Robots with Different Embodiments Can Express and Influence Carefulness in Object Manipulation

Linda Lastrico†‡, Luca Garello†‡, Francesco Rea§, Nicoletta Noceti†, Fulvio Mastrogiovanni†, Alessandra Sciutti†, Alessandro Carfi†
†Department of Informatics, Bioengineering, Robotics, and Systems Engineering (DIBRIS), University of Genoa, Italy
‡Robotics, Brain and Cognitive Science Department (RBCS), Italian Institute of Technology, Genoa, Italy
§Cognitive Architecture for Collaborative Technologies Unit (CONTACT), Italian Institute of Technology, Genoa, Italy

Abstract—Humans have an extraordinary ability to communicate and read the properties of objects by simply watching them being carried by someone else. This level of communicative skills and interpretation, available to humans, is essential for collaborative robots if they are to interact naturally and effectively. For example, suppose a robot is handing over a fragile object. In that case, the human who receives it should be informed of its fragility in advance, through an immediate and implicit message, i.e., by the direct modulation of the robot’s action. This work investigates the perception of object manipulations performed with a communicative intent by two robots with different embodiments (an iCub humanoid robot and a Baxter robot). We designed the robots’ movements to communicate carefulness or not during the transportation of objects. We found that not only this feature is correctly perceived by human observers, but it can elicit as well a form of motor adaptation in subsequent human object manipulations. In addition, we get an insight into which motion features may induce to manipulate an object more or less carefully.

Index Terms—Implicit Communication, Object Manipulation, Human-Robot Interaction, Motor Adaptation, Motion Generation

I. INTRODUCTION

To establish an effective interaction between humans and robots, the robot must generate readable and legible movements [1]. As a broad review by Venture and Kulic points out [2], robots can exploit their embodiment to be communicative while performing a task to transmit more information to their partner. In particular, research on implicit communication focuses mainly on two goals: conveying the robot’s intentionality or affective/emotional attitude. Another context where robots’ communicative potential has been explored is the co-ordinary robots if they are to interact naturally and effectively. For example, suppose a robot is handing over a fragile object. In that case, the human who receives it should be informed of its fragility in advance, through an immediate and implicit message, i.e., by the direct modulation of the robot’s action. This work investigates the perception of object manipulations performed with a communicative intent by two robots with different embodiments (an iCub humanoid robot and a Baxter robot). We designed the robots’ movements to communicate carefulness or not during the transportation of objects. We found that not only this feature is correctly perceived by human observers, but it can elicit as well a form of motor adaptation in subsequent human object manipulations. In addition, we get an insight into which motion features may induce to manipulate an object more or less carefully.

Research Question (RQ1): Are human-derived motions enough for the robot to communicate the carefulness required to transport an object?

Research Question (RQ2): Can the carefulness demonstrated by a robot influence how humans handle objects?

Research Question (RQ3): Do different robot embodiments affect the carefulness perceived by humans?

To explore these research questions, we deployed an already proposed solution [10], [11] to generate transportation motions on two different robotic platforms: iCub (a humanoid) by the Italian Institute of Technology [12] and Baxter (a two-arms manipulator) by Rethink Robotics [13]. Due to their different conformation, these two platforms are suitable to...
explore the possible effect of the embodiment (RQ3). Videos recorded with the two robots have been used to investigate (RQ1) through an online questionnaire and (RQ2) with an in-presence experiment.

II. METHODS

We decided to address (RQ1) by preparing a questionnaire including a few videos of Baxter and iCub transporting a blurred object, with what we programmed as a Careful (C) or Not Careful (NC) attitude. Participants were asked to evaluate the movement carefulness, on a five-points Likert scale from “not careful” to “careful” (carefulness score). Moreover, we presented a list of common-use objects and asked how careful we should be in transporting such items on the same scale. This is to clarify the concept of carefulness and to understand which attributes make an object more or less delicate to carry from the participant’s perspective.

Concerning (RQ2), we opted for an in-presence experiment. In this case, participants transported a sensorized cube after watching videos of one of the robots carrying the same item, again either with a C or NC attitude. The experiment allowed us to investigate if participants, with no prior knowledge of the robots’ attitude, would be influenced by their movement to the point of modulating their kinematics.

Both the experiments, relying on videos realized with different robotics platforms, will also allow quantifying (RQ3). In the following part, we will give a short overview of how we generated iCub and Baxter movements. Then, we will describe the questionnaire. Finally, we will present the in-presence experiment setup.

A. Generating communicative movements on robots

Previous literature showed that human movements naturally embed implicit information that allows an observer to deduce additional properties of the action being witnessed. For instance, information on the weight or the carefulness of the handled object can be deduced from the movement velocity [5], [6], [8], [9], [15]. For this reason, in previous studies [10], [11], we focused on generating communicative robots’ movement by modulating the velocity of their end-effector. Our goal was not to simply replicate human kinematics, but to exploit Generative Adversarial Networks to produce novel and meaningful velocity profiles associated with a careful, or not so careful, manipulation. We will now explain how, starting from one desired velocity profile, we controlled the robots to follow it.

1) Baxter: We used MoveIt!, a ROS motion planning framework widely used by several robotics systems [16], to control Baxter. We remapped the selected velocity profile to meet the intended trajectory length, keeping constant the temporal duration of the action while modulating the magnitude of the velocity. Due to the MoveIt! controller implementation, we used a fixed spatial step to cover the trajectory while modulating the temporal execution. Hence, MoveIt! planned a trajectory defined by a sequence of poses in the joints space. We imposed the time with which each spatial point was supposed to be reached to determine the velocity norm of each spatial step, mapping the desired velocity profile. Baxter trajectories had a range of 60 cm.

2) iCub: We controlled the iCub robot with its default “Cartesian Controller” [17]. However, differently from Baxter, on iCub, the end-effector controller required to fix the execution time of each sub-step, varying the Euclidean distance between them. Due to its kinematic configuration, iCub trajectories were shorter, around 20 cm, and, since the movement’s original duration was preserved, slower than Baxter ones. To prepare the stimuli for the questionnaire and the in-presence experiment, we recorded videos of iCub and Baxter moving the same object, a green cube, along the frontal plane. In particular, we used 2 distinct velocity profiles per condition (C/NC) per robot, resulting in 8 videos for the online questionnaire, while 6 videos for each combination for the experiment, resulting in 24 different videos. The movements in the videos where characterized by a maximum velocity reached by the end-effector in the Not Careful condition of $0.281 \pm 0.037 \text{ m/s}$ for iCub, and of $0.871 \pm 0.080 \text{ m/s}$ for Baxter. In the Careful videos instead, iCub reached a maximum velocity of...
The object they had to move was the iCube, a green sensorized cube of side 5 cm [18]. Each participant was asked to read the same instructions, which were carefully designed so as not to introduce bias in the experiment. In particular, we explained that they would watch videos of the robots transporting the iCube. We used several prototypes of this cube, aesthetically indistinguishable but more or less delicate depending on their assembly state and internal components. At this point, we showed a box containing three iCubes to be used for the experiment. Even though we actually used only one of them, the purpose of having three was to avoid participants assuming a priori a binary interpretation of the videos (C or NC, even though this was exactly what we implemented in the robot actions). At the end of each trial, the participants were asked to close their eyes, for the experimenter to place a new cube in the starting position without influencing them with how the cube was moved. A sound informed the subject to open their eyes and prepare for watching a new video. We explained that the purpose of the experiment was to simulate a collaboration between the robot in the video and the participant. So, they were asked to imagine that, in every trial, the cube on the table was the one manipulated by the robot in the video, who virtually deposited it on the table for them to grab. It should be noted that to avoid a simple copy of the observed movement, while the robots’ actions occurred in the frontal plane, the participants moved the iCube in the sagittal direction. Before starting with the video reproduction, we ran a baseline of 6 trials with no stimuli, for the participants to familiarize with moving the cube from the starting position to the ones in A or B. To precisely record the hand kinematic during the transport of the cube, we used active infrared markers from the Optotrak Certus®, NDI, motion capture (MoCap) system, which gave the 3D position of each marker with respect to a common reference frame, with a frequency of 100 Hz. We positioned four markers on the metacarpal bones of the right hand and considered for further analyses the most visible one in each trial. For the purpose of this study and the sake of space, we report the results of the analysis of the kinematics data coming from the motion capture system and not those recorded through the iCube.

### III. Results

To analyse the acquired data, we used Jamovi software [20], in particular the GAMLj module [21] for mixed models.

#### A. Online questionnaire

To assess if the implicit information embedded in the robots movements is communicative of their carefulness, we first evaluated the responses regarding the perceived carefulness in the videos observed by the participants. Considering that two videos were presented for each robot in the C and NC condition, each participant evaluated 8 videos, resulting in a total of 392 carefulness scores, ranging from 1 (NC) to 5 (C). From these data, we ran a mixed model assuming the Carefulness Score as dependent variable, the robot type, the
condition and their interaction as factors, and the subjects as cluster variables. The resulting metrics are shown in Figure 3. The effect of condition resulted significant ($NC - C$, estimate = $-1.041$, $SE = 0.0878$, $t = -11.85$, $p < 0.001$) as well as the effect of the robot ($iCub - Baxter$, estimate = $0.429$, $SE = 0.0878$, $t = 4.88$, $p < 0.001$). Moreover, also the interaction between robot and condition is significant ($Careful \cdot Baxter$, estimate = $0.531$, $SE = 0.1756$, $t = 3.02$, $p = 0.003$). These results indicated that the videos showing what we planned as Careful movements were significantly perceived as more careful, getting a score higher by 1.041. Also, iCub was generally perceived as more careful than Baxter. The carefulness score attributed to Baxter videos covered a wider range, with a stronger difference between the C and NC condition (see Figure 3). For both the robots, the Not Careful movements were perceived as rather neutral on the Likert scale.

Regarding the carefulness score associated with the proposed 12 items, the results are shown in Figure 4. The answers allowed to create a scale going from the objects which were considered not critical to be transported, such as the rubber ball, to the one which requires the highest attention in the manipulation, that is the glass of water full till the brim. Bringing together the results of the carefulness scores attributed to the robots’ videos, in Figure 3 and to the proposed objects, in Figure 4 we could evaluate for which items on the list the robots’ movements might be considered suitable. On average, none of the robots’ movements seem adequate to move the items evaluated as delicate as the Lit Candle, onwards. iCub NC movements obtained a Carefulness Score which could match the Scissors and Pair of Glasses. Even though we did not ask specifically about the item, we can posit that iCub NC movements are suitable, in general, to manipulate objects with an equivalent fragility. Following the same reasoning, Baxter NC attitude seem suitable for those objects which obtained a neutral evaluation, as the Origami.

B. In-presence experiment

From the results presented in the previous part, we verified that the movements produced with iCub and Baxter were indeed perceived as more or less careful, coherently with what we planned. The aim of the in-presence experiment was to go a step further and to investigate whether the carefulness modulation of robot motion was sufficient to affect how participants performed their own movements. To answer this question, we computed the hand velocity norm from the components of the velocity, obtained by deriving the three-dimensional position of the considered marker. As metric to quantify the modification of the human movement we considered the maximum velocity reached and the duration of
the transport movement performed after watching the video. Indeed, previous studies showed that the velocity peak is a significant measure to distinguish between careful and not careful movements \cite{9,11}; in particular, careful movements are characterized by lower velocities and longer durations.

In the in-presence experiment, six videos were presented for each robot in the C and NC condition, for a total of 24 trials for each of the 11 participants. We ran a mixed model assuming the magnitude of the velocity peak as dependent variable and the robot type, the condition and their interaction as factors, and the subjects as cluster variable. The effect of condition resulted significant (\( NC - C, \) \( estimate = 0.0841, SE = 0.0127, t = 6.60, p < 0.001 \)) as well as the effect of the robot (\( i\text{Cub} - \text{Baxter}, \) \( estimate = -0.0608, SE = 0.0128, t = -4.77, p < 0.001 \)). The interaction between robot and condition was instead not significant (\( \text{Careful} \cdot \text{Baxter}, \) \( estimate = -0.0266, SE = 0.0255, t = -1.04, p = 0.297 \)). The results are shown in Figure 5. It can be noticed that the transport movements performed after a careful video stimuli have indeed lower peaks of velocity, with an estimate of the velocity reduction equal to 0.841 cm/s. There is also an effect of the robot, with the videos of iCub soliciting slower movements, with an average reduction in the peak speed of -0.608 cm/s, respect to Baxter. This is in agreement with what we found from the questionnaires, where iCub movements got a higher carefulness score (see Figure 3). For the duration of the movement, we ran the same mixed model using the temporal duration as dependent variable. Again, we found a significant effect of the condition (\( NC - C, \) \( estimate = -0.1796, SE = 0.0233, t = -7.70, p < 0.001 \)) and of the robot (\( i\text{Cub} - \text{Baxter}, \) \( estimate = 0.0886, SE = 0.0233, t = 3.80, p < 0.001 \)), not of their interaction (\( \text{Careful} \cdot \text{Baxter}, \) \( estimate = 0.0758, SE = 0.0466, t = 1.63, p = 0.105 \)). When moving the cube from the starting position to the final one, the covered distance was always of about 30 cm, so the time difference detected depends on how the participants modulated their movement speed. Confirming our hypothesis, after observing careful robot manipulations, participants moved the cube in a slower way, with an estimate increase of duration of 179.6 ms. Analogous considerations were verified as well for other kinematics metrics, such as the mean velocity of the transport or its maximum acceleration.

As an additional control, we also compared the movement duration and the velocity peak of the transport after the video stimuli with those of the cube manipulation in the initial baseline, when no videos were presented and participants performed a simple pick and place of the cube. We ran two mixed models, one for each kinematic metrics as dependent variable, the condition (baseline, C or NC) as factor, and the subjects as cluster variable. The results for the velocity
peak metrics are shown in Figure 6. For both the metrics, we obtained a significant difference between the Baseline and the Careful condition (Movement duration, $C - Baseline$: estimate $= 0.1296$, $SE = 0.0282$, $t = 4.59$, $p < 0.001$; Velocity peak, $C - Baseline$: estimate $= -0.0644$, $SE = 0.0163$, $t = -3.94$, $p < 0.001$). This shows that when the robots displayed an attentive attitude, the participants moved significantly slower than they would have naturally. Instead, the models did not evidence a significant difference between the Baseline and the NC condition, even though, as shown in Figure 3, participants produced slightly faster and shorter movements than during the baseline (Movement duration, $NC - Baseline$: estimate $= -0.0497$, $SE = 0.0282$, $t = -1.76$, $p = 0.08$; Velocity peak, $NC - Baseline$: estimate $= 0.0196$, $SE = 0.0163$, $t = 1.20$, $p = 0.231$).

IV. DISCUSSION

In the proposed work, we aimed to answer three main research questions: (RQ1) assess if the implicit information embedded in the robots’ movements is perceived; (RQ2) evaluate the potential modulation produced by the robot movements on how humans handle objects and, (RQ3), investigate whether the robots’ embodiment affects how their actions are perceived.

Even though the idea of carefulness associated with object manipulation has already been explored, it is difficult to give an unequivocal definition of the term. For instance, unlike the weight, carefulness cannot be measured with a defined scale. From the questionnaire answers in Figure 4, about grading the carefulness required for manipulating common-use objects, what emerges is that the concept of carefulness does exist and it is shared among people. Moreover, the emerging gradation correlates with distinct attributes of the items, which may contribute to making an object delicate to handle: the content (glass with water), the material (crystal), the preservation of the object state (lit candle), the value (pair of glasses) or the potential danger (scissors). The questionnaires outcomes positively confirmed that the movements we produced, with both the robotic platforms, were indeed communicative of the desired feature, that is the presence or absence of a careful attitude in the action (see Figure 3). Interestingly, we found an effect of the robot: iCub was perceived as generally more careful in its movements, while Baxter presented a more marked difference between the two conditions. The same result also emerged when comparing the kinematics metrics of the in-presence experiment, presented in Section II-B: iCub videos solicited slower actions by the participants, i.e., more careful manipulation of the cube they found on the table. We can ascribe this to two main factors: how the movements were generated and the embodiment of the robots. As explained in Section IIA and in [10] with more detail, when controlling the robot end-effector to follow the desired trajectory with a velocity profile of choice, in our implementation, the duration of the original movement is preserved; this means that for longer trajectories the velocity of the movement needs to be higher, whereas for shorter paths the original velocity profile is stretched. The different sizes of the two robots cause the trajectories covered by iCub to be shorter and lower the speed of its movements. iCub’s appearance is more human than Baxter’s, not only for the hand shape, but also for the kinematic configuration of its arm; this could lead people to identify themselves more with its movements and perceive them, regardless, as more delicate. However, the difference between the C and NC conditions in both robots was significant for the carefulness score attributed to the videos and for the kinematic measures of motor contagion. Hence, even with a platform not designed to have an anthropomorphic shape - such as Baxter - a bioinspired modulation of movement speed effectively communicates motion carefulness. Finally, Figure 6 allows to understand better how a C or NC movement is modulated with respect to standard baseline actions. In particular, moving objects in a not careful manner seems to be the natural way of transporting them, whereas we change the movement kinematics when we want to show caution. This seems reasonable to optimize the actions’ efficiency and should also be considered when programming robots’ behavior.

V. CONCLUSION

This study showed that the movements we generated, associated with object manipulation, are indeed perceived as characterized by different degrees of carefulness for both the robots iCub and Baxter. In particular, the movements we implemented were rated as going from a rather careful attitude to a neutral one. Indeed, what we labeled as Not Careful movements, were perceived as neutral. This comes not by surprise, as confirmed by the comparison with the baseline movements since in everyday life, we tend to act as not careful when manipulating objects, and the kinematics modulation emerges when we are challenged with a difficult
task. Moreover, the robots’ gestures proved to be efficient in soliciting a motor change in the participants: the simple request to imagine a collaboration with the robot, such as virtually carrying the same object, was enough to bring out a motor adaptation, even along a different direction with respect to the plane where the robot action took place. In a real one-to-one human-robot collaboration, the modulation of human movements in response to the robot ones could become even stronger.

Future studies should assess whether the perception of iCub as more careful than Baxter depends on their embodiment or on the different velocities their end-effectors reached in the transport movements. Finally, from the list of items provided in the questionnaires, we plan to evaluate in a more granular manner the features - and their combination - that allows an item to be classified as careful or not to be handled. This to arrive at an explicit definition of the concept of carefulness, aware that an implicit and shared understanding of it already exists.

REFERENCES

[1] A. D. Dragan, K. C. T. Lee, and S. S. Srinivasa, “Legibility and predictability of robot motion,” in Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction, Tokyo, Japan, Mar 2013.

[2] G. Venture and D. Kulić, “Robot expressive motions: A survey of generation and evaluation methods,” J. Hum.-Robot Interact., vol. 8, no. 4, nov 2019.

[3] N. T. Viet Tuyen, A. Elibol, and N. Y. Chong, “Learning from humans to generate communicative gestures for social robots,” in 2020 17th International Conference on Ubiquitous Robots (UR), 2020, pp. 284–289.

[4] M. Salem, S. Kopp, I. Wachsmuth, K. Rohlffing, and F. Joublin, “Generation and evaluation of communicative robot gesture.” International Journal of Social Robotics, vol. 4, pp. 201–217, 04 2012.

[5] A. Hamilton, D. Joyce, J. Flanagan, C. Frith, and D. Wolpert, “Kinematic cues in perceptual weight judgment and their origins in box lifting,” Psychological research, vol. 71, pp. 13–21, Feb 2007.

[6] G. Bingham, “Kinematic form and scaling: further investigations on the visual perception of lifted weight,” Journal of experimental psychology. Human perception and performance, vol. 13, no. 2, p. 155–177, May 1987.

[7] J. Shim, H. Hecht, J.-E. Lee, D.-W. Yook, and J.-T. Kim, “The limits of visual mass perception,” Quarterly Journal of Experimental Psychology, vol. 62, no. 11, pp. 2210–2221, 2009.

[8] A. Sciutti, L. Patané, and G. Sandini, “Development of visual perception of others’ actions: Children’s judgment of lifted weight.” PLOS ONE, vol. 14, pp. 1–15, Nov 2019.

[9] N. F. Duarte, K. Chatzilygeroudis, J. Santos-Victor, and A. Billard, “From human action understanding to robot action execution: how the physical properties of handled objects modulate non-verbal cues,” in Proceedings of the 10th International Conference on Development and Epigenetic Robotics, Valparaiso, Chile, Oct 2020, pp. 1–6.

[10] L. Garello, L. Lastrico, A. Sciutti, N. Noceti, F. Mastrogiavanni, and F. Rea, “Synthesis and execution of communicative robotic movements with generative adversarial networks,” 2022, Unpublished, submitted to the Special Issue on Emerging Topics on Development and Learning, IEEE TCDS. [Online]. Available: https://arxiv.org/abs/2203.15640

[11] L. Garello, L. Lastrico, F. Rea, F. Mastrogiavanni, N. Noceti, and A. Sciutti, “Property-aware robot object manipulation: a generative approach,” in 2021 IEEE International Conference on Development and Learning (ICDL), 2021, pp. 1–7.

[12] C. Fitzgerald, “Developing Baxter,” in 2013 IEEE Conference on Technologies for Practical Robot Applications (TePRA). IEEE, 2013, pp. 1–6.

[13] J. Li, “The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents,” International Journal of Human-Computer Studies, vol. 77, pp. 23–37, 2015.

[14] L. Lastrico, A. Carfì, F. Rea, A. Sciutti, and F. Mastrogiavanni, “From movement kinematics to object properties: Online recognition of human carefulness,” in Social Robotics, 2021, pp. 61–72.

[15] D. Coleman, I. Sucan, S. Chitta, and N. Correll, “Reducing the barrier to entry of complex robotic software: a MoveIt! case study,” arXiv, 2014.

[16] U. Pattacini, F. Nori, L. Natale, G. Metta, and G. Sandini, “An experimental evaluation of a novel minimum-jerk cartesian controller for humanoid robots,” in 2010 IEEE/RSJ international conference on intelligent robots and systems. IEEE, 2010, pp. 1668–1674.

[17] A. Sciutti and G. Sandini, “The role of object motion in visuo-haptic exploration during development,” in 2019 Joint IEEE 9th International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob), 2019, pp. 123–128.

[18] G. Lombardi, J. Zouzen, G. Belgiovine, F. Vannucci, F. Rea, A. Sciutti, and G. Di Cesare, “The influence of vitality forms on action perception and motor response.” Scientific Reports, vol. 11, 2021.

[19] The Jamovi Project, 2021, Jamovi (Version 1.6). [Online]. Available: https://www.jamovi.org/.

[20] M. Gallùcci, GAMLJ: General analyses for linear models, 2019. [Online]. Available: https://gamlj.github.io/