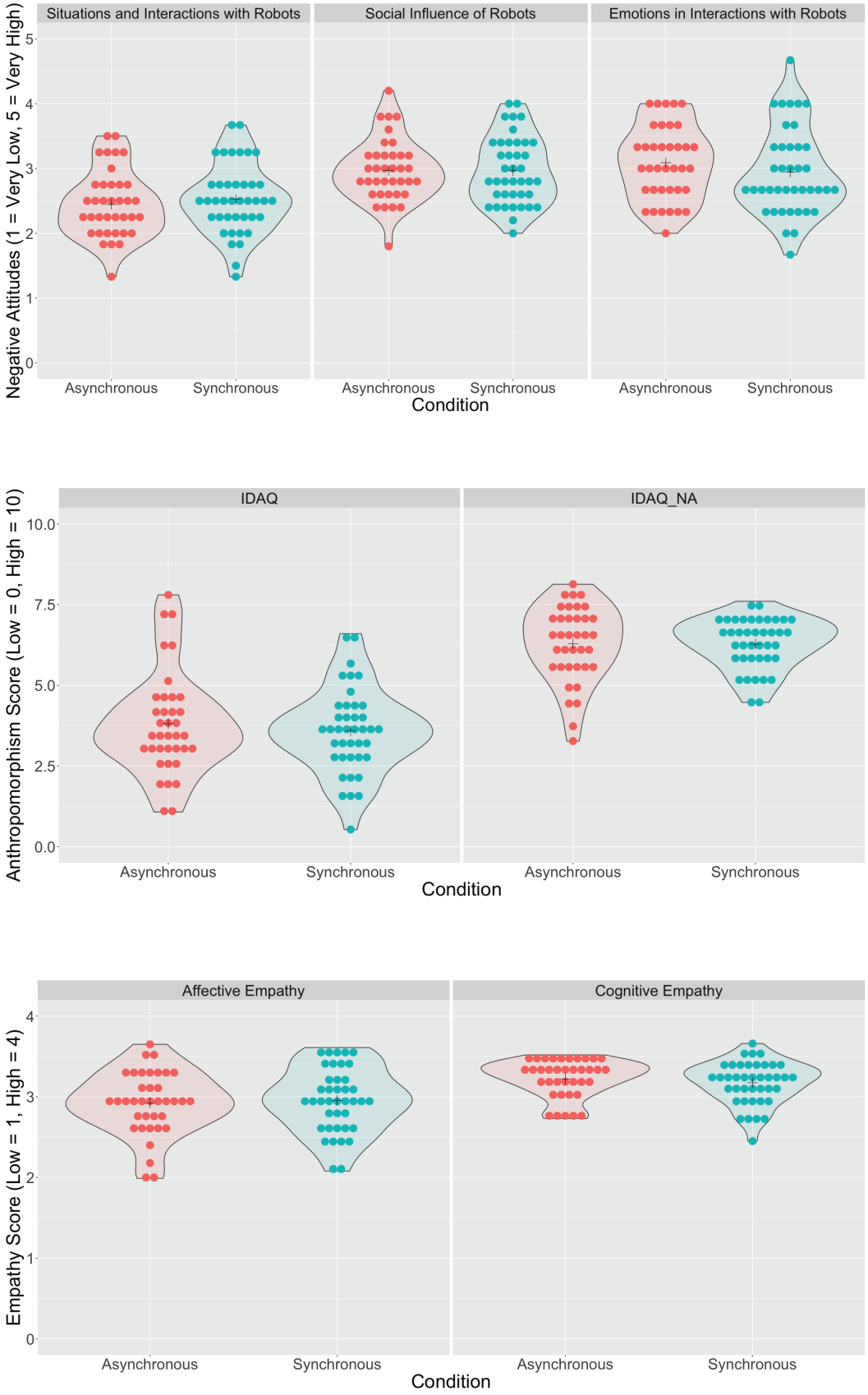


Figure 1. Plots illustrating the two groups' distributions regarding negative attitudes towards robots (top), anthropomorphic tendencies (middle), and general empathetic concern (bottom).

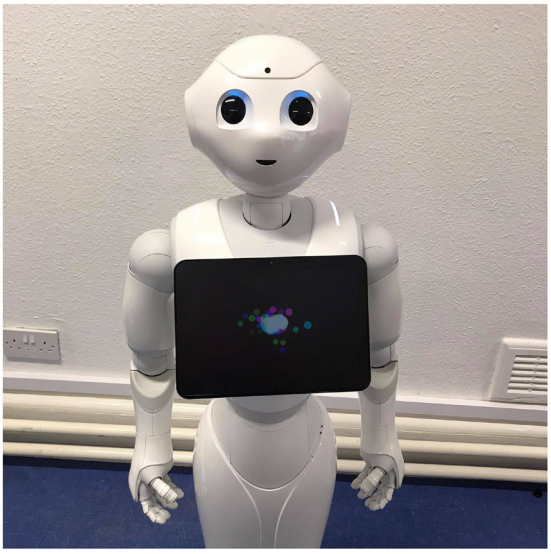


Figure 2. The Pepper robotic system.

Pepper's behavior remotely. Specifically, for one part of the experiment we triggered a sequence of speech and movements via the 'Choreograph' software (see 'Procedure' for further details). In another section we created a panel of key phrases using html. Upon clicking a speech button, a corresponding line of Python code is triggered and Pepper speaks and moves accordingly. To maintain experimental control, for each interaction the experimenter systematically clicked from the first phrase "Hi there", through a series of closed questions and responses, to the final phrase "Thank you". Closed questions (e.g. do you prefer tea or coffee?) and option-specific responses were used to create the illusion that the robot was responding to the specific words of the participant.

In addition to controlling the speech and movement of the robot, it is also possible to co-opt the lights in the shoulder panels. Specifically, rather than flashing to indicate the 'mode' or 'state' of Pepper, we programmed the lights to illuminate either synchronously or asynchronously with participants' heart rate, depending on group assignment.

Heart rate monitor. To capture heart rate information, we used a Polar OH1 optical heart rate sensor. The Polar OH1 heart rate sensor was chosen due to its reliability, sport-focussed design (allowing for freedom of movement without diminished accuracy), Bluetooth Low Energy (BLE) capabilities (necessary for relaying the heart rate information to external devices), and ability to be worn on the wrist (as opposed to uncomfortable or obstructive devices such as finger clips or chest straps). The data collected by the heart rate monitor were relayed via a laptop to the robot, allowing the robot's shoulder lights to illuminate at the same rate as the participant's heart rate ('synchronous'), or 20% slower ('asynchronous').

Each second, the PolarOH sensor sampled the participant's pulse rate (indicative of the number of times the heart beats in a minute). Using the BLE capabilities of the PolarOH sensor, we used a laptop to extract the pulse rate data in real time. The pulse rate data were then processed using Python code – allowing us to dictate whether the lights of the robot should illuminate at the same rate as the pulse (synchronous group) or 20% slower (asynchronous group). After processing the pulse data, this information was relayed to the robot in real time, via WiFi.

We contemplated using random or stable rates as the control for heartrate synchrony, but ultimately decided against this as we wished to ensure that the effect was the result of the synchrony manipulation *per se* – and not other factors such as effects of dynamic vs stable pulsing. We also considered binding participants' heartrates to other aspects of the robot's behavior (such as breathing movements or gestures), however we decided against this in order to study to impact of the illuminating lights alone.

To avoid exposing the aim of the experiment, the experimenter made no mention of the robot's shoulder lights, nor their illuminating nature. We also made the decision to use gradual illumination, opposed to spikes, to create a subtle effect – opposed to an obvious and distracting light manipulation. During the human-robot interaction however, two participants did enquire why the shoulder lights of the robot were flashing. To this, the experimenter briefly commented that the lights were simply a visual indication that the robot was on and functioning correctly. Both participants appeared to accept the cover story, and continued to ask questions about different features of the robot. At the end of the experiment, each participant was asked if they had any thoughts or feelings about the illuminating lights on the robot. No participants suspected a link between the lights on the robot and their heart rate.

Measures

Questionnaires. To probe liking, participants completed the validated Liking scale of the Godspeed Questionnaire (Bartneck *et al.*, 2009a) before and after interacting with the robot. The questionnaire has received some criticism due to the overlapping nature of the "Anthropomorphism" and "Animacy" scales (Carpinella *et al.*, 2017), however the high internal validity of the likeability scale specifically (Bartneck *et al.*, 2007a; Bartneck *et al.*, 2009b), paired with the benefits of the semantic differential format (Friborg *et al.*, 2006), motivated our decision to use the scale in this experiment.

To explore if and how participants' perception of the robot is affected, we also administered questionnaires probing the extent to which participants perceived the robot as a social agent; specifically, the Inclusion of Self in Other task (Aron *et al.*, 1992) and the Robotic Social Attributes Scale (Carpinella *et al.*, 2017). To check for between-group differences in attitudes towards robots, participants also completed questionnaires to probe their history with robots via the Exposure to Cinematic Depictions of Robots (Riek *et al.*, 2011), and Negative Attitudes towards Robots (Nomura *et al.*, 2008) questionnaires. To

account for between-group differences with regards to anthropomorphic tendencies and general empathy, the Individual Differences in Anthropomorphism Questionnaire (Waytz *et al.*, 2010), and the Empathy Components Questionnaire (Batchelder *et al.*, 2017) were also administered. Participants were given the option of completing the questionnaires on paper, or via the online questionnaire platform form{'r}.

Gaze cueing. To investigate the extent to which participants perceive the Pepper robot as a social agent, opposed to a mindless object, a computer-based gaze cueing paradigm (Driver *et al.*, 1999; Friesen & Kingstone, 1998) was administered before and after interacting with the robot. This task operates under the assumption that when we believe an agent has knowledge and is behaving intentionally, our attention is directed by their eye gaze and we exhibit slower reaction times (Friesen & Kingstone, 1998). Such 'gaze cueing effects' have been demonstrated in studies using robots – with participants being more distracted and slower to respond when the robot is thought to be controlled by a human (Morgan *et al.*, 2018; Wiese *et al.*, 2012). By comparing participants' reaction times in response to images of robots, compared to humans or other objects, we should be able to determine the extent to which a person perceives the robot as a mere object, or an intentional agent like a human. We can then compare between the asynchronous and synchronous groups to determine the extent to which the two groups differ in their perceptions, before and after interacting with the robot.

In the task participants saw images of an arrow, or the faces of Pepper, a different robot, or a human (dimensions: approximately 600mm x 800mm). The robot images are original - taken in the Social Brain in Action Laboratory. The face of the human was chosen from the Karolinska Emotional faces database (Lundqvist *et al.*, 1998), on the basis of its resemblance to the robot in both form and contrast. All images were changed to black and white, to control for potential influences of colour on attention. To retain ecological validity, the decision was made not to control for contrast and composition between the agent types.

Classically, gaze cueing paradigms are conducted with arrows or eyes, however the limited eye movements of the robots forced us to adopt a whole-head shift in direction. Studies have demonstrated that gaze cueing effects are still present when stimuli depict the whole head of the agent (Frischen *et al.*, 2007; Langton & Bruce, 2000), further validating this decision. The task was created using the PsychoPy3 experiment builder and was presented to participants on a 21.5-inch iMac desktop computer.

In the gaze cueing task, participants see a fixation cross, followed by a front-facing image. See 'Neutral' images on Figure 3. They then see said agent orientating in the left or right direction. See 'Direction Cueing' on the same figure. The participant then sees a target (in this case, an asterix) appear congruent, incongruent, or neutral relative to the face. The sequence of

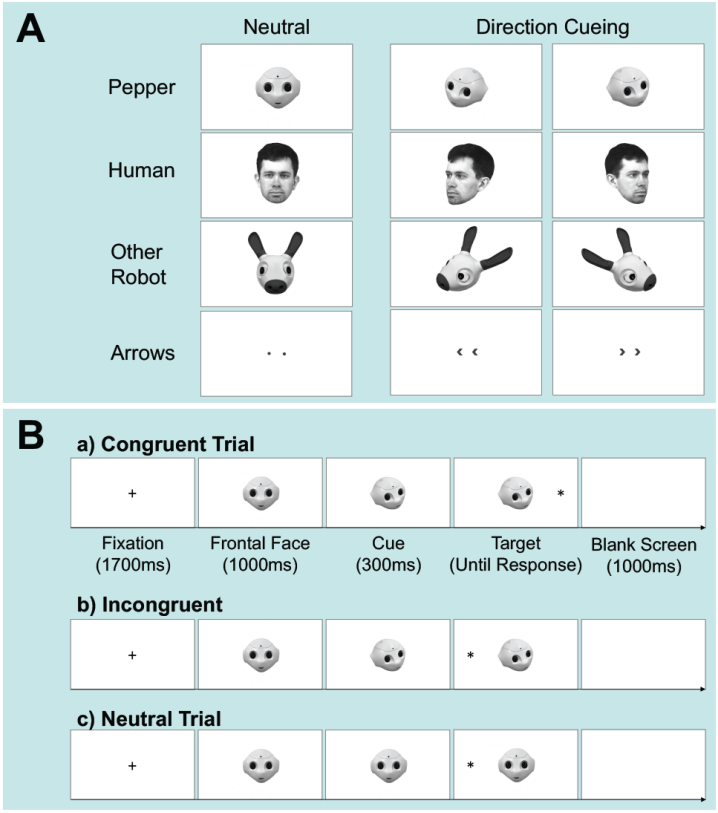


Figure 3. Neutral and Direction Cueing stimuli (**A**) and stimuli presentation timings and examples (**B**).

images was presented in quick succession, leading to the illusion of apparent motion of the head/face. Refer to 'Introduction' for further details, see Figure 3 for illustration of trial types and timings. The participants are instructed to focus on the fixation cross, then respond "as quickly and accurately as possible" upon seeing the 'target' appear. The participant responds to the left or right targets by pressing the corresponding arrow key on the keyboard (left, or right, respectively).

Responses were removed if slower than 150ms ('preemptive') or above 2500ms ('unusually slow'). This is in line with previous work scrutinising attention cueing differences in young and older adults (Gayzur *et al.*, 2014) and pieces focussed on the 'effective analysis of reaction time data' (Whelan, 2008). These criteria led to the exclusion of approximately 0.26% of the dataset (35 data points).

Hesitance to hit. To probe attachment towards the robot, we used a modification of the "hesitance to hit" task (Bartneck *et al.*, 2005; Darling *et al.*, 2015) - asking participants to strike the robot on the head with a mallet. The time between being given the instruction to hit and agreeing to do so (indicating intention to hit) was measured and compared between the asynchronous and synchronous groups. To determine why participants hesitated, and better understand what hesitance to hit reflects (a gap in previous research) we administered a semi-structured interview at the end of the study, asking participants why they hesitated and what they were thinking and feeling during the task.

Note: in previous experiments, relatively simple robots (e.g., inexpensive robotic bugs) were used – allowing participants to actually hit and break the robot (Bartneck *et al.*, 2005; Darling *et al.*, 2015). In this experiment however, breaking the robot was unjustifiable due to the costs associated (both financially, and in terms of physical waste). We considered other platforms (e.g., animal-like, mechanical-looking…), however, significant costs and waste were associated with breaking any commercially available social robot. As a result, the paradigm was adapted to measure participants' *intention* to hit a robot, as opposed to actually hitting the robot. Different intention signals were considered (e.g., a button press, a verbal command…) however we decided that standing up (to walk towards the robot) was the best measure, due to the physical effort required.

To better understand what the hesitance reflects, and validate the technique as a measure of social attribution, we interviewed participants regarding why they hesitated as part of an extensive debriefing procedure. The results from this are described in Riddoch & Cross (2021).

Procedure

Upon entering the room (see Figure 4) participants took a seat in the control space, in front of the iMac (depicted in blue). After participants provided written informed consent, they undertook the computer-based gaze-cueing paradigm, then saw a video of Pepper (Tech Insider, 2018) and were asked to complete questionnaires probing how much they liked the robot (see the previous Questionnaires section, for details). During this time the experimenter turned on the robot and sat in the testing space. Upon completion of the aforementioned tasks, the experimenter returned to the control space and asked the participant to don the wrist-based heart rate monitor. As a cover story, the participant was told that the device encourages the robot to focus on them (as opposed to the experimenter). No participants asked for clarification regarding the brief cover story. The robot was triggered to illuminate in a synchronous or asynchronous manner (group-dependent). The participant was then taken to the testing space and was seated in front of the robot. A small video recording device (Logitech webcam) mounted on the room separator was triggered, which allowed the experimenter to monitor and record the human-robot interaction. Participants were not made aware of the webcam recording at this point – to avoid potential influence on their behaviour.

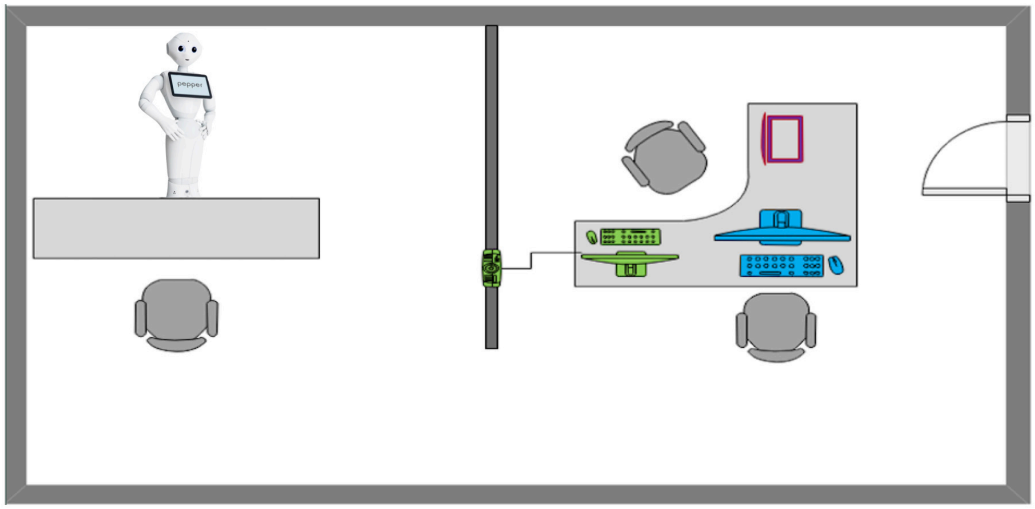


Figure 4. Room setup illustrating a room divider separating the testing space (left) and control space (right).

Participants drew and made notes about the robot (5 minutes), then observed as the robot performed the 'Tai Chi', 'Vacuuming', and 'Disco Dancing' movements (2 minutes). Participants were then informed that Pepper would ask them about their food preferences, and they would create a shopping list together. The experimenter used the excuse "I'll get out of the way so that Pepper doesn't try to talk to me as well" and returned to the control space to control Pepper (as described in 'Robotic Platform'). The participant was asked a series of questions about their food preferences, and what they would like on a food shopping list. This sequence of events was designed to 1) reflect a typical first interaction with a social robot (observation and evaluation, followed by a two-way interaction), and 2) to subtly encourage attention towards the synchronous/asynchronous shoulder lights.

After repeating the liking questionnaires and gaze cueing task in the control area, participants were invited to take a seat in front of the robot and don a pair of safety goggles. The experimenter then stood next to the robot and proceeded to read the following script:

> *"Right, there is something I haven't told you about this experiment. This Pepper is one of ten specially designed robots that I was given as part of a large research grant. By 'specially designed' I mean that they're totally shatterproof – so if you hit one, the robot will break in a safe way that's easy to repair. The reason Pepper is designed this way is because our lab is interested in what happens when someone has to hit a robot – for example, if a robot was to malfunction and you had to hit and disable one. Does that make sense? Great." *Experimenter passes participant the hammer*. "So, for this part, your task is to give the robot one hard hit on the head. So, when you're ready, come round the table and I'll get out of the way."*

After standing up to hit the robot (indicating the intention to hit the robot), the participant is told to pass the hammer to the experimenter, and is informed that they will not actually be hitting the robot. The participant is prompted to remove their safety goggles and is invited back to the control area for a task debriefing. If the participant verbally protests against hitting the robot, they are told "It's just part of the experiment". Upon protesting three times, the task is ended as indicated previously, and the participant is deemed to have 'refused' to hit the robot.

Upon completion of the task (either by agreeing or refusing to hit the robot) the participant is asked a series of open questions (e.g. "After I asked you to hit the Pepper, what was going through your mind?") to probe their thoughts and feelings during the hesitance to hit task. For the full list of questions see extended data. To conclude the experiment, participants completed demographic and personality questionnaires (see Questionnaires section for details), and then were debriefed and compensated for their time. The debriefing included: 1) a statement that the robot was under the experimenter's control, 2) informing the participant that webcam recording had been taken for the purpose of validating the length of hesitation, and 3) signposting to key contact details should they have questions or concerns following the study.

Data collection and analysis software. The computer-based Gaze Cueing task was created using the PsychoPy Builder Interface (Peirce *et al.*, 2022). The Qualtrics survey platform (Qualtrics, 2005) was used to collect questionnaire responses. For those unable to use a computer, a paper version of the questionnaires was provided (see "QuestionnaireBooklet_Paper" file on the study OSF). To conduct statistical testing on the quantitative data we used Jamovi software (Jamovi, 2021). To manage and code the qualitative data we utilised NVivo11 (NVivo11, 2015). The data analysis we conducted (see Results section for details) aligned to the methods preregistered on the Open Science Framework (Riddoch & Cross, 2020).

Results

This study was designed to evaluate three main hypotheses relating to the impact of cardio-visual synchrony between a human and robot. Specifically, we predicted that:

1. after interacting with the robot, the synchronous group will rate the robot higher on the 'Liking' scale of the Godspeed questionnaire, compared to the asynchronous group.

2. the synchronous group, compared to the asynchronous group, will hesitate for longer after being asked to hit the robot with a mallet.

3. more individuals in the synchronous group, compared to the asynchronous group, will refuse to hit the robot.

In the following sections, each hypothesis is addressed in turn.

Hypothesis 1: Liking

To determine whether the robot was rated as more likable, depending on whether it illuminated in a manner that was synchronous or asynchronous with a participant's heart rate, scores on the 'Likeability' scale of the Godspeed questionnaire were considered. As each of the 5 items in the Likeability scale is rated from 1–5, the minimum score that could be given is 5 and the maximum is 25. 25 indicates high scores on all dimensions – Dislike-Like, Unkind-Kind, Unfriendly-Friendly, Unpleasant-Pleasant, and Awkward-Nice.

Before the interaction, descriptive statistics indicated little difference between the scores of the synchronous (M=20.61, SD=3.89) and asynchronous group (M=20.39, SD=3.65). After the interaction, both groups rated the robot slightly higher on the likeability scale, however little difference between the synchronous (M=21.76, SD=4.00) and asynchronous (M=21.14, SD=3.76) groups was found (see Figure 5 for visualisation).

The results of a 2 x 2 ANOVA indicated a Bayes Factor of 0.215 when considering the difference between the synchrony

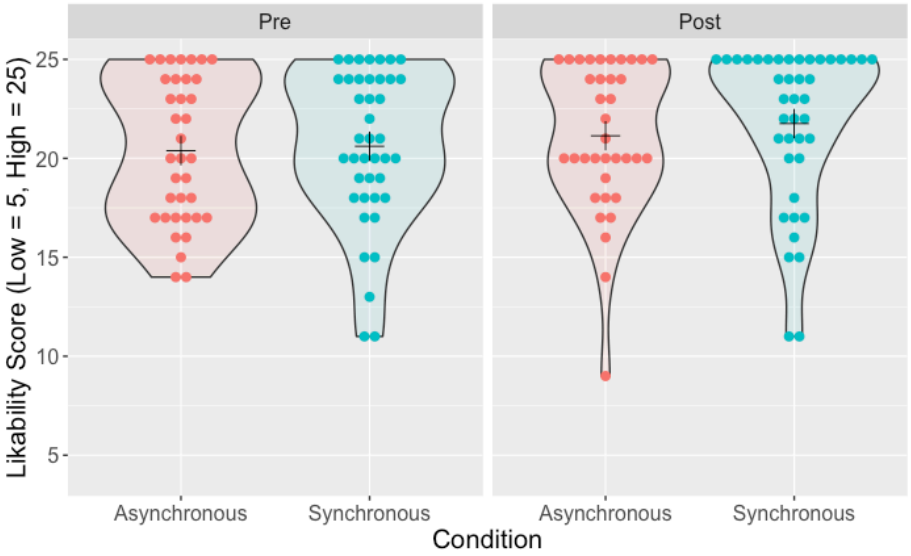


Figure 5. Ratings on the Likeability Scale. Each point represents the score of an individual participant. The crosshair represents the mean of the participants in that condition, at the specified time point.

groups. The analyses also indicated a Bayes Factor of 0.532 for the effect of Time (Before vs After interacting with the robot). These Bayes factor values are substantially lower than 6 – the value proposed to indicate "strong" evidence that the alternate hypothesis is true (Best *et al.*, 2018; Schönbrodt & Wagenmakers, 2018). This indicates that neither condition, nor time, changed the extent to which participants liked the robot in our particular manipulation.

Hypothesis 2: Hesitance to hit

Compared to the asynchronous group, we hypothesised the synchronous group would hesitate for longer between being asked to hit the robot with a mallet, and agreeing to do so. To account for individuals who refused to hit the robot, we adopted the method used by Darling *et al.*, (2015) - adding 1 second onto the maximum measured hesitation. This method was originally justified as follows: "if a subject did not strike the robot, we considered this to be greater than the maximum measured hesitation" (Darling *et al.*, 2015). This led to three values of 101 in the dataset. For clarity, these values were *not* plotted in Figure 6, however they *were* included in the descriptive statistics and statistical tests. The 'refusals' will be discussed more in the later Results section 'Hypothesis 3: Refusal to hit'.

Descriptive statistics indicated that the asynchronous group (M = 18.50s, SD = 28.92) hesitated for longer, on average, than the synchronous (M = 14.75 seconds, SD = 22.14) group. See Figure 6 for visualisation.

To compare statistical significance of these results we had intended to perform an independent samples T-Test, however the assumptions of normality and equal variances were violated (Shapiro-Wilkov and Levene's p<.05). This led to the use of a two-tailed Mann-Whitney U test instead. This test indicated no significant differences between hesitation time of the two, U = 698.0, p = 0.864.

Hypothesis 3: Refusal to hit

The final hypothesis states that more individuals in the synchronous group, compared to the asynchronous group, will refuse to hit the robot. A reminder: a 'refusal' results from a participant protesting (e.g., "I don't want to...", "Do I have to...?") three times. The results of a Mann-Whitney U test indicated no significant difference between the number of protests made by the synchronous group (M = 0.52, SD = 0.68), compared to the asynchronous group (M = 0.60, SD = 1.14), t(39) = 262.50, p=0.917. A Mann Whitney U test, opposed to an Independent T Test, was used because a Levene's test indicated inequality of variances between the groups (p < .05). To visually compare the number of protests made by participants, split by condition (synchronous and asynchronous), refer to the chart in Figure 7.

Exploratory analyses

Self-Other overlap. To better understand the pattern of null effects, we explored results of the Self-Other questionnaire (Aron *et al.*, 1992) – due to prior evidence highlighting the relevance of self-other identification for reaping the prosocial benefits of synchrony (Hove & Risen, 2009; Sel *et al.*, 2017; Suzuki *et al.*, 2013). Important to note; participants used a series of overlapping circles to indicate how they closely perceived themselves as overlapping with the robot - with 1 (two separate circles) indicating zero overlap, and 7 (two fully overlapping circles) indicating high self-identification with the robot.

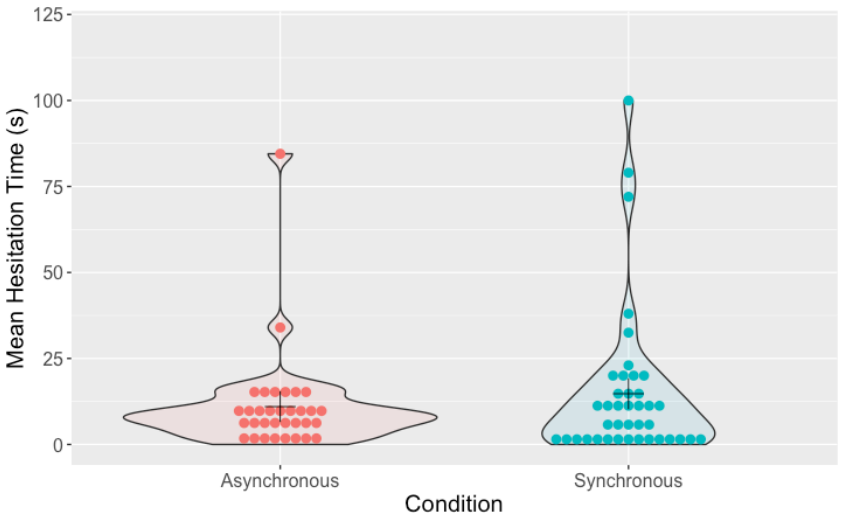


Figure 6. Length of hesitation between being asked to hit the robot, and agreeing to do so. Each circle represents an individual participant and the crosshairs indicate the mean for each group. Note: three participants refused to hit the robot and have been excluded from this plot as a result. The next section of results will discuss these individuals in greater detail.

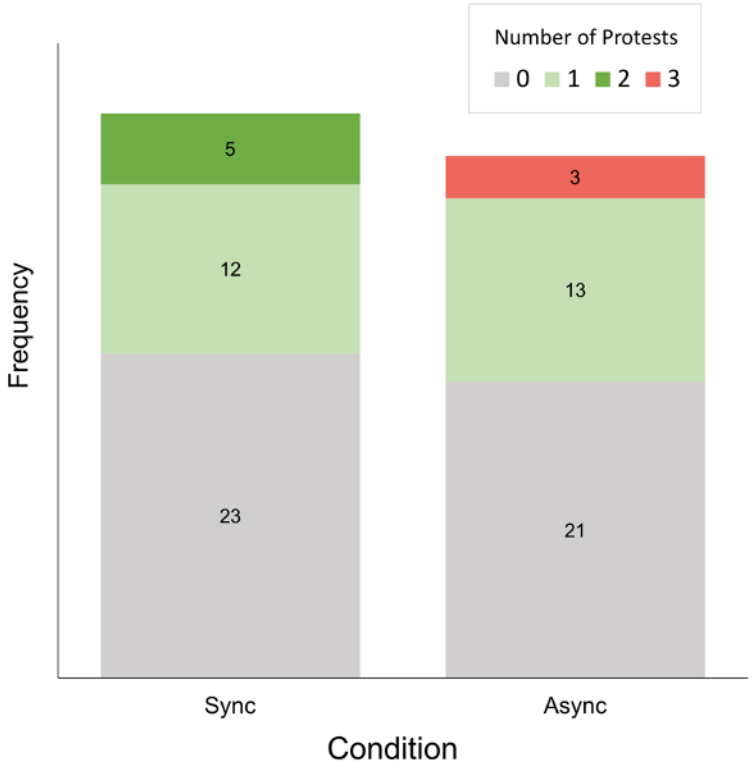


Figure 7. Chart to illustrate the frequency of individuals exhibiting the distinct number of protests (0, 1, 2 , or 3 protests). Note: After protesting 3 times the experiment was terminated, and the participant was deemed to have 'refused' to hit the robot.

Before interacting with the robot, the synchronous (M=1.78, SD=1.35) and asynchronous (M=1.89, SD=1.06) groups both demonstrated low scores on the Self-Other questionnaire. After interacting with the robot, the average scores of both groups increased slightly, but were still low (synchronous: M= 2.24, SD 1.36. asynchronous: M=2.03, SD = 1.34.). A 2x2 ANOVA provided no evidence for a significant effect of condition (synchronous vs asynchronous; $F_{1,\ 150} = 0.07$, $p = 0.796$, $n^2 = 0$) or time (Pre-Interaction vs Post Interaction; $F_{1,\ 150} = 2.09$, $p = 0.150$, $n^2 = 0.014$).

Gaze cueing

As part of our pre-registered exploratory analyses, we also evaluated whether gaze cueing effects differed between the two groups. To do so we performed a 2 x 2 x 4 x 3 mixed ANOVA: group (synchronous vs asynchronous), time (pre-interaction vs post-interaction), agent (human vs Pepper robot vs other robot vs arrows), congruency (congruent vs incongruent vs neutral). The assumption of sphericity was violated, therefore the Greenhouse-Geisser correction was used. See Figure 8 for a visual representation of the data split by the various factors (group, time, and condition).

In contrast to our hypothesis, the results also indicated that these effects were not significantly different between the two groups [agent x group: $F(2.470,167.989)= 2.382$, $p=.083$, $\eta P2 = .034$; congruency x group: $F(1.836,124.852)= 3.092$, $p= .053$, $\eta P2 = .043$].

The results of the ANOVA did indicate a main effect of agent ($F(2.470,167.989) = 8.355$, $p <.001$, $\eta P^2 = .109$). Descriptive statistics further reveal that on average, participants were slower to respond to the arrow condition ($M = 0.4043$, $SD = 0.1179$) compared to the human ($M = 0.3937$, $SD = 0.1837$), other robot ($M = 0.3963$, $SD = 0.1575$), and Pepper ($M = 0.3888$, $SD = 0.1087$) trials.

The results also indicate a main effect of congruency ($F(1.836,124.852) = 52.969$, $p <.001$, $\eta P2 = .438$). Descriptive statistics indicated that participants were, on average, fastest to respond to congruent trials ($M = 0.3795$, $SD = 0.1678$) compared to incongruent ($M = 0.3957$, $SD = 0.1271$) and neutral ($M = 0.4121$, $SD = 0.1357$) trials.

No significant interaction emerged between agent x congruency after applying the Greenhouse Geisser correction, $F(3.305, 224.734)= 2.161$, $p = .087$, $\eta P^2 = 0.020$. Additionally, time had no significant effect on the effects of agent ($F(2.324,158.006) = .926$, $p = .429$, $\eta P^2 = .013$) or congruency ($F(1.723,117.144)= 1.170$, $p = .309$, $\eta P^2 = .017$.

Discussion

Our aim in the present study was to determine whether synchrony between a participant's heart rate and the illumination of lights on a robot's torso (cardio-visual synchrony) would lead participants to exhibit increased prosocial attitudes and behaviors towards the robot. Our results demonstrate that overall, this particular kind of cardio-visual synchrony did not impact quantitative or qualitative measures of positive attitudes toward the Pepper robot, including self-report questionnaire responses, or hesitancy to hit the robot when instructed. Furthermore, we did not find evidence that heart rate synchrony influenced a cognitive measure of social attribution – specifically, gaze cueing effects.

Liking

The Godspeed questionnaire results indicate that participants in the synchronous group did not rate Pepper as more "likable", "friendly", "kind', "pleasant', or "nice", compared to the asynchronous group. As a result, no differences between the groups on the "Likeability" scale of the Godspeed questionnaire emerged. As can be seen in Figure 4, it would appear that both groups report high liking ratings for Pepper both before and after the manipulation. As such, liking ratings that cluster near the ceiling of the scale do not have much room for improvement.

An important limitation to note, regarding the Likeability scale of the Godspeed questionnaire, is that it cannot currently account for nuances between social liking and liking in other forms. For example, you might not perceive your kettle as friendly and kind, however a new add-on might make it more likable as a result of improved usefulness and usability. Consequently, it is not possible to rule out whether other factors of likeability not probed by this scale changed based on our manipulation. A challenge for future work, therefore, will be to explore ways of probing liking with more nuance and range. The focus on "likeability" results from our interest in perceptions of Pepper as a social agent, however it would also be insightful to probe factors affecting the usage and uptake of technologies as Pepper is inherently a machine. Davis *et al.* (1989) identified that "perceived usefulness" and "perceived ease" were positively correlated with current usage and future use of computers, so measures that take account of these factors would appear to be useful.

Hesitance to hit

Moving to the objective measures we collected regarding prosocial behaviors toward the Pepper robot, our reaction time data indicated that participants' hesitance to hit the robot did not vary significantly between the synchronous and asynchronous heart rate manipulation groups. It is of note, however, that the number of 'absolute refusals' (protesting against hitting the robot three times, resulting in the termination of the task) did differ between the groups. Contrary to predictions, in the synchronous group, no individual absolutely refused to hit the robot, whereas in the asynchronous group, three individuals absolutely refused to do so. Although it is impossible to conclude anything meaningful about our manipulation leading to this result, given such small numbers, this finding could nonetheless be followed up by future work.

If this finding were to be replicated, one possibility explaining why individuals in the asynchronous group refused to hit Pepper comes from literature regarding the effects of engineered music on relaxation (Leslie *et al.*, 2019). It has been demonstrated that slow music relative to an individual's breathing rate can induce slowed breathing and a calm state (Leslie *et al.*, 2019). In a similar way, the observation of the slow pulsing lights relative to the person's heart rate (as in the asynchronous condition) could have given rise to a similar relaxed state. This is supported by a statement from one participant in the asynchronous condition who commented that the pulsing lights were "calming". This calm state might have then led to an aversion to hit the robot, along the lines of

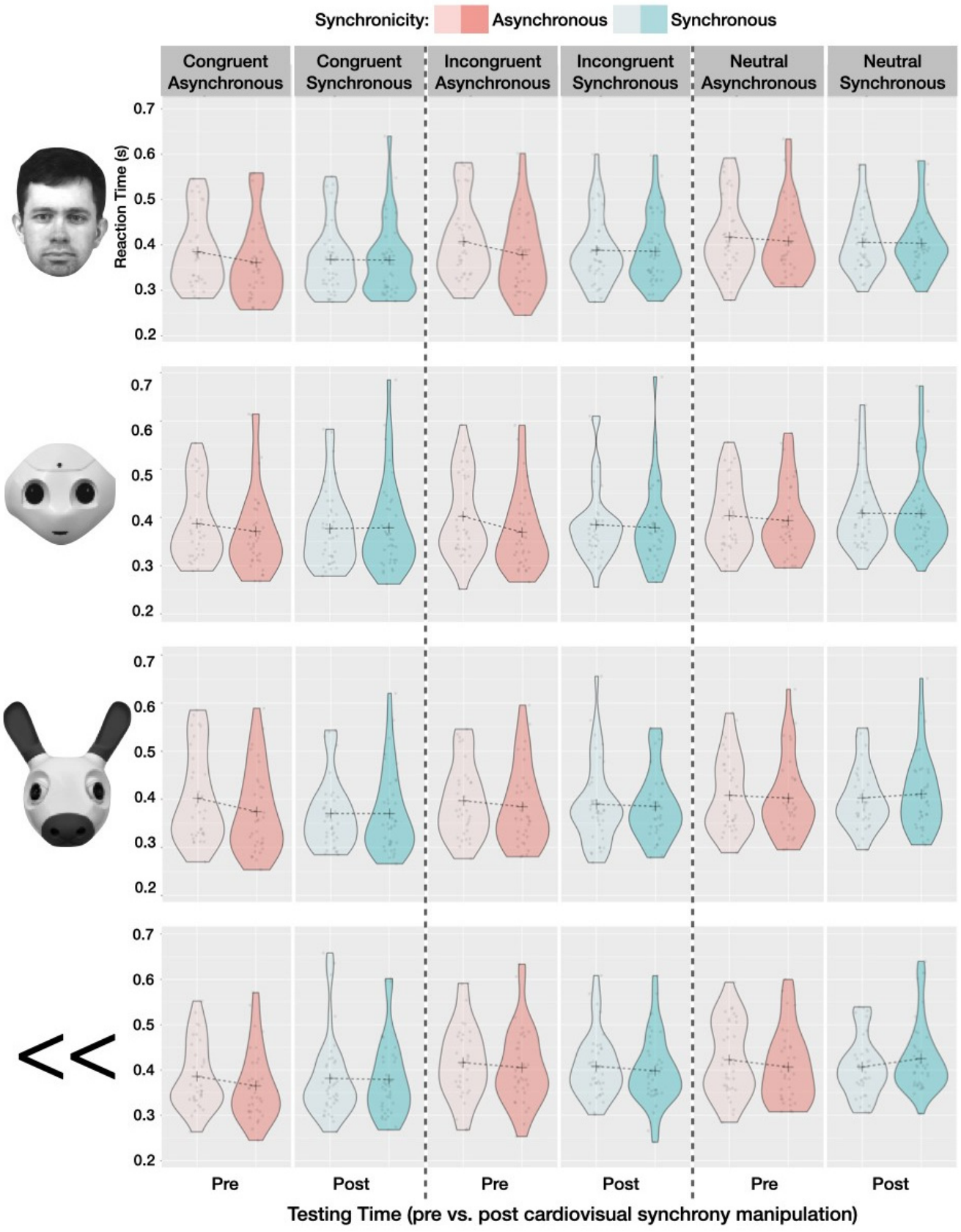


Figure 8. Plots to illustrate the gaze cueing reaction time data, split by agent type. Violin plots indicate the variance of the data within each condition.

empirical work that demonstrates induced relaxation can reduce aggression and increase prosocial behaviors (Whitaker & Bushman, 2012).

Many questions remain though, including why only three out of 37 people in the asynchronous group refused to hit Pepper, whereas the others agreed. Individual differences in attention to the lights, and susceptibility to the relaxing properties of the slow pulsing, offer suggestions as to why this might be the case. To gain greater insight regarding this theory though, it is necessary to undertake further experimentation targeting the use of illumination for relaxation, and using both subjective and physiological measurements of arousal.

It is important to note that although refusals and hesitation occurred, all but three participants (n=74) agreed when asked to hit Pepper on the head with the hammer. These results are comparable to those in a study by Rosalia and colleagues (2005), in which all participants complied when asked to administer the high level of electric shocks to a robot. This study was modelled on the (in)famous work by Milgram (1965), wherein participants were asked to administer electric shocks to a human confederate. In the original work with humans performed by Milgram, he reported that the majority of participants (24 out of 40) did not comply - refusing to administer the highest level of shock to the other person. It could be argued that compliance was lower in the Milgram study due to the verbal protests of the confederate, however Rosalia and colleagues (2005) also included vocal protests ("Please, please stop" etc.) in their robotic replication. While the Rosalia *et al.* (2005) study was very different from the current study in a number of important ways (differences in the setup, robot used, the nature of the harm, etc...) it nonetheless corroborates the suggestion that when instructed, people will more readily comply when asked to harm a robot, compared to a human. Such findings clearly underscore that we perceive living and artificial agents differently, and hold important ethical implications for human-robot interactions (Fosch-Villaronga & Albo-Canals, 2019; Ozcana *et al.*, 2016).

Qualitative insights

Despite finding no between-group difference in the hesitance to hit quantitative measures, the qualitative data gained from debriefing procedures was particularly insightful. Upon asking participants what they were thinking and feeling after being asked to hit the robot, participants indicated numerous reasons for hesitating. These included (but were not limited to) fondness for Pepper, worries about the cost of damaging the robot, concern for their own safety, disbelief about the instruction, feeling uncomfortable about being asked to commit an act of 'violence', social desirability effects, and feeling pressure from an authority figure. This diverse range of feelings and responses to the request to hit a robot leads us to suggest that it is not advisable to use reaction time alone as a measure of liking for this paradigm, and, returning to considerations presented in the Questionnaires section, that researchers should explore other methods of behaviorally probing liking and attachment.

For a detailed overview of the qualitative findings of this study, please refer to the work of Riddoch and Cross (2021). In addition to exploring participants' thoughts and feelings during the period of hesitation, the paper explores factors influencing apparent feelings of connection towards the robot.

Gaze cueing

A final task we performed to assess the extent to which participants viewed Pepper as a social agent, and whether the heart rate synchrony manipulation had any impact on such perceptions, was a gaze cueing task. The results from this task indicated no differences in gaze cueing effects between the two groups, or between pre- and post-manipulation testing points. While this finding is ultimately not surprising, given the null effects reported for other measures in this study, it is nonetheless valuable to consider what the gaze cueing findings suggest about social perception.

Previous literature indicates slowed responses for agents perceived as "mindful", compared to those considered to be lacking knowledge or insight (See Introduction for details). Given these previous findings, we expected clear increases in response times when participants responded to the human face compared to the arrow stimuli. However, our findings indicated an almost opposite pattern. Participants were slower to respond to the arrow, compared to both the human and robot faces. While this finding is certainly at odds with what has frequently be reported by prior literature (Morgan *et al.*, 2018; Teufel *et al.*, 2010; Teufel *et al.*, 2009; Wiese *et al.*, 2012), we suggest two reasons why people were on average slower to respond to the arrow than the human.

One explanation for why gaze cueing effects were stronger for the arrow, relative to the faces, is because the arrow is an extremely strong and salient direction cue, and the task in our study was to respond solely to the direction of the target. In previous studies participants have been asked to respond to the nature of the target (e.g. what letter is presented) in addition to the direction (Driver *et al.*, 1999; Greenwood *et al.*, 1993), offering one explanation why previous results differ to our own. It is possible that with the addition of a target-discrimination element, participants would be less distracted by the direction of the arrow, and we would see a pattern of results more reminiscent of previous findings (reduced gaze cueing effects for the arrow relative to eye gaze). Another reason for the unexpected finding could be because, contrary to many previous gaze cueing tasks, we used a full head shift, opposed to a movement of just the eyes.

Langton and Bruce (1999) demonstrated that head orientation and gaze direction both yield robust gaze cueing effects, leading us to initially deem the current stimuli as appropriate. However, upon reflecting on the result, we found evidence in the literature that when the face is averted by 30 degrees (similar to our stimuli), gaze effects were reduced (Hietanen, 1999). Hietanen (1999) argued that this is because when both head and eyes are averted, the agent is perceived to be oriented away from the participant and gazing straight ahead. As a result,

the behavior of the perceived agent is perceived as being less socially engaged or relevant to the task at hand, and the direction of attention has less signal value. The findings of the study by Hietanen offer insight as to why gaze cueing effects in response to the robot and human faces might have been reduced, rather than amplified, relative to the arrow.

With that being said, an emerging number of studies exploring gaze cueing effects with avatars and physical robots demonstrate that full shifts in visual attention (as demonstrated by a head and gaze shift) by embodied and virtual agents can indeed have powerful impacts on human behavior when performed *in situ* (e.g., Kompatsiari *et al.*, 2018; Schilbach *et al.*, 2010; Wilms *et al.*, 2010). Perhaps the most compelling research to combine the gaze cueing paradigm with embodied robots comes from work with the iCub robot, which indeed demonstrates that a physically co-located robot's gaze cues can be followed and used to induce prosocial feelings in a human interaction partner towards a robot (e.g., Ghiglino *et al.*, 2020; Kompatsiari *et al.*, 2021). Taken together, our failure to replicate the general human > arrows gaze cueing effect complicates any interpretation we may wish to make regarding the heart rate synchrony manipulation on following Pepper's gaze. However, we would suggest that probing more deeply into reasons why this is the case might not be as instructive or useful as further work looking to understand the role of eye gaze and prosociality with embodied artificial agents, as these, and not screen-based images, are the focus of research and development into socially assistive robotics.

Next steps

Cardio-visual synchrony. Perceiving the "other" as similar to the "self" is proposed to be a crucial mechanism underpinning the prosocial benefits of synchrony (Hove & Risen, 2009). Previous cardio-visual synchrony studies have supported this suggestion – finding that participants feel greater self-identification towards objects and images which are flashing in a synchronous (opposed to asynchronous) manner relevant to their heart rate (Sel *et al.*, 2017; Suzuki *et al.*, 2013). In this study however, average scores on the Self-Other overlap questionnaire were very low, regardless of whether Pepper was flashing in a synchronous or asynchronous manner. The lack of change before and after interacting with Pepper, regardless of group allocation, suggests that the manipulation was ineffective in terms of altering the perception of self-other overlap. Based on Hove & Risen's (2009) theory that prosocial behaviors are the result of identifying with another individual, it could be that we did not see positive effects of the manipulation in terms of liking, hesitation time, or any of the other measures due to the negligible perceived overlap between individual participants and Pepper. Questions still remain, though, whether the lack of change in self-other identification scores (and the general pattern of null findings) are most attributable to the robotic nature of the "other", the subtlety of the light manipulation, or both.

To reduce the subtlety of the manipulation, facilitate greater attention towards the lights, and potentially an increased impression of synchrony, future work could incorporate a task design that seeks to accelerate and decelerate participants' heartrate in response to some sort of external stimulus, while ensuring the robots' illumination also matches these changes in heartrate. Alternatively, future work could attempt to induce measurable effects by increasing the salience of the lights (e.g., by illuminating a whole limb, or face, as in previous studies) (Sel *et al.*, 2017; Suzuki *et al.*, 2013). We would also advise that researchers carefully consider how they define and measure self-other overlap, as we found that some participants were unfamiliar with the concept (and asked for clarification when answering questions).

To gain a more complete understanding of the effects of cardio-visual synchrony as a potential facilitator of prosocial behaviours and self-identification with a robot, it will be useful to conduct further work using different equipment and approaches. For example, in this experiment we used a commercially-available wrist-based sensor – as our goal was to create a manipulation which was inexpensive, feasible to introduce in real world settings, and subtle. By doing so however, we were limited to the measurement of pulse rate (opposed to heart rate, as measured in the chest) and it was not possible to precisely match the phases of the heart with the lights of the robot as a result. More accurate and reliable heart monitoring equipment (e.g., ECG) would be valuable in the pursuit of such heart-robot matching (while keeping in mind that such equipment also introduces other challenges for real-world applications as a result of increased costs and reduced ease of use).

A number of further limitations to the present study warrant discussion and will require further investigation in order to determine whether our failure to find the predicted effects of cardio-visual synchrony are due to the ineffectiveness of such a manipulation with robots in general, or peculiarities of our particular approach. For example, one issue to bottom out is whether the success of the human—robot synchronisation manipulation depends on the algorithm used. It would be valuable for future studies to explore different types of algorithms (for example, an oscillator model, which could help remove sources of noise) to test whether such approaches impact the effects of cardio-visual synchrony between humans and robots. Furthermore, the success of the experiment could also be related to the particular type of task used. At this stage, it would be useful to explore the extent to which synchronisation effects are amplified in sensorimotor tasks, or where sensory processing is related to interaction frequency (see Henschel & Cross, 2020). Along these same lines, unlike the control of arm or leg movements, or even breathing, most people cannot (easily) exert direct control over their heart rate, and heartrate itself is not easily visible to spectators. Even though an increasing number of studies exploring cardio-visual synchrony are emerging from a number of leading psychology laboratories around the world (Azevedo *et al.*, 2022; Galvez-Pol *et al.*, 2022; Heydrich *et al.*, 2018), it could be that the subtlety of such effects and manipulations remain ill-suited for implementation in human—robot interaction contexts.

Attachment to robots

The hesitance to hit paradigm, and the associated qualitative results, proved insightful – not only in terms of method validation (showing that participants hesitate to hit a robot for many reasons), but also for identifying individuals who quickly felt a connection the robot in this context. Unfortunately, there were unexpected consequences associated with the paradigm. Specifically, some participants reported feeling anxious and stressed whilst deciding whether to hit the robot. Additionally, the experimenter experienced high levels of stress as a result of repeatedly observing individuals in a state of distress (whilst having to remain composed and neutral).

Moving forwards, a clear need to balance maintaining realism while minimising participants' stress emerges (Geiskkovitch *et al.*, 2016; Rea *et al.*, 2017). As part of Geiskkovitch and colleagues' (2016) guidelines for researchers hoping to do so, they recommend that experimenters take the time to revisit and read the major codes of ethics that govern research with human participants, such as the Declaration of Helsinki (1964; 1996), the Belmont Report (1979), and the British Psychological Society (2014) Code of Human Research Ethics. It is also recommended that if researchers choose to conduct "robot abuse" studies in the future, they consider the implications of doing so – not only for the wellbeing of participants and the research team, but for the field of human-robot interaction more broadly. Studying moral behavior with robots presents new ethical challenges, and researchers working on the front lines of this discipline have a duty to actively reflect on and refine ethical guidelines (Chrisley, 2020; Rea *et al.*, 2017; Williams *et al.*, 2020).

Conclusions

Here we report that cardio-visual synchrony had no significant effect on prosocial attitudes and behaviors towards the Pepper robot. The subtlety of the manipulation offers one explanation as to why no difference emerged, and it would be beneficial for our understanding of perceived synchrony to conduct further research using a more obvious signal or a robot that illuminated more uniformly (such as the MiRo robot developed by Consequential Robotics). Although the current study failed to yield affirmative answers regarding the utility of cardio-visual synchrony for inducing prosocial thoughts and behaviors toward a robot, it nonetheless highlights potential issues with certain methods used to measure perceptions of, and behavior towards, robotic agents. Through qualitative data we show that the hesitance to hit paradigm is not only probing liking, but it also taps into other factors such as perceived cost and a person's concern for their own safety. We also find huge variation between participants with regards to how they refer to Pepper, and perceived feelings of connection towards the system. Based on current and other recent findings (e.g., Henschel & Cross, 2020), we would also argue that measures such as the Godspeed Liking scale are overly specific - focussing on social liking and neglecting other forms of liking in the process. In addition to advising the use of embodied human-robot interaction for probing other cognitive measures of social perception, such as gaze cueing (e.g., Kompatsiari *et al.*, 2018; Kompatsiari *et al.*, 2021; see also Henschel *et al.*, 2020), we also suggest that future studies also make more use of open-ended questioning of participants' subjective experiences. In doing so it is possible to gain new insights regarding measurement techniques and individual differences, and generate new research questions that are highly relevant to our future social engagements with robotic technologies.

Ethics and consent

All study procedures were approved by the College of Science and Engineering Ethics Committee (University of Glasgow, Scotland) – approval number 300180265. Participants provided written informed consent prior to taking part in the study.

Data availability

Underlying data

Open Science Framework: HeartBots - Investigating the Effect of Heart Rate Synchrony on Prosocial Behaviour towards a Social Robot. http://doi.org/10.17605/OSF.IO/D7C8T (Riddoch & Cross, 2020)

This project contains the following underlying data:

Folder: Main Study Data

- GazeCueing_ReactionTime. Raw Gaze Cueing data, collected in PsychoPy3.
- Hesitance_ReactionTime. Raw behavioural data – hesitance to hit the robot.

Folder: Qualitative Study Data

- BetweenCoderComparison. Files illustrating the codes created by the individual qualitative coders, followed by between-coder comparison.
- FinalCodesandThemes. Files containing the final qualitative data codes and themes.
- HighLowConnectionGroups. Raw data, allowing comparison between individuals expressing high and low connection to the robot.

Extended data

Open Science Framework: HeartBots - Investigating the Effect of Heart Rate Synchrony on Prosocial Behaviour towards a Social Robot. http://doi.org/10.17605/OSF.IO/D7C8T (Riddoch & Cross, 2020)

This project contains the following extended data:

Folder: Experiment Materials

- GazeCue_Experiment_PsyPy3. Gaze Cueing experiment setup which can be loaded and used in PsychoPy3.
- GazeCue_StimuliOnly. Stimuli image files, for the Gazecueing experiment.
- Pepper_RobotShoppingSpeech. html file, containing the panel used to trigger the robot's speech.

- PostHit_InterviewQs. Interview questions used after the Hesitance to Hit task.
- Procedure_Checklist. Procedure list used by the experimenter, to ensure all tasks and aspects were conducted in the correct order.
- QuestionnaireBooklet_Paper. Paper questionnaire booklet used for participants who were unable to use the computerised version.

Folder: Manuscripts & Figures

- MainManuscript_Preprint. Preprint manuscript.
- QualitativePiece_Figures. Figures used in the Qualitative Study Manuscript.

Software availability

Source code available from: https://github.com/SocialBrainInActionLab/Heartbot

Archived source code at time of publication: https://doi.org/10.5281/zenodo.7227615 (Bishakha & Blanco, 2022)

License: MIT.

References

Aron A, Aron EN, Smollan D: **Inclusion of other in the self scale and the structure of interpersonal closeness.** *J Pers Soc Psychol.* 1992; **63**(4): 596–612.
Publisher Full Text

Azevedo R, Von Mohr M, Tsakiris M: **From 'gut feelings' to first impressions: Phase dependent cardio-visual signals bias the perceived trustworthiness of faces**. 2022.
Publisher Full Text

Bartneck C, Kanda T, Ishiguro H, *et al.*: **Is the uncanny valley an uncanny cliff?** In: *RO-MAN 2007-The 16th IEEE international symposium on robot and human interactive communication*. IEEE, 2007a; 368–373.
Publisher Full Text

Bartneck C, Kanda T, Ishiguro H, *et al.*: **My robotic doppelgänger - A critical look at the Uncanny Valley**. In: *RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE. 2009b; 269–276.
Publisher Full Text

Bartneck C, Kulić D, Croft E, *et al.*: **Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots.** *Int J Soc Robotics.* 2009a; **1**(1): 71–81.
Publisher Full Text

Bartneck C, Rosalia C, Menges R, *et al.*: **Robot abuse-a limitation of the media equation**. 2005.
Reference Source

Batchelder L, Brosnan M, Ashwin C: **The development and validation of the Empathy Components Questionnaire (ECQ).** *PLoS One.* 2017; **12**(1): e0169185.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Benefield LE, Holtzclaw BJ: **Aging in place: merging desire with reality.** *Nurs Clin North Am.* 2014; **49**(2): 123–131.
PubMed Abstract | **Publisher Full Text**

Bharucha AJ, Anand V, Forlizzi J, *et al.*: **Intelligent assistive technology applications to dementia care: current capabilities, limitations, and future challenges.** *Am J Geriatr Psychiatry.* 2009; **17**(2): 88–104.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Bishakha, Blanco AM: **SocialBrainInActionLab/Heartbot: v0.1.0.** (v0.1.0). Zenodo. 2022.
http://www.doi.org/10.5281/zenodo.7227615

Bradwell HL, Winnington R, Thill S, *et al.*: **Longitudinal diary data: Six months real-world implementation of affordable companion robots for older people in supported living**. In: *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. 2020; 148–150.
Publisher Full Text

Broekens J, Heerink M, Rosendal H: **Assistive social robots in elderly care: a review.** *Gerontechnology.* 2009; **8**(2): 94–103.
Reference Source

Cacioppo JT, Hughes ME, Waite LJ, *et al.*: **Loneliness as a specific risk factor for depressive symptoms: cross-sectional and longitudinal analyses.** *Psychol Aging.* 2006; **21**(1): 140–51.
PubMed Abstract | **Publisher Full Text**

Carpinella CM, Wyman AB, Perez MA, *et al.*: **The Robotic Social Attributes Scale (RoSAS): Development and Validation**. In: *Proceedings of the 2017 ACM/IEEE International Conference on human-robot interaction*. 2017; 254–262.
Publisher Full Text

Chrisley R: **24 A Human-Centered Approach to AI Ethics: A Perspective from Cognitive Science.** *The Oxford Handbook of Ethics of AI.* 2020; 462–474.
Publisher Full Text

Darling K, Nandy P, Breazeal C: **Empathic concern and the effect of stories in human-robot interaction**. In: *2015 24th IEEE international symposium on robot and human interactive communication (RO-MAN)*. IEEE, 2015; 770–775.
Publisher Full Text

Davis FD, Bagozzi RP, Warshaw PR: **User acceptance of computer technology: a comparison of two theoretical models.** *Manage Sci.* 1989; **35**(8): 982–1003.
Publisher Full Text

Driver J IV, Davis G, Ricciardelli P, *et al.*: **Gaze perception triggers reflexive visuospatial orienting.** *Vis Cogn.* 1999; **6**(5): 509–540.
Publisher Full Text

Fosch-Villaronga E, Albo-Canals J: **"I'll take care of you," said the robot.** *Paladyn J Behav Robot.* 2019; **10**(1): 77–93.
Publisher Full Text

Foster ME, Alami R, Gestranius O, *et al.*: **The MuMMER project: Engaging human-robot interaction in real-world public spaces**. In *International Conference on Social Robotics*. Springer, Cham, 2016; **9979**: 753–763.
Publisher Full Text

Fozard JL, Rietsema J, Bouma H, *et al.*: **Gerontechnology: Creating enabling environments for the challenges and opportunities of aging.** *Educ Gerontol.* 2000; **26**(4): 331–344.
Publisher Full Text

Friborg O, Martinussen M, Rosenvinge JH: **Likert-based vs. semantic differential-based scorings of positive psychological constructs: A psychometric comparison of two versions of a scale measuring resilience.** *Pers Individ Differ.* 2006; **40**(5): 873–884.
Publisher Full Text

Friesen CK, Kingstone A: **The eyes have it! Reflexive orienting is triggered by nonpredictive gaze.** *Psychon Bull Rev.* 1998; **5**(3): 490–495.
Publisher Full Text

Frischen A, Bayliss AP, Tipper SP: **Gaze cueing of attention: visual attention, social cognition, and individual differences.** *Psychol Bull.* 2007; **133**(4): 694–724.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Galvez-Pol A, Antoine S, Li C, *et al.*: **People can identify the likely owner of heartbeats by looking at individuals' faces.** *Cortex.* 2022; **151**: 176–187.
PubMed Abstract | **Publisher Full Text**

Gayzur ND, Langley LK, Kelland C, *et al.*: **Reflexive orienting in response to short- and long-duration gaze cues in young, young-old, and old-old adults.** *Atten Percept Psychophys.* 2014; **76**(2): 407–419.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Geiskkovitch DY, Cormier D, Seo SH, *et al.*: **Please continue, we need more data: an exploration of obedience to robots.** *J Hum-Robot Int.* 2016; **5**(1): 82–99.
Reference Source

Ghiglino D, Willemse C, De Tommaso D, *et al.*: **At first sight: robots' subtle eye movement parameters affect human attentional engagement, spontaneous attunement and perceived human-likeness.** *Paladyn, J Behav Robot.* 2020; **11**(1): 31–39.
Publisher Full Text

Greenwood PM, Parasuraman R, Haxby JV: **Changes in visuospatial attention over the adult lifespan.** *Neuropsychologia.* 1993; **31**(5): 471–485.
PubMed Abstract | **Publisher Full Text**

Gurley K, Norcio AF: **A Systematic Review of Technologies Designed to Improve and Assist Cognitive Decline for Both the Current and Future Aging Populations**. In: *International Conference on Internationalization, Design and Global Development*. Springer, Berlin, Heidelberg, 2009; 156–163.
Publisher Full Text

Hawkley LC, Thisted RA, Cacioppo JT: **Loneliness predicts reduced physical activity: cross-sectional & longitudinal analyses.** *Health Psychol.* 2009; **28**(3): 354–63.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Henschel A, Cross ES: **No evidence for enhanced likeability and social motivation towards robots after synchrony experience.** *Interact Stud.* 2020; **21**(1): 7–23.
Publisher Full Text

Henschel A, Hortensius R, Cross ES: **Social Cognition in the Age of Human-Robot Interaction.** *Trends Neurosci.* 2020; **43**(6): 373–384.
PubMed Abstract | **Publisher Full Text**

Heydrich L, Aspell JE, Marillier G, *et al.*: **Cardio-visual full body illusion alters bodily self-consciousness and tactile processing in somatosensory cortex.** *Sci Rep.* 2018; **8**(1): 9230.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Hietanen JK: **Does your gaze direction and head orientation shift my visual attention?** *Neuroreport.* 1999; **10**(16): 3443–3447.
PubMed Abstract | **Publisher Full Text**

Hortensius R, Cross ES: **From automata to animate beings: the scope and limits of attributing socialness to artificial agents.** *Ann N Y Acad Sci.* 2018; **1426**(1): 93–110.
PubMed Abstract | **Publisher Full Text**

Hortensius R, Hekele F, Cross ES: **The perception of emotion in artificial agents.** *IEEE Transactions on Cognitive and Developmental Systems.* 2018; **10**(4): 852–864.
Publisher Full Text

Hove MJ, Risen JL: **It's all in the timing: Interpersonal synchrony increases affiliation.** *Soc Cognition.* 2009; **27**(6): 949–960.
Publisher Full Text

Insider Tech. 2018.
Reference Source

Kelly PA, Cox LA, Petersen SF, *et al.*: **The effect of PARO robotic seals for hospitalized patients with dementia: A feasibility study.** *Geriatr Nurs.* 2021; **42**(1): 37–45.
PubMed Abstract | **Publisher Full Text**

Kompatsiari K, Ciardo F, Tikhanoff V, *et al.*: **On the role of eye contact in gaze cueing.** *Sci Rep.* 2018; **8**(1): 17842.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Kompatsiari K, Ciardo F, Tikhanoff V, *et al.*: **It's in the eyes: the engaging role of eye contact in HRI.** *Int J Soc Robot.* 2021; **13**: 525–535.
Publisher Full Text

Langton SR, Bruce V: **You *must* see the point: automatic processing of cues to the direction of social attention.** *J Exp Psychol Hum Percept Perform.* 2000; **26**(2): 747–757.
PubMed Abstract | **Publisher Full Text**

Langton SRH, Bruce V: **Reflexive Visual Orienting in Response to the Social Attention of Others.** *Visual Cognition.* 1999; **6**(5): 541–567.
Publisher Full Text

Lehmann H, Saez-Pons J, Syrdal DS, *et al.*: **In good company? perception of movement synchrony of a non-anthropomorphic robot.** *PLoS One.* 2015; **10**(5): e0127747.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Leite I, Martinho C, Paiva A: **Social robots for long-term interaction: a survey.** *Int J Soc Robot.* 2013; **5**(2): 291–308.
Publisher Full Text

Leslie G, Ghandeharioun A, Zhou D, *et al.*: **Engineering Music to Slow Breathing and Invite Relaxed Physiology**. In: *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 2019; 1–7.
Publisher Full Text

Lundqvist D, Flykt A, Öhman A: **The Karolinska Directed Emotional Faces - KDEF, CD ROM.** from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, 1998.
Reference Source

Milgram S: **Some conditions of obedience and disobedience to authority.** *Hum Relat.* 1965; **18**(1): 57–76.
Publisher Full Text

Mogan R, Fischer R, Bulbulia JA: **To be in synchrony or not? A meta-analysis of synchrony's effects on behavior, perception, cognition and affect.** *J Exp Soc Psychol.* 2017; **72**: 13–20.
Publisher Full Text

Morgan EJ, Freeth M, Smith DT: **Mental State Attributions Mediate the Gaze Cueing Effect.** *Vision (Basel).* 2018; **2**(1): 11.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Mörtl A, Lorenz T, Hirche S: **Rhythm patterns interaction--synchronization behavior for human-robot joint action.** *PLoS One.* 2014; **9**(4): e95195.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Munafò M, Nosek B, Bishop D, *et al.*: **A manifesto for reproducible science.** *Nat Hum Behav.* 2017; **1**: 0021.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Niemelä M, Heikkilä P, Lammi H: **A social service robot in a shopping mall: expectations of the management, retailers and consumers.** In: *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction.* 2017; 227–228.
Publisher Full Text

Nomura T, Kanda T, Suzuki T, *et al.*: **Prediction of human behavior in human--robot interaction using psychological scales for anxiety and negative attitudes toward robots.** *IEEE Trans Robot.* 2008; **24**(2): 442–451.
Publisher Full Text

NVivo11: **NVivo qualitative data analysis (Version 11) [Computer software].** QSR International Pty Ltd, 2015.

O'Luanaigh C, O'Connell H, Chin AV, *et al.*: **Loneliness and cognition in older people: the Dublin Healthy Ageing study.** *Aging Ment Health.* 2012; **16**(3): 347–352.
PubMed Abstract | **Publisher Full Text**

Ozcana B, Caligiore D, Fosch-Villaronga E, *et al.*: **Designing socially assistive robot (sar) for cognitive child-robot interaction (ccri) with children with autism spectrum disorder-the case of "+ me.** In: *2nd International Conference on Social Robots in Therapy & Education.* 2016.
Reference Source

Peirce JW, Hirst RJ, MacAskill MR: **Building Experiments in PsychoPy.** 2nd Edn, London: Sage, 2022.
Reference Source

Perissinotto CM, Cenzer IS, Covinsky KE: **Loneliness in older persons: a predictor of functional decline and death.** *Arch Intern Med.* 2012; **172**(14): 1078–83.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Piau A, Campo E, Rumeau P, *et al.*: **Aging society and gerontechnology: a solution for an independent living?** *J Nutr Health Aging.* 2014; **18**(1): 97–112.
PubMed Abstract | **Publisher Full Text**

Pinquart M, Sörensen S: **Risk factors for loneliness in adulthood and old age--a meta-analysis.** Nova Science Publishers, 2003.
Reference Source

Qualtrics: **Qualtrics.** Provo, Utah, USA. 2005.

Rantz MJ, Marek KD, Aud M, *et al.*: **A technology and nursing collaboration to help older adults age in place.** *Nurs Outlook.* 2005; **53**(1): 40–45.
PubMed Abstract | **Publisher Full Text**

Rea DJ, Geiskkovitch D, Young JE: **Wizard of awwws: Exploring psychological impact on the researchers in social HRI experiments.** In: *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction.* 2017; 21–29.
Publisher Full Text

Rennung M, Göritz AS: **Prosocial Consequences of Interpersonal Synchrony: A Meta-Analysis.** *Z Psychol.* 2016; **224**(3): 168–189.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Riddoch K, Cross ES: **HeartBots - Investigating the Effect of Heart Rate Synchrony on Prosocial Behaviour towards a Social Robot.** [Dataset]. 2020.
http://www.doi.org/10.17605/OSF.IO/D7C8T

Riddoch KA, Cross ES: **"Hit the Robot on the Head With This Mallet" - Making a Case for Including More Open Questions in HRI Research.** *Front Robot AI.* 2021; **8**: 603510.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Riek LD, Adams A, Robinson P: **Exposure to cinematic depictions of robots and attitudes towards them**. In: *Proceedings of International Conference on Human-Robot Interaction, Workshop on Expectations and Intuitive Human-Robot Interaction*. 2011.
Reference Source

Rosalia C, Menges R, Deckers I, *et al.*: **Cruelty towards robots**. 2005.
Reference Source

Schilbach L, Wilms M, Eickhoff SB, *et al.*: **Minds made for sharing: initiating joint attention recruits reward-related neurocircuitry.** *J Cogn Neurosci.* 2010; **22**(12): 2702–2715.
PubMed Abstract | **Publisher Full Text**

Schönbrodt FD, Wagenmakers EJ: **Bayes factor design analysis: Planning for compelling evidence.** *Psychon Bull Rev.* 2018; **25**(1): 128–142.
PubMed Abstract | **Publisher Full Text**

Sel A, Azevedo RT, Tsakiris M: **Heartfelt self: cardio-visual integration affects self-face recognition and interoceptive cortical processing.** *Cereb Cortex.* 2017; **27**(11): 5144–5155.
PubMed Abstract | **Publisher Full Text**

Simmons JP, Nelson LD, Simonsohn U: **False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant.** *Psychol Sci.* 2011; **22**(11): 1359–1366.
PubMed Abstract | **Publisher Full Text**

Simmons JP, Nelson LD, Simonsohn U: **A 21 Word Solution.** *SSRN Electronic*

Journal. 2012; 14.
Publisher Full Text

Suzuki K, Garfinkel SN, Critchley HD, *et al.*: **Multisensory integration across exteroceptive and interoceptive domains modulates self-experience in the rubber-hand illusion.** *Neuropsychologia.* 2013; **51**(13): 2909–2917.
PubMed Abstract | **Publisher Full Text**

Takayanagi K, Kirita T, Shibata T: **Comparison of verbal and emotional responses of elderly people with mild/moderate dementia and those with severe dementia in responses to seal robot, PARO.** *Front Aging Neurosci.* 2014; **6**: 257.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Tanioka T: **Nursing and Rehabilitative Care of the Elderly Using Humanoid Robots.** *J Med Invest.* 2019; **66**(1.2): 19–23.
PubMed Abstract | **Publisher Full Text**

Teufel C, Alexis DM, Clayton NS, *et al.*: **Mental-state attribution drives rapid, reflexive gaze following.** *Atten Percept Psychophys.* 2010; **72**(3): 695–705.
PubMed Abstract | **Publisher Full Text**

Teufel C, Alexis DM, Todd H, *et al.*: **Social cognition modulates the sensory coding of observed gaze direction.** *Curr Biol.* 2009; **19**(15): 1274–1277.
PubMed Abstract | **Publisher Full Text**

The jamovi project: **jamovi (Version 1.6) [Computer Software]**. 2021.
Reference Source

Waytz A, Cacioppo J, Epley N: **Who sees human? The stability and importance of individual differences in anthropomorphism.** *Perspect Psychol Sci.* 2010; **5**(3): 219–232.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Whelan R: **Effective analysis of reaction time data.** *Psychol Rec.* 2008; **58**(3): 475–482.
Publisher Full Text

Whitaker JL, Bushman BJ: **"Remain calm. Be kind." Effects of relaxing video games on aggressive and prosocial behavior.** *Soc Psychol Personal Sci.* 2012; **3**(1): 88–92.
Publisher Full Text

Wiese E, Wykowska A, Zwickel J, *et al.*: **I see what you mean: how attentional selection is shaped by ascribing intentions to others.** *PLoS One.* 2012; **7**(9): e45391.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Williams T, Zhu Q, Grollman D: **An experimental ethics approach to robot ethics education.** *Proceedings of the AAAI Conference on Artificial Intelligence.* 2020; **34**(09): 13428–13435.
Publisher Full Text

Wilms M, Schilbach L, Pfeiffer U, *et al.*: **It's in your eyes--using gaze-contingent stimuli to create truly interactive paradigms for social cognitive and affective neuroscience.** *Soc Cogn Affect Neurosci.* 2010; **5**(1): 98–107.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Woo H, LeTendre GK, Pham-Shouse T, *et al.*: **The use of social robots in classrooms: A review of field-based studies.** *Educ Res Rev.* 2021; **33**: 100388.
Publisher Full Text

Wykowska A, Chaminade T, Cheng G: **Embodied artificial agents for understanding human social cognition.** *Philos Trans R Soc Lond B Biol Sci.* 2016; **371**(1693): 20150375.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Yang GZ, Nelson BJ, Murphy RR, *et al.*: **Combating COVID-19—The role of robotics in managing public health and infectious diseases.** *Sci Robot.* 2020; **5**(40): eabb5589.
PubMed Abstract | **Publisher Full Text**

Young JE, Hawkins R, Sharlin E, *et al.*: **Toward acceptable domestic robots: Applying insights from social psychology.** *Int J of Soc Robotics.* 2009; **1**(1): 95–108.
Publisher Full Text

Open Peer Review

Current Peer Review Status: ? ?

Version 1

Reviewer Report 26 June 2023

https://doi.org/10.21956/openreseurope.16218.r32129

© **2023 Bossi F.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Francesco Bossi
Information Engineering Department, University of Pisa, Pisa, Tuscany, Italy

The present manuscript describes a study investigating the effect of synchronous vs. asynchronous cardio-visual stimulation on the perceived likeability of a robot.
The study may show interesting results, but there is a major concern regarding the actual effectiveness of the manipulation the authors used, which needs to be discussed carefully or could affect the whole interpretation of the results. Moreover, statistical analyses are often inconsistent and the introduction is vastly off-topic.
I believe the manuscript needs a major revision before meeting the publication standards.
Specific issues are listed hereafter:

Introduction:

- The Introduction has a strong focus on the use of robots in elder people. Given the absence of a focus on older participants (or at least a comparison between younger and older participants), this introduction is deeply off-topic.

- Cardio-visual synchrony is not so apparent as movement synchrony or other forms of interpersonal synchrony: many participants may be unaware of their heartbeats if they do not bring voluntary attention to it. Moreover, previous literature shows that cardio-visual synchrony can affect the self-identification of body parts, not their likeability and not considering an external agent. Moreover, participants are not explicitly instructed to pay attention to the robot's shoulders, so the manipulation may not be effective in this sense.

- It might be that the slower rate in the asynchronous condition could create possible biofeedback relaxation effects. In this case, the manipulation would create the opposite effects compared to the authors' hypotheses. Please discuss the impact of this possibility.

Methods:

- Please justify the unusual choice of a power of .8 in the power analysis (typically a power of .9 or .95 is used).

- Please correct "alternate hypothesis" to "alternative hypothesis".
- What was the BF for the null hypothesis when you stopped data collection? This could represent a valid criterion instead of the BF for the alternative hp.
- The age range of the participants is huge. Why did the authors not consider age as an independent variable in the design? Considering the Introduction, it would seem that the authors hypothesized a difference related to age or different impact of the robot between younger and older participants.
- Heart rate: if lights are activated at the same heart rate as measured each second (and not with a higher sampling rate), you cannot be completely confident that they will be exactly synchronous to the participant's heartbeat in phase. Paradoxically, this manipulation in the synchronous condition could generate a visual stimulus with a constant offset to the participant's heartbeat, resulting in perceived asynchrony. In the same way, you cannot be completely sure that the asynchronous condition is not corresponding to (at least some of) the heartbeats. Please highlight and discuss this limitation carefully, as you started doing in "Next steps".
- Gaze cueing: the difference related to perceived human likeness should be found in the DIFFERENCE between congruent and incongruent trials, which is not considered in this paper.
- "asterix" -> "asterisk"
- "Responses were removed if slower than 150ms" -> "FASTER than 150ms", I assume.
- Were outliers not removed from the gaze cueing dataset? Removed responses still do not account for the normality assumption of the response curve. This should be controlled in some way (outliers removal or log-transformation).
- Did you check how many participants reported having already understood that the robot was controlled by the experimenter during the debriefing?
- Did the authors check for any differences among the other Godspeed subscales apart from Likeability?

Results:

- GENERAL: I honestly appreciate the use of Bayesian analyses in this paper. However:
  1) The meaning of the Bayes Factor should be explained in the methods for readers who are not familiar with Bayesian statistics.
  2) The BF10 (in favour of the alternative hp) is extremely low. In these results, it would be extremely useful to show the BF01 instead (in favour of the null hp), which represents evidence in favour of a null effect.
  3) In some tests the authors report Bayesian analyses, while in other tests they report frequentist inferential tests (es: t and p-values). It would be best to uniform the analyses to one of the two approaches, or at least justify the choices between the two. I would suggest adopting a Bayesian approach for all analyses since you obtained several null results.

- HP1: Did you check the 2x2 interaction effect in the ANOVA? That would be the effect that can test whether the difference between before and after the interaction is different between the two conditions.
- HP2: How did you measure the time of hesitation? When did the count start exactly? And how did you measure this time?
- HP3: The authors specify they used a Mann-Whitney U test and then report results from a t-test.
- Self-other: see comment about HP1: you should most importantly test the interaction.
- Gaze cueing: given the hypotheses, I would expect a 4-way interaction effect of group x time x agent x congruency (i.e., between the two groups, the congruency effect changes before and after the exposition to the robot, differently for each agent - specifically for Pepper). Did the authors check for all 3-way and 4-way interactions?
- Gaze cueing: post-hoc tests should be run on the significant main effect of agent, in order to assess what agents have significantly different response times.

Discussion:

- Please discuss the fact that this specific manipulation might have some intrinsic issues in eliciting an effect. The authors should discuss in a more extensive way why their manipulation did not work as hypothesized.
- Gaze cueing: before discussing differences between agents, post-hoc tests should be run on the significant main effect of agent, in order to assess what agents have significantly different response times. Moreover, the gaze cueing effect is assessed as the difference between incongruent and congruent condition (i.e., agent x congruency interaction effect), which is not the case in this study.

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and does the work have academic merit?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
No

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Cognitive neuroscience; Social neuroscience; Human-robot interaction

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 10 March 2023

https://doi.org/10.21956/openreseurope.16218.r30929

© 2023 Galvez-Pol A. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Alejandro Galvez-Pol
Psychology Department, University of the Balearic Islands, Palma de Mallorca, Spain

Riddoch and Cross studied the effect of cardio-visual synchronisation on: i) participants' liking towards a robot, ii) self-report measures, iii) a gaze-cueing task (to test pre- and post- effects of the main manipulation above). Also, iv) they tested whether participants, after being asked to do so, would strike the robot with a mallet.

Cardio-visual a/synchronisation was achieved by manipulating the lighting of the robot's shoulders. These could lighten according to the participants' heartbeats or slightly slower (async).

The results show no significant differences in all pre- and post- effect measures (between groups); regardless of the group and manipulation of lighting, participants equally liked the robot. This no difference in outcomes also encompassed the rest of the measures (reluctance to hit the robot and reaction time in the cue gaze paradigm).

Overall, the manuscript is clear and well-written. The rationale behind the theory (background) and methods (paradigm and results) are well explained and seem justified. This is an informative study and I do think that it opens many forthcoming and interesting questions. Obviously, I think the study is relevant regardless of the lack of significant differences.

Below are some comments that I hope might improve the scope of the manuscript:

Regarding the likeability scores/scale: I notice that many participants scored 25 (max.) or around the max. Would it be possible that such a measure suffers a ceiling effect? This could potentially mask whether they like Pepper more or less, and before and after. Also concerning these data:

1. Could the authors use the single dimensions of the self-report? Dislike-Like, Unkind-Kind, Unfriendly-Friendly, Unpleasant-Pleasant, and Awkward-Nice. I'm not familiar with the questionnaire though, but I believe it might be possible.

2. The plot/s is/are not that intuitive and transparent. I recommend using lines depicting the change in likeability between the baseline (i.e., pre-) and post- interaction stages. Much alike this:

   https://www.google.com/search?q=raincloudplots+repeated+measures&sxsrf=AJOqlzV_DviLJmtexA5IqpM
   CLMQ_AUoAnoECAEQBA&biw=1366&bih=797&dpr=1#imgrc=JDdJ5JpXIf_XeM

   *I don't mean the authors should use raincloud plots, but depict the trajectory of subjects along the experiment.

3. A suggestion: the authors could show the absolute difference in likeability (and potentially other measures). This is to say: if someone scores 25 in pre- and 25 in post-, the score is 0. Here the data can be shown in a single point and differences can appropriately visualised. In this context, given that many participants scored almost max. in this questionnaire (around 23-25). The score can be tested against zero, etc.

4. As exploratory analysis, is it possible to disregard these 'ceiling participants' and statistically analyse the ones that could show a variation? Given the smaller n, a boostrapping might be included.

In addition, about accounting for individual differences: participants' features might mediate the target measures (those expected to vary as a function of the lighting synchronisation). There might not be a direct and single influence between one self-reported variable and the dependent variable, but a relationship between two or three of them on the DV. I suggest the authors at least explore this possibility. Here, the authors might check these examples: i) Palser *et al.* (2018[1]) and Arslanova *et al.* (2022[2]) (besides, it would be nice to include the reference to this latter paper when citing at the end of the manuscript the related paper in Cortex).

Overall, the above comments follow these recommendations: i) work a bit more on the analysis, ii) try to uncover not so straight forward trends, iii) maybe work on the clarity of the plots.

About the methods: it is not clear to me if the HB synchronisation worked correctly. Did the authors have previous experience with the setting? Did they make any checks about potential delays in the signal? Proper communication between devices, etc? When they mean 20% slower, do the authors mean delaying each HB 20% of the time required to complete the recorded heartbeat?

Also, could the recording of the HBs on the wrist affect the subjective feeling of sync? If someone feels (even unconsciously) the HB on the chest, the feeling of sync might be different because I imagine the blood needs some time to travel around the body [e.g., heart to wrist, one could feel the HB on the chest while the experimental signal is recorded from the wrist (this is just a hypothesis)].

Minor comments:

- In the discussion, the figure about liking is Fig. 5 not 4.

- In the methods: 'Responses were removed if slower than 150ms (‘preemptive’)' - I think the authors meant 'faster'.

- Regarding the lighting of the shoulders: the lighting areas seems quite small. Do the authors expect different results if the lighting happened around the face are for instance? The shoulders, blue lighting, and the thin lights do not seem, a priori, clear socially relevant physiological cues. This is just a hypothesis. The authors could comment on this.

References

1. Palser ER, Palmer CE, Galvez-Pol A, Hannah R, et al.: Alexithymia mediates the relationship between interoceptive sensibility and anxiety.*PLoS One*. 2018; **13** (9): e0203212 PubMed Abstract | Publisher Full Text
2. Arslanova I, Galvez-Pol A, Kilner J, Finotti G, et al.: Seeing Through Each Other's Hearts: Inferring Others' Heart Rate as a Function of Own Heart Rate Perception and Perceived Social Intelligence. *Affect Sci*. 2022; **3** (4): 862-877 PubMed Abstract | Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and does the work have academic merit?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Psychology, physiology, cognitive neuroscience, interoception

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.