

Human performance during joint action with iCub: a successful adaptation of a paradigm of cognitive neuroscience in HRI*

J. Pérez-Osorio, D. De Tommaso, E. Baykara, & A. Wykowska

Istituto Italiano di Tecnologia, Genova, Italy

Abstract— Robots will soon enter social environments shared with humans. We need robots that are able to efficiently convey social signals during interactions. At the same time, we need to understand the impact of robots' behavior on the human brain. For this purpose, human responses to the robot behavior should be quantified offering feedback to improve and adjust robot behavior. Under this premise, we aimed to adapt a classical paradigm of experimental cognitive psychology into a naturalistic human-robot interaction scenario. This allowed for maintaining excellent experimental control offered by a paradigm borrowed from cognitive psychology, and for ecological validity of an interactive protocol. We showed the feasibility of such an approach with a validation study, which demonstrated that our design yielded similar results to what has been previously observed in experiments within the area of cognitive psychology. Our approach allows for addressing specific mechanisms of human cognition that are elicited during human-robot interaction, and thereby, in a longer-term perspective, it will allow for designing robots that are well-attuned to the workings of the human brain.

*This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant awarded to AW, titled "InStance: Intentional Stance for Social Attunement". Grant agreement No: 715058). All authors are with the Istituto Italiano di Tecnologia, Genova, 16145.

I. INTRODUCTION

During social interactions, we attend to others' movements, facial expressions, and gaze direction. These social signals offer valuable information to infer others' mental states and intentions, and help to anticipate future events. That is because humans are very good detecting subtle social signals to predict others' behavior in complex social context. On the brink of the robot era, robots will soon share human social contexts. Either as social companions, bar tenders or tour guides in museums, robots will face rich social environments in which robots will be expected to convey the most basic social signals. In order to evaluate the efficiency of those signals, we believe that HRI research should use a systematic approach based on well-documented experimental protocols, grounded in cognitive neuroscience theories. This perspective allows for evaluate the impact of robot behavior on humans beyond self-reported measures and also offers valuable data to robot designers regarding the skills expected from the robot. Importantly, this evaluation/adjustment enables a cycle towards more socially capable artificial agents. Starting from these premises, we implemented a pilot study based on a experimental protocol to evaluate participants' performance in a joint action scenario with the humanoid robot iCub [1]. In detail, we measured how participants responded to the social signals of the robot as a function of the expectations related to a social context. In the following, we briefly describe the literature on the topic of how humans predict others' behavior in social interactions and how these factors have been addressed in HRI research.

A. Gaze cueing

Discriminating and following others' gaze direction is an essential component of establishing a common social context and pivotal to the ability of infer others' mental states (e.g., [2–5]). In the lab setting this phenomenon has been studied using the gaze-cueing paradigm. Here a face (with eyes looking straight ahead) is presented in the centre of the screen. After 500-1200 ms the eyes look towards one of the sides. Then, a target is presented in the same or in different location. Typically, participants responses are faster when the target tallies with the gaze-at location (valid cue), compared to targets that appear in the opposite location (invalid cue). This has been observed for detection (i.e., determine whether the target is present

or absent) and discrimination tasks (i.e., differentiate targets based on a given feature). The difference in the reaction times (RTs) is referred to as the gaze-cueing effect (GCE). Extensive research reveals that gaze following is a very robust and easy to deploy attentional mechanism, and it has been thought to involve an automatic/reflexive component [6], [7]. Neuronal correlates of GCE reveal that the observed shift of gaze direction summonses the focus of attention of the observer to the indicated location in space. Therefore, it facilitates the subsequent processing of the target on that particular spot, relative to different locations.

GCE has been investigated in HRI with mixed results. While some studies suggest that the effect is completely absent [8,9], other suggest that following a robot face is more like following the direction of an arrow [10]. However, most recent results show that robots can elicit GCE [11-12]. Furthermore, in a gaze cueing paradigm involving an embodied humanoid agent (the iCub robot), Wykowska et al. [13] showed GCE, independent of whether participants perceived behavior of the robot as human-like or more mechanistic. Collectively, findings suggest that embodiment and appearance might play a significant role in GCE during HRI.

B. Prediction of action steps

Follow others' gaze direction is also useful to predict their upcoming actions. Recordings of eye movements of participants observing someone performing a common action (i.e. preparing a sandwich) revealed that eye fixations predicted action steps in a sequence [14, 15]. Interestingly, eye movements during the same task were strongly coupled to the task-relevant objects and preceded their use. These findings strongly suggest that gaze behavior provides reliable hints regarding successive action steps of others, helps to pick up critical information for performing the task, and supports high-precision movements. Thus, following others' eye movements might facilitate prediction observed actions in natural situations.

Similarly, studies in HRI reveal that humans attended to the robot gaze as a cue for predicting behavior. In a "guess the object" game with a humanoid robot, participants were able to determine relevant locations

on a table following a humanoid robot gaze [16]. Observers used gaze behavior to make predictions about which objects to select and showed improved performance in the task (i.e. guess which object the robot had chosen) relative to a condition without gaze-related information. Interestingly, participants reported that they were not aware of following such cues. Consequently, it is plausible that humans would be able to use also robots' gaze to predict action steps in complex social human-robot interactions.

In the light of the state-of-the-art research presented here, it seems reasonable to suggest that during complex action scenarios gaze following and expectations regarding goal-directed actions might be closely related. In fact, we showed that gaze-guided orienting of attention can be modulated by expectations about observed action sequences [17,18]. Those studies embedded a traditional gaze-cueing paradigm in a context involving a representation of a complex high-level action. The paradigm aimed to evoke expectations regarding action sequences and gaze behavior similar to a natural situation. To achieve that, we used a series of naturalistic photographs depicting a person (i.e. April) completing a goal-oriented task. In each trial April had an action goal: either someone asked her for something to drink or someone asked her to fetch fabric softener to do the laundry. Then April was shown behind the counter with two relevant objects (two bottles, one with orange juice and the other with fabric softener). Then she gazed at either the action-congruent or action-incongruent bottle (e.g., orange juice was congruent in the context of "bringing something to drink"). After the gaze shift, some of the liquid (either orange juice or softener) appeared in a plastic glass placed next to the bottle (the target). Participants' task was to respond whether the level of liquid in the plastic cup was high or low –a target discrimination task. We found that participants were faster when the target coincided with the observed gaze-at locations relative to the opposite locations, in line with the classical GCE. Importantly, the GCE was modulated by whether the observed agent's gaze was directed to the object congruent or incongruent with the action context. In more detail, when the actor's gaze behavior violated observer's expectations (i.e. April looked at the bottle with fabric softener when she was asked to bring something to drink), the GCE was strongly attenuated or even completely suppressed, relative to when gaze behavior confirmed expectations. These findings were

replicated and supported by electrophysiological evidence [18]. In sum, our findings confirmed that action prediction, expectations regarding action sequences and gaze cueing are closely linked. We showed that the fundamental mechanisms of social cognition can be modulated by higher-order expectations regarding others' behavior.

II. AIM OF OUR APPROACH

Gaze following, predictive gaze and action prediction have been investigated separately in research on HRI. However, to the best of our knowledge, these interplay of crucial cognitive mechanisms have not yet been measured together in human-robot or human-human joint action scenarios.

The aim of the present pilot study was to demonstrate the feasibility of implementing an experimental protocol in a naturalistic interaction with iCub. This approach has multiple significant advantages as it provides for (i) maximum experimental control, as we follow well-documented experimental protocols, which allow to (ii) isolate specific social cognitive mechanisms, based on previous findings reported on cognitive psychology, and grants (iii) a unique design to evaluate GCE within the context of action expectations in real action scenarios for the first time. All together, this approach offers high reliable objective measures of performance that can be used to adjust and improve social skills on the robots.

Despite of the clear benefits, adapting previous studies to a naturalistic HRI scenario certainly represents a considerable challenge. It requires to integrate into one setup human performance measures, stimulus presentation and robot behavior, with great temporal resolution and remarkable synchronization. In order to show the feasibility of the implementation, the present paper focuses mostly on the validation of the approach. That is, we present in great details the technical requirements, method, materials, and procedure. Additionally, as a validation, we present data collected from three participants in order to show that the setup actually works as intended.

Collectively, we aim to show the feasibility of adapting well-established protocols of cognitive neuroscience and experimental psychology to naturalistic HRI scenarios with the utmost methodological rigorousness and high replicability.

III. METHOD

A. Materials

The main processing unit (SSD 250 Gb, 32 Gb RAM, Intel i7 7700CPU@3.60 GHz 4 cores, GPU GeForce GTX 1060 3Gb, running Ubuntu 12.04 LTS) was connected to: (1) the iCub robot through a peer-to-peer Ethernet connection, to (2) the table projector through HDMI and to (3) the eye-tracker glasses through a Wi-Fi connection. All the software modules in the experiment have been integrated into a single workstation. Based on OpenSesame (an open-source, graphical experiment builder written on Python [19]), all iCub's movements, the collection of eye tracker and performance data were controlled and synchronized. For this purpose, a set of custom-made plugins was developed.

A combination between OpenSesame and the iCub middleware YARP (Yet Another Robot Platform [20]) controlled the movements of the robot. Based on the requirements of the experiment, the iCub robot executed two kinds of tasks: looking at one object placed on the table (Task 1) and grasping one of those objects in order to hand it to the participant (Task 2). For Task 1 we implemented a Python wrapper for a 6-DOF gaze controller [21]. The controller is able to track a 3D Cartesian fixation point calculating neck and eyes poses to look at that point. Indeed we configured four different 3D target points: a) object on the left, b) object on the right, c) participant's face and d) center of the table. iCub moved both its eyes and its neck to indicate the bottle. The eyes and the neck of iCub were controlled by the YARP Gaze Interface, iKinGazeCtrl [22]. The vergence of the robot's eyes was set to 2 degrees, and the tilt was +/- 3 degrees, both maintained constant through the experiment. To ensure the same robot movement for all the conditions the y-coordinate of "left" and "right" was the same. The trajectory time for the movement of eyes and neck was set to 1000 ms and 1400 ms respectively, to maintain the impression of a smooth and naturalistic movement. For the Task 2 we implemented a Python wrapper for an iCub module called Constant Time Position Service. This service allows sending target joint positions and the trajectory time for executing the movement. Then, the velocity profiles and the joints trajectory are generated for reaching

that target position in the required time. All the robot movements for manipulating the objects were pre-recorded to guarantee the minimum variability in terms of time and joint positions.

For the monitoring of the eye movements, we used Tobii Pro Glasses 2 [26]. This commercial mobile eye tracker is lightweight and records ocular data up to 100Hz. The plugin implemented to control the Tobii Pro Glasses 2 allowed to start/stop the recordings for each participant, to calibrate the eye-tracker at the beginning of each session, and to send triggers related to the different conditions of the experiment. Markers used to characterize events during the recording were sent via WiFi to OpenSesame. The markers allowed to analyze offline the conditions of interest. For the eye tracking analysis we processed the data using Tobii Pro Lab software.

The initial sentence was created using the free online Text To Speech (TTS) service with natural sounding voices (<http://www.fromtexttospeech.com/>). All the sentences were matched for volume (45 dB).

B.Procedure

Three healthy volunteers (1 left-handed, mean age = 26, all women) took part in the validation. All provided their informed written consent prior to participation. Data were collected at the Istituto Italiano di Tecnologia, IIT, Genova. The study was approved by the local ethical committee (Comitato Etico Regione Liguria). We implemented a modified version of the experimental protocol used by Perez-Osorio et al. [17,18]. Participants were standing in front iCub (See Figure 1).

The figure 2 shows an example of a trial sequence. Every trial started with two bottles placed on the translucent table on top of a dark grey placeholder. After that, an auditory stimulus (sentence) informing participants about the action context (drink, laundry or neutral) was presented. Then, the robot looked at the participant for 2 s while the sentence was reproduced through the speakers (“something to drink”, “something to do the laundry” or “the weather is nice”). Then, the identity of the bottles was revealed (projected from below). One bottle was a beverage (yellow), and the other a fabric softener (blue). Subsequently, the robot looked either to the left or to the right bottle. The time from looking straight ahead

to the end of the head and eyes movement was 1.2 seconds. At the end of the robot movement, the target appeared for 300 ms under one of the bottles. Simultaneously to target appearance, an octothorp (“#”) was presented, on the opposite side. Participants responded to target with the left index and middle fingers using the letters “z” for “V” and “x” for “T”, marked accordingly. After 300 ms the target was masked with a octothorp until response or up to 1.3 s (target timeout). Participants had to respond which letter appeared, as accurate and as fast as possible. After target offset, the robot looked at the participant for 1s, then gazed to one the bottles, approached, grasped and offered it to the participant. Participants were asked to receive it and put it back on the placeholder (colored circle). At the end of each trial, participants evaluated whether iCub took the correct bottle depending on the request made at the beginning. The question “Did iCub take the correct bottle” appeared at the centre of the table. Participants responded to the question using the right hand over the number pad on the right of the keyboard. The options were “1”

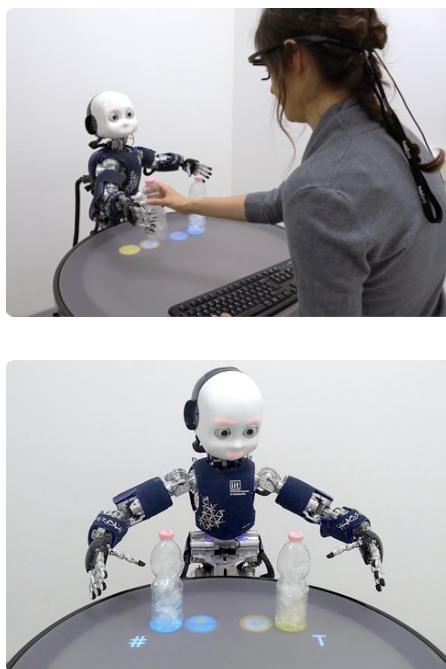


Figure 1. Experimental setup. The lower panel shows a valid cue (from the participant’s perspective) - where the robot’s gaze is directed to the same location as the appearance of the target.

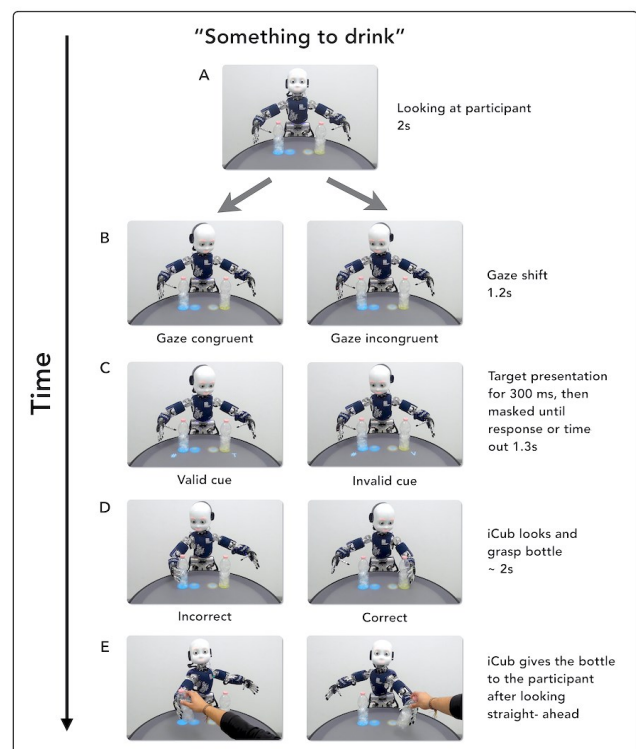


Figure 2. Example of an experimental sequence from the participant’s perspective. The objects on the table were illuminated from below using the translucent table. The identity of the bottles changed randomly between trials. The blue bottle indicated the softener and the yellow bottle the drink. Similar to a naturalistic situation, iCub looked at the bottle before grasping, and towards the participant to deliver the object.

for “yes”, “2” for “no” or “3” for “NA”. NA was for “does not apply” when the sentence was related to the weather. The questions were on the screen for 5s.

All conditions were pseudo-randomized across trials; also the side on which each bottle was presented (drink left or right), the target type (F or T), and the target position (low or high) were pseudo-randomized. The total number of trials was 48, presented in 4 blocks of 12 trials each. All participants received identical instructions and had a block of practice with the robot and using the eye tracking glasses to facilitate adaptation to the task.

C.Experimental design

We manipulated congruency by embedding an action scenario in each trial sequence. Both factors (action congruency and spatial validity) are defined as (a) iCub's gaze landing on an object congruent or incongruent with the action context, and (b) the target object being spatially validly or invalidly cued by iCub's gaze direction. Importantly, the critical manipulation was that iCub's gaze was directed to either the bottle congruent or the bottle incongruent with the action context. We added a congruency-neutral condition in which there is no request regarding the bottles. This means that iCub's gaze behavior was not related to any action context. To make the protocol more naturalistic, we made iCub actually grasp one of the bottles and hand it to the participant. Following previous findings, we hypothesized that we should observe GCE modulated by congruency of the gaze with respect to action context and that this interactive effect might even be enhanced in a real interaction protocol, as compared to screen-based experiments.

In addition to the behavioral measures, we decided to evaluate eye movements during the task. These measures would reveal how attentional mechanisms are deployed in such task. Importantly, eye tracking would allow us to observe whether participants follow iCub's gaze, whether action context information guides attentional mechanisms and whether predictions regarding action sequences have an impact on gaze following.

Gaze behavior was analysed on three different segments: 1) *Action-related context*: from the beginning of the trial until right before the gaze movement started; the aim was to determine whether participants looked at the bottle congruent with the action context or not; 2) *Gaze cueing*: from the starting of the gaze-head movement to letter response; this segment is crucial for examining whether participants followed the gaze of the robot. Finally, 3) *grasping and handing-over the bottle*: from the letter response to the moment when the participant receives the bottle; here, we examined gaze behavior during the joint action task. For each segment, we determined three AOI based on global fixations (face, left and right). From those AOIs we extracted total fixations and accumulated fixation duration.

To the best of our knowledge, this is the first setup integrating and synchronizing all the software and hardware components to implement in a naturalistic HRI scenario a precisely-controlled experimental gaze cueing paradigm used in cognitive neuroscience to measure gaze cueing embedded in action predictions.

III. VALIDATION

We tested our setup only with three volunteers. We report data from this pilot validation study in order to show that our setup worked properly. Meaning, data collection worked as intended and performance patterns were similar to what was hypothesised based on previous research. However, as the main aim of this pilot study focuses on the successful implementation and validation of this unique setup, the sample

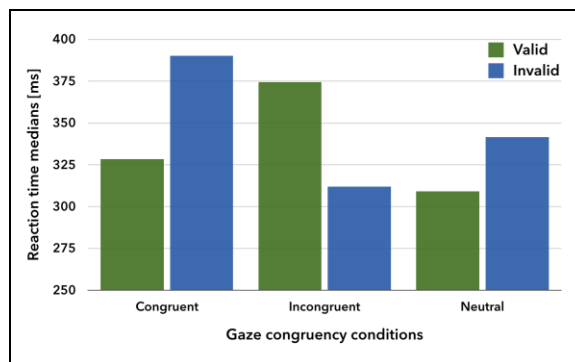


Figure 3 Median reaction times per condition, only trials with correct responses (90.28%)

of participants was not sufficient to draw any statistical inferences. Therefore, we only report data patterns

that suggest expected effects. Follow-up studies with appropriate sample sizes will determine the statistical significance of the tendencies observed here.

A. Reaction times

Median reaction times showed a pattern similar to what we observed in previous studies: participants tended to respond faster to validly cued targets relative to invalidly cued targets in congruent and neutral conditions (see Figure 3). However, the pattern of responses tended to be inverted in the incongruent condition, with faster RTs for incongruent invalid trials, relative to the incongruent valid condition (a pattern suggesting an inverse GCE).

B. Eye tracking data

On average, 96% percent of fixations were collected by the eye tracker. The dominating feature of performance was that participants kept their gaze on the robot's face (between 60 and 80% of all fixations) during all analyzed segments.

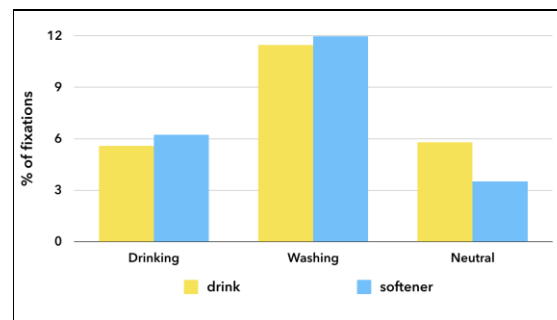


Figure 5. Percentage of fixations after the sentence and before the gaze cueing.

1) Social context (eye tracking data)

There was a smaller proportion of fixations on either of the bottles, relative to the robot's face (see Figure 5). Participants seemed not to look more often at the bottle congruent with the action context, in the segment preceding the directional gaze shift. Interestingly, when the context was related to *laundry*, participants tended to look more often to either of the bottles, as compared to *drinking* and *neutral*. In the neutral context, the drink had a higher number of fixations.

2) Gaze cueing (eye tracking data)

Accumulated time of fixations duration and percentage of fixations revealed that participants tended to look for a longer time and more often at the direction indicated by the robot's gaze, as compared to the opposite direction (see Figure 6). Interestingly, these differences tended to be larger in the congruent, as compared to incongruent conditions. Another interesting observation is that for the incongruent invalid condition there is almost the same percentage of fixations for both directions. This might explain the negative GCE. Taken together, this pattern of results mirrors the reaction time data and reveals that iCub gaze might have effectively guided participants attention.

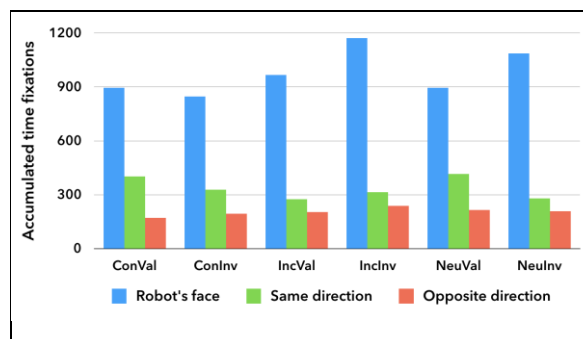


Figure 6. Accumulated time of fixations . The lower panel displays percentage of fixations across conditions.

*ConVal (congruent valid), ConInv (congruent invalid), IncVal (incongruent valid), IncInv (incongruent invalid), NeuVal (neutral valid), NeuInv (neutral invalid)

3) Grasping and handing over (eye tracking data)

As expected, participants gaze behavior seemed to have anticipated robot movements. Fixations and occurred mostly between the robot's face and the bottles, and these areas of interest were also looked at the most time (as indicated by fixation durations). The sequences of fixations alternated between these two areas of interest (see Figure 7).

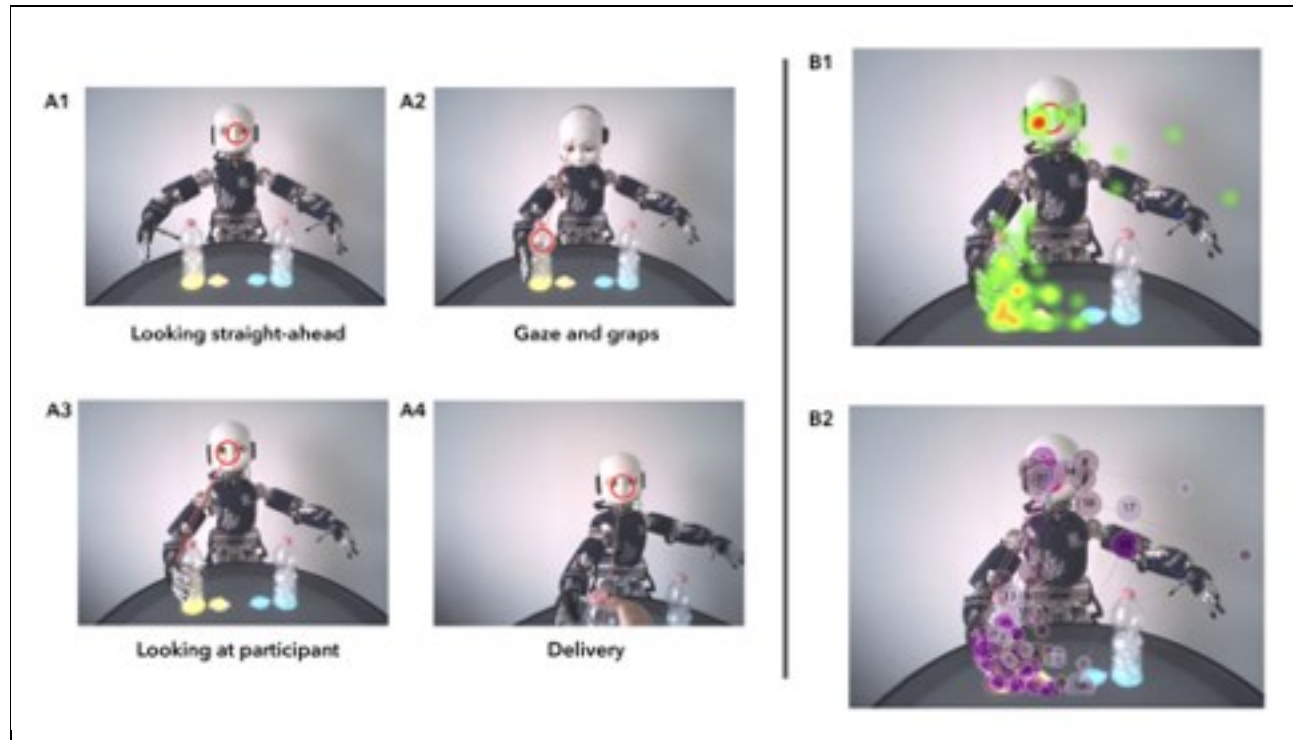


Figure 7. In the left panel, sequence of grasping from the participants point of view. The red circle indicates the fixation point and the red lines the eye movements. We observe alternating eye movements between iCub's face and the object. On the right panel, B1, shows a heat map of fixation durations and B2 shows number and sequence of fixations during the grasping and giving of the bottle located on the left side of the observer. Eye movements of one participants are plotted on a reference image for the purpose of illustration.

IV.DISCUSSION

The present pilot study aimed at validating the feasibility of transferring well-established experimental protocols of cognitive neuroscience to more naturalistic human-robot interaction scenarios that maintain the excellent control of the original designs but add more ecological validity through the interactive nature. We adapted a previous protocol designed in experimental psychology studies to test the gaze-cueing effect. The human performance was evaluated as a function of the social behavior of the iCub robot in a relatively complex context (iCub was requested to reach for either something to drink or for a laundry softener). In detail, we examined whether participants followed the gaze of the robot in a gaze cueing task embedded in the interaction. Additionally, we investigated whether expectations regarding the upcoming behavior of the robot (i.e. gazing at the bottle related with the task of reaching for either something to drink or for laundry softener) would modulate gaze following. Our preliminary results revealed a similar

pattern to previously reported results. That is, in general, participants seemed to have followed iCub's gaze during the task, showing faster responses when the robot gaze was directed to the target position. Likewise, the effect was increased when the robot behavior was in line with expectations regarding action sequences, relative to no social context. Interestingly, when the robot gazed at the incongruent bottle, participants seemed to have attended the object congruent with the action, rather than the robot's gaze.

Eye movement recordings supported the behavioral data, as participants fixated more often, and for a longer time, at locations where the robot gazed. However, they remained attentive to both objects when the robot behavior violated their expectations regarding the upcoming action sequence. Additionally, eye movements revealed that during the exchange of the object, participants looked at the robot face and the bottle, alternating fixations between them. This suggests that participants constantly monitored iCub's face during the task, presumably looking for social signals. Collectively, the data of this pilot validation study suggest that during interactions with a humanoid robot, humans monitor the robot for detecting social signals (gaze/head direction in this case) to efficiently coordinate in joint action.

The main aim of this validation study was to show that it is possible to transfer classical paradigms of experimental psychology/cognitive neuroscience into more naturalistic interaction scenarios, and thereby maintain excellent experimental control with increased ecological validity (relative to the classical paradigms). Translating classical paradigms into a naturalistic human-robot interaction setup is not a trivial task, as it requires integration of contingent and reliable social behavior of a humanoid robot with apparatus used for measuring human behaviour (reaction time data, eye tracking). Together with stimulus presentation software, all these components need to be synchronized with excellent temporal resolution in order to measure human reactions to specific events (robot's gaze/head shift, onset of arm movement, etc.) with millisecond precision. Our validation study was successful in meeting this challenge. However, this validation study was performed only as a proof of concept. Future studies should be conducted with increased sample sizes as well as larger number of trials per condition, in order to draw statistical inferences from the observed data.

V.CONCLUSION

In general, the present study is a successful example of the implementation of an experimental paradigm of cognitive neuroscience in an HRI protocol. We observed a similar pattern of results as in well-documented effects in cognitive neuroscience literature. Such an approach allows casting light on the human social cognitive mechanisms involved HRI through isolating specific mechanisms of human cognition and allowing to address them with appropriate design of the robot behaviors. This should result in the creation of robots that are well-attuned to the workings of the human brain.

REFERENCES

- [1] G. Metta, G. Sandini, D. Vernon, and L. Natale, "The iCub humanoid robot: an open platform for research in embodied cognition," in *Proceedings of the 8th Workshop on Performance Metrics for Intelligent System*, 2008, pp. 50–56.
- [2] S.R.H. Langton, V. Bruce, "Reflexive Visual Orienting in Response to the Social Attention of Others," *Vis. Cogn.*, 1999, pp. 541–67.
- [3] S. Baron-Cohen, "Mindblindness. Learning, development, and conceptual change", 1995, xxii, pp. 171.
- [4] N.J. Emery, "The eyes have it: the neuroethology, function and evolution of social gaze," *Neurosci. Biobehav. Rev.*, Aug 2000, 24(6), pp. 581–604.
- [5] C. Friesen and A. Kingstone, "The eyes have it! Reflexive orienting is triggered by nonpredictive gaze," *Psychon. Bull. Rev.*, 1998.
- [6] A. Frischen, A. P. Bayliss, and S. P. Tipper, "Gaze cueing of attention: visual attention, social cognition, and individual differences," *Psychol. Bull.*, 2007, vol. 133, no. 4, pp. 694.
- [7] J. Driver, G. Davis, P. Ricciardelli, P. Kidd, E. Maxwell, and S. Baron-Cohen, "Gaze perception triggers reflexive visuospatial orienting," *Vis. cogn.*, 1999, vol. 6, no. 5, pp. 509–540.
- [8] H. Admoni, C. Bank, J. Tan, M. Toneva, and B. Scassellati, "Robot gaze does not reflexively cue

- human attention,” *Proc. Annu. Meet. Cogn. Sci. Soc.*, , 2011 vol. 33, no. 33, pp. 1983–1988.
- [9] Y. Okumura, Y. Kanakogi, T. Kanda, H. Ishiguro, and S. Itakura, “Infants understand the referential nature of human gaze but not robot gaze,” *J. Exp. Child Psychol.*, Sep. 2013, vol. 116, no. 1, pp. 86–95.
- [10] H. Admoni, C. Bank, J. Tan, M. Toneva, and B. Scassellati, “Robot gaze does not reflexively cue human attention,” *Proc. Annu. Meet. Cogn. Sci. Soc.*, 2011, vol. 33, no. 33, pp. 1983–1988.
- [11] E. Wiese, A. Wykowska, J. Zwickel, and H. J. Müller, “I See What You Mean: How Attentional Selection Is Shaped by Ascribing Intentions to Others,” *PLoS One*, Sep. 2012, vol. 7, no. 9, p. e45391.
- [12] A. Wykowska, E. Wiese, A. Prosser, and H. J. Müller, “Beliefs about the minds of others influence how we process sensory information,” *PLoS ONE*, 2014, vol. 9, no. 4.
- [13] A. Wykowska, J. Kajopoulos, and K. Ramirez-Amaro, “Autistic traits and sensitivity to human-like features of robot behavior,” *Interact. Stud.*, vol. 16, no. 2, pp. 219–248, 2015.
- [14] M.M. Hayhoe, A. Shrivastava, R. Mruczek, JB. Pelz, “Visual memory and motor planning in a natural task,” *Journal of Vision*, 2003, 3(1), pp.49–63. pmid:12678625
- [15] J.R. Flanagan, R.S. Johansson, “Action plans used in action observation,” *Nature*, 2003, 424(6950), pp. 769–771. pmid:12917683
- [16] B. Mutlu, F. Yamaoka, T. Kanda, H. Ishiguro and N. Hagita, "Nonverbal leakage in robots: Communication of intentions through seemingly unintentional behavior," *2009 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, La Jolla, CA, 2009, pp. 69-76. doi: 10.1145/1514095.1514110
- [17] J. Perez-Osorio, H.J. Müller, E. Wiese, A. Wykowska, “Gaze Following Is Modulated by Expectations Regarding Others’ Action Goals,” *PLoS ONE*, 2015, 10(11): e0143614.
<https://doi.org/10.1371/journal.pone.0143614>
- [18] J. Perez-Osorio, H.J. Müller, A. Wykowska, “Expectations regarding action sequences modulate

- electrophysiological correlates of the gaze- cueing effect,” *Psychophysiology*, 2017, 54, pp. 942–954. <https://doi.org/10.1111/psyp.12854>
- [19] S. Math, D. Schreij, and J. Theeuwes, “OpenSesame: An open-source, graphical experiment builder for the social sciences,” *Behav. Res. Methods*, Jun 2012, vol. 44, no. 2, pp. 314–324.
- [20] G. Metta, P. Fitzpatrick, and L. Natale, “YARP: Yet Another Robot Platform,” *Int. J. Adv. Robot. Syst.*, Mar. 2006, vol. 3, no. 1, p. 8.
- [21] A. Roncone, U. Pattacini, G. Metta, and L. Natale, “A Cartesian 6-DoF Gaze Controller for Humanoid Robots,” *Proceedings of Robotics: Science and Systems*, 2016.
- [22] <https://www.tobiipro.com/product-listing/tobii-pro-glasses-2/>