

# If You Are Careful, So Am I! How Robot Communicative Motions Can Influence Human Approach in a Joint Task

Linda Lastrico<sup>1,2,\*</sup>, Nuno Ferreira Duarte<sup>3</sup>, Alessandro Carff<sup>2</sup>, Francesco Rea<sup>1</sup>, Fulvio Mastrogiovanni<sup>2</sup>, Alessandra Sciutti<sup>4</sup>, and José Santos-Victor<sup>3</sup>

<sup>1</sup> Robotics, Brain and Cognitive Science Department (RBCS), Italian Institute of Technology, Genoa, Italy

<sup>2</sup> Department of Informatics, Bioengineering, Robotics, and Systems Engineering (DIBRIS), University of Genoa, Genoa, Italy

<sup>3</sup> Institute for Systems and Robotics (ISR), Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

<sup>4</sup> Cognitive Architecture for Collaborative Technologies Unit (CONTACT), Italian Institute of Technology, Genoa, Italy

**Abstract.** As humans, we have a remarkable capacity for reading the characteristics of objects only by observing how another person carries them. Indeed, how we perform our actions naturally embeds information on the item features. Collaborative robots can achieve the same ability by modulating the strategy used to transport objects with their end-effector. A contribution in this sense would promote spontaneous interactions by making an implicit yet effective communication channel available. This work investigates if humans correctly perceive the implicit information shared by a robotic manipulator through its movements during a dyadic collaboration task. Exploiting a generative approach, we designed robot actions to convey virtual properties of the transported objects, particularly to inform the partner if any caution is required to handle the carried item. We found that carefulness is correctly interpreted when observed in the robot movements. In the experiment, we used identical empty plastic cups; nevertheless, participants approached them differently depending on the attitude shown by the robot: humans change how they reach for the object, being more careful whenever the robot does the same. This emerging form of motor contagion is entirely spontaneous and happens even if the task does not require it.

**Keywords:** Communicative robots' movement · Implicit Communication · Generative Adversarial Networks · Robots' motion generation · Objects' manipulation · Human-Robot Interaction.

---

\* Corresponding author [linda.lastrico@iit.it](mailto:linda.lastrico@iit.it)

This paper is supported by the European Commission within the Horizon 2020 research and innovation program, under grant agreement No 870142, project APRIL (multipurpose robotics for mAniPulation of defoRmable materIaLs in manufacturing processes). N. F. Duarte is supported by FCT-IST fellowship grant PD/BD/135116/2017.

## 1 Introduction

Humans routinely engage in joint actions and coordinate their movements with others, e.g. working together, playing a team sport, or merely moving objects. These tasks involve a collaborative process to coordinate attention, communication, and actions to achieve a common goal. During this process, humans observe the behavior of their partners to anticipate their actions and plan their own accordingly. Verbal communication is not the only means to express intentions. Since verbalizing every step of the interaction would be time-consuming and cognitively expensive, humans also exploit their bodies and movements to exchange information. While executing an intended action, we also implicitly communicate to others our goal, its urgency, and the required effort. This ability is referred to as non-verbal communication (i.e., non-verbal cues), and it can be expressed with our body: from turning the head or torso to a simple eye movement.

In ordinary life, humans are very proficient at monitoring different components of other people’s kinematics, which they leverage to disclose hidden qualities of a handled item. For instance, studies on human non-verbal cues found that joints kinematics and dynamics of hand manipulation are crucial features to estimate the weight of a manipulated object [1,16,19] or predicting action duration [7].

Given the importance of implicit cues in human-human communication, we believe it should be taken into account also in the robotic field. A robot meant to interact with humans, able to exploit the same communication channels as the partner, would guarantee a natural and less cumbersome experience. Indeed, numerous channels of communication may be employed to convey information between robots and people (such as synthetic speech, light-based, digital display, mixed or augmented reality [5,15,13]). However, a valuable alternative that does not require any training or explicit instruction is mediated by movement, and it should be sought whenever feasible. Human non-verbal cues from eyes, head, and arm movements encode the intention driving the action; when such cues are embedded onto a robot, they similarly allow to read the robot’s intention [4]. Regarding object manipulation, object affordances was popularized in robotics and linked to (i) the action associated with the object, (ii) a physical property, or (iii) the type of behaviour required to manipulate the object [18,9]. Specifically, works on affordance reasoning examine the object’s properties [21,8,23], e.g., how to infer the water level in cups [14], although trying to directly detect such property from the object appearance, making it only possible with transparent cups and glasses.

To overcome the need to understand the properties of objects from their appearance alone, it is relevant to quantify the effect their features have on the kinematics of the action during manipulation. In our previous works, we exploited human kinematics to infer the impact of cup water content on human motion, irrespective of the cup’s transparency [3,11]. Indeed, it has been shown that humans alter their behavior, adapting to the properties of the object they transport, such as weight, fragility, or content. Additionally, depending on the

type of cup, these behaviors may be more predominant or less, which may indicate that the difficulty of the manipulation impacts the human motion [17]. Knowing from the mentioned studies that humans reveal some object properties through movements, in this work, we investigate if it is possible to modulate the movements adopted by a robot end-effector during the transport of an object to communicate some of its hidden properties. In a previous study [12], we assessed the communicative potential of movements on different humanoid robots, by asking participants to explicitly judge the robotic motion’s carefulness after observing it in videos. This study proposes a dyadic interaction with a new robotic manipulator in a realistic collaborative context. We used Generative Adversarial Networks (GANs) to synthesize and design the robot movements to convey a particular style feature associated with object manipulation: carefulness [6]. By using a generative approach, we can consistently produce novel but meaningful robot actions. In this study, we explore (i) whether the attitude conveyed by the robot’s movements is perceived as expected, i.e., if the carefulness (or its absence) is correctly expressed by our controller, and (ii) if a robot transporting objects and expressing the appropriate human-like behavior can invoke motor adaptation in the human response.

## 2 Materials and Methods

The objective of our study is to assess whether the generated robot’s movements are informative of the properties of the transported object. Moreover, we evaluate if the robot behavior affects how humans perform their tasks.

We will now explain how we synthesized the required velocity profiles and controlled a Kinova Gen3 robot with 7 degrees of freedom to execute them; then, we will describe the experimental setup and design.

### 2.1 Generation of robot movements

To create robot’s movements communicative about the object properties, we used Generative Adversarial Networks (GANs) and controlled the robot end-effector to follow the velocity profiles produced by such model. Our interest is in generating movements to convey whether the transported object requires caution and care to be transported (*careful* movement) or it is safe to move without any particular concern (*not careful* movement). Previous studies assessed this kind of object manipulation and showed a marked difference in the kinematics of the human hand associated with the two classes of motions [11,3]. The velocity profile is mainly affected, where actions associated with delicate objects, e.g., a cup full of water, are characterized by lower maximum velocity, prolonged deceleration phase a longer duration. This modulation is so marked that it is possible to discriminate between careful movement and not automatically [3,10].

**GAN** To define meaningful movements associated with the properties of the carried object, we decided to modulate the velocity profile adopted by the robot

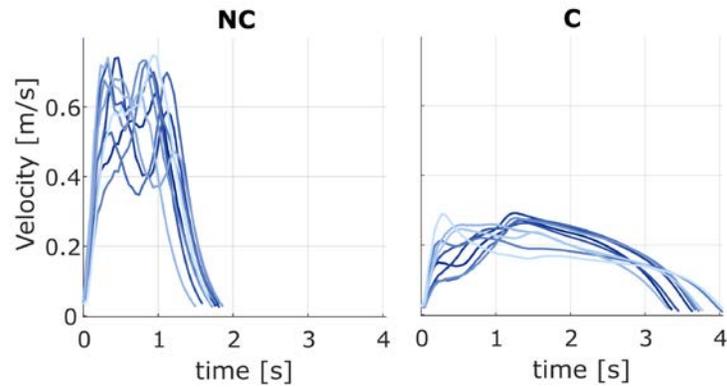


Fig. 1: Velocity profiles generate by the GANs for associated to Not Careful (NC) and Careful (C) transportation of objects. These velocity profiles were used to control the robot and provide the stimuli to the participants during the interaction experiment

end-effector. We used Generative Adversarial Networks to synthesize novel velocity profiles using an approach already tested [6]. The details on the Time-GAN model [22] and its training are described in [6]. The original data used to feed the GANs consisted of hand velocity profiles recorded with a Motion Capture System during the transport of glasses, either empty or containing water, at two possible weight levels. The trajectories followed by participants during the manipulations were designed to grant a good degree of variability. After training, each generative model can produce novel yet meaningful velocity profiles belonging to the distribution of the human data used during the training. This approach provides new and unlimited synthetic data, always falling in the desired class of motion (careful or not), avoiding a trivial copy of the human velocity profiles. Moreover, learning the velocity norm is useful in generalization terms since the same pattern of motion can be applied to multiple spatial trajectories. For this specific study, from the trained GANs, we synthesized ten velocity profiles for each of the two classes to be replicated by the Kinova robot. A representation of the generated data is available in Figure 1.

**Robot Controller** The Kinova Gen3 robot is controlled using ROS and the package `kortex_ros`<sup>1</sup>. Such package provides a velocity controller in Cartesian space, which moves the end-effector at 40 Hz in linear (m/s) and angular (rad/s) velocities. Attached to the end-effector is the Robotiq 85 two-finger gripper<sup>2</sup> used to grasp the cups. This work applies two high-level controllers: (i) a velocity PI

<sup>1</sup> Official repository of the Kinova Gen3 ROS package: [https://github.com/Kinovarobotics/ros\\_kortex](https://github.com/Kinovarobotics/ros_kortex)

<sup>2</sup> Official website of the gripper: <https://robotiq.com/products/2f85-140-adaptive-robot-gripper>

controller and (ii) a velocity GAN controller. The first controller is responsible for picking the cups from the table, and the second is for transporting and handing over the cups to the participant. The former generates a constant velocity profile throughout the trials, while the latter follows one of the 20 GAN velocity profiles selected (10 careful and 10 not careful) during the experiment. For each GAN motion trajectory, the velocity profile is decomposed into the 3D Cartesian velocity coordinates by setting the current location and final location (handover point) at each time step. The handover location was fixed in advance to avoid any variability that could influence participants during the experiment. The position of the participant’s wrist was tracked with a motion capture system, and the position of the robot gripper was estimated the same way (more details on the sensors used are stated in Section 2.2). The handover release moment was obtained by applying a threshold: the robot opened the gripper to release the cup whenever the distance between its end-effector and the participant’s wrist was below a fixed value. This simple design was enough to grant a smooth and reactive handover required for our experiment.

## 2.2 Setup, sensors and experiment design

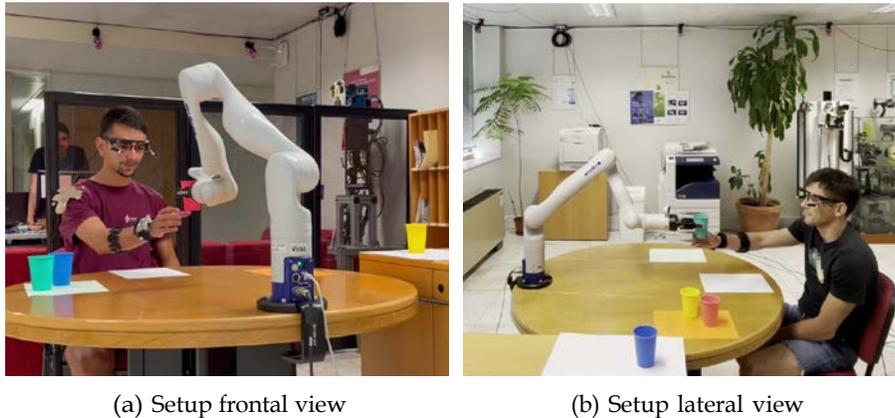


Fig. 2: *Setup*: when interacting with the Kinova Gen3 robot, participants seated at a table. Once grasped the cup from the robot gripper, they had to put it down on one of the three areas delimited on the table. The motion capture markers used to analyse the human kinematics are visible on the participants’ right wrist

Participants were asked to sort the items handled by the Kinova Gen3 robot by positioning them on the appropriate areas marked on the table where they were seated. We designed the experiment for the participants to focus on the robot behaviour and not on the characteristics of the items. For this reason, we

used identical plastic cups: in the instructions, we explained that we were simulating a bar-like scenario, where the robot and the human had to collaborate in sorting the glasses between those full to be served to the clients, and the used and empty ones, to be washed; in such context, the cups were meant to be either full of a liquid or empty: however, we explained to the participants that due to the danger of having a robot transporting water, all the cups were empty. This granted that participants could not rely on any visual cue or the actual object features to decide where to place the cup. In every trial, the Kinova robot grasped a cup from the table next to it (see Figure 2b) and transported it towards the participant, following either a careful (or not) velocity profile generated by the GANs (associated respectively, to the transport of a full or empty glass). The task for the participants was then to grasp the cup from the robot gripper and place it in the appropriate area on the table: on the “To be served” area, on the right, if they thought that the cup was actually meant to be full, or on the “To be washed” area, in case they assumed the cup was indeed empty. A third area, in the middle, was available to place the cups whose virtual content was not clear to the participant to avoid forcing them into making a decision. They were not informed about the modulation of the robot transport movements and, since the cups were all the same, they had to rely on the robot behavior to make their decision<sup>3</sup>.

We used Optitrack<sup>4</sup> motion capture system, with an acquisition frequency of 120 Hz, to track the position of the participant wrist and shoulder.

Twelve healthy participants, all members of Instituto Superior Técnico, voluntarily took part in the experiment. Each evaluated 20 robot movements, where the sequence of careful and not careful modulation was randomized once and then maintained for every participant. The interactions were organized as five blocks of a sequence of four trials. At the end of each block, the experimenter put the cups on the table next to the robot. This resulted in a total of 240 movements evaluated, equally balanced between careful or not robot behavior.

### 3 Results

One of the aims of this study was to verify whether modulating the robot end-effector velocity to express carefulness can inform participants about the virtual content of the manipulated glasses. Figure 3 shows the participants’ accuracy in evaluating, for each trial, if the observed transportation motion was meant to be associated with a delicate object, i.e., careful robot movement, or not. We represented with a dark bar the percentage of correct answers given by the participants, i.e., when they correctly interpreted the robot’s attitude. Considering the total number of evaluated trials (240), 189 were correctly classified with no indecision, resulting in an accuracy of 78.75%. The transparent colored bars represent the misclassified movements. For instance, when we generated a robot

<sup>3</sup> Sample video of the human-robot interaction: <https://www.youtube.com/watch?v=HVahS-0tn6g>

<sup>4</sup> Optitrack website: <https://optitrack.com/cameras/flex-13/>

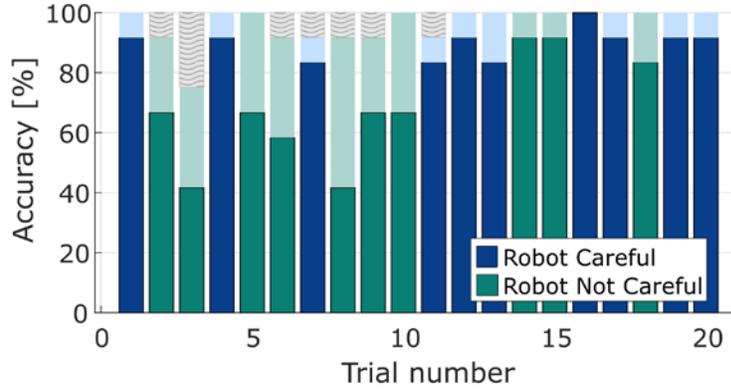


Fig. 3: *Perception of robot's movements*: Percentage of correct interpretation of the robot's transportation movements during the experiment. When the robot performed careful movements, in blue, they were correctly perceived 90% of the times. NC motions required more trials to be consistently classified. The dark bars represent the percentage of correct classification of the movement from the participant, the transparent bars the percentage of wrong attribution; finally, the light gray bars with a wavy pattern, the percentage of "Unknown" answers in each trial

action modulated to communicate a not careful attitude, while participants associated it with the transport of a full cup. As it can be noticed, misunderstanding a not careful action for a careful one was the most frequent occurrence, especially in the first trials. In detail, 90% of the careful robot movements were perceived as such, whereas 75% of the not careful ones were correctly interpreted. Finally, the grey bars with a wavy pattern represent those trials where participants preferred not to make a choice and placed the cup on the neutral area on the table. Also, these occurrences, which happened in 9 trials out of 240, decrease as the experiment progresses.

Another aspect we were interested in investigating is if the two attitudes shown by the robot had any effect on how participants performed their tasks. An exploratory inspection of the hand velocity data encouraged us to deepen this intuition: Figure 4 reports an example of the velocity adopted by one participant when reaching for the cup in the robot gripper. There is a noticeable modulation in the participant's movements that correlates with the attitude shown by the robot. When the robot handled the cup with a careful attitude, also the participant reached for it with slower and prolonged action compared to the not careful situation. To assess this modulation quantitatively in human actions, we considered the duration and median velocity of the movements as relevant features. To perform statistical analyses on the acquired data, we used

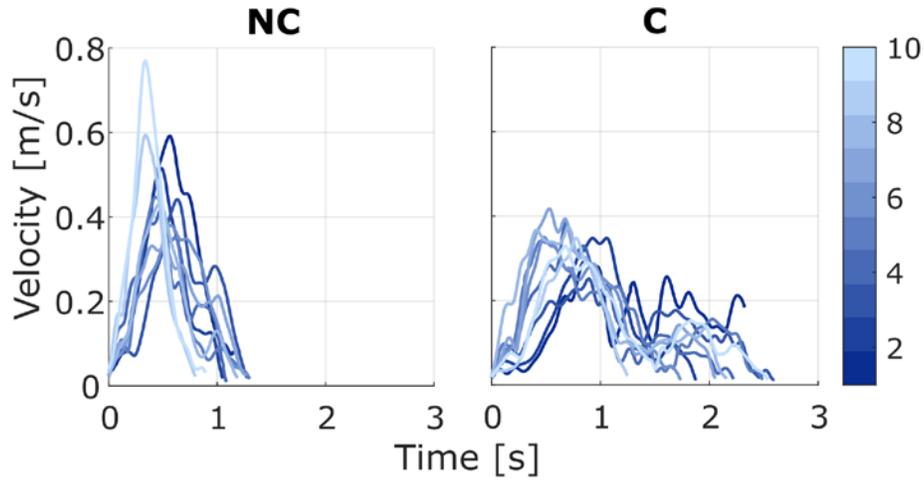


Fig. 4: *Velocity reaching movement*: profiles adopted by one participant when reaching for the cup in the robot gripper. It is noticeable a modulation of both the duration and the maximum values depending on the style of the movement adopted by the robot: not careful (NC) or careful (C). The colormap is associated to the trial numbers, in order

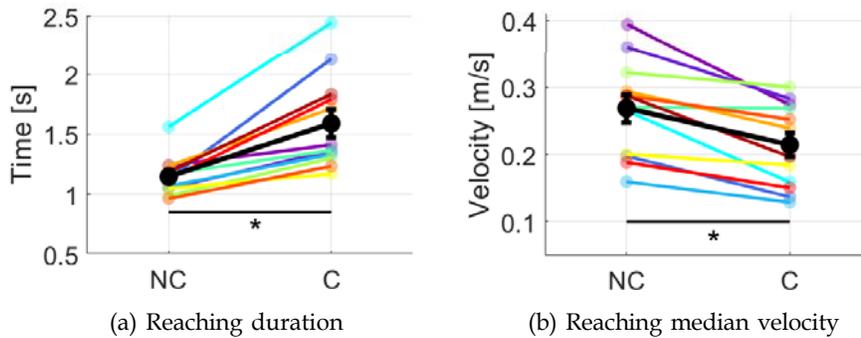


Fig. 5: *Reaching movement*: in (5a) mean duration of the participants reaching movements towards the robot's gripper. When the robot performs a Careful (C) transportation movement, participants are significantly slower in reaching for the cup. Also the median velocity adopted in the reaching movements (5b) is modulated by how the robot moved in the transport action. The mean values for each participant are represented in a different color. The thick black lines represent the mean over the twelve participants, with the standard error. The star indicates a significant difference with  $p < 0.001$

Jamovi software<sup>5</sup>, in particular the GAMLj module<sup>6</sup> for mixed models. Figure 5a shows the mean durations of the participants reaching movements toward the robot gripper. We ran a mixed model assuming the duration of the participants' reaching movements as the dependent variable, the carefulness in the robot movement as a factor, and the subjects as cluster variables. The effect of condition resulted significant ( $C - NC$ ,  $estimate = 0.443$ ,  $SE = 0.055$ ,  $t = 8.00$ ,  $p < 0.001$ ), indicating that when the robot end-effector was following a careful velocity profile, the subsequent human reaching action was longer, with an extended duration estimate of 0.443 seconds. A second mixed model was used to evaluate the median velocity adopted by the participants when reaching the robot gripper (see Figure 5b), using this time the median velocity as dependent variable: when the robot was careful, participants significantly diminished their median velocity, with an estimated reduction in speed of  $0.055m/s$  ( $C - NC$ ,  $estimate = 0.055$ ,  $SE = 0.012$ ,  $t = -4.60$ ,  $p < 0.001$ ). These findings prove that the modulation of the robot movements affected how participants moved to reach the cup and take it from the robot gripper. This happened even if there was no reason to adapt to the object properties since we consider a reaching movement without any object directly involved; moreover, all the cups had exactly the same characteristics. We also verified, for both the duration and the median velocity of the reaching movements, if there was an interaction with the participants' accuracy in evaluating the robot's behavior in every trial. We used the accuracy in their classification as an additional factor in the mixed model, but we found no interaction with how they performed the reaching duration or velocity. The modulation in response to the robot attitude also occurred when participants did not recognize it explicitly.

## 4 Discussion

In our study, we exploited a generative approach to produce robot movements that could implicitly communicate if a handled object required or not carefulness to be transported. To avoid influencing the choice, all the items transported by the Kinova robot were identical (empty plastic cups). The participants had to decide if they were supposed to be virtually full or empty, without any particular hint or instruction on how to proceed. Firstly, we assessed (i) whether our controller can express caution in the gestures or its absence. According to the results shown in Figure 3, we notice that the *careful* robot actions have been perceived as such since the first trials of the experiment. Regarding the *not careful* actions, there is a learning curve in how they were perceived during the experiment. In the first trials, they were sometimes mistaken for actions associated with transporting a full cup. As the experiment progressed, the difference between the two modulations became more evident, with an accuracy in the participants' choices above 80%. Reflecting on the original dataset of human movements used to train the GANs, associated with the transport of full and empty cups [11],

<sup>5</sup> Jamovi software website: <https://www.jamovi.org>

<sup>6</sup> General analyses for linear models Jamovi module: <https://gamlj.github.io/>

we can observe that a *not careful* attitude is standard in our actions. Indeed, when no particular circumstances are forcing us, for instance when picking and placing an ordinary object, we tend to move in a “neutral” way, and we can shortly describe our gesture as not careful. On the contrary, a strong kinematics modulation appears when we are paying attention to not spill the content of a glass [3,10]. This careful kinematics shaping is what we truly modeled in the communicative robot’s movements, and it is rewarding that careful movements were perceived correctly from the beginning.

This study also allows us to evaluate (ii), the effect that the implicit modulation of the robot actions has on the interaction. Even though participants knew from the beginning that the plastic cups were all the same and all empty, there was a modulation in how they approached the robot gripper. We gave an overview of this phenomenon in Figure 4 and a quantitative assessment in Figure 5. If the robot manifested a careful attitude, adopting a lower magnitude in the velocity profile and a longer duration of the movement (see Figure 1 for reference), also the reaching movement of the humans was significantly slower. Interestingly, this also happened when participants had trouble explicitly recognizing the motion style and classifying the cup: the contagion in how they performed the reaching task was still present. This result emphasizes how important it is to modulate the actions of robots appropriately, with a view to collaborative interaction. Indeed, we proved a motor contagion from the robot to the human, even if there was for the participants no need directly associated with the task to adapt their motor strategies. We observed natural coordination emerging from such a simple task, where the pace of the human spontaneously adapted to the robot one, mimicking, even unconsciously, the attitude observed. Human-robot motor contagion on velocity was already observed, as far as the robot velocity profile is biologically plausible [2,20]. In our approach, the reasonableness of the velocity profiles was granted using a generative network trained on human examples. The findings in our study extend the existing evidence of motor contagion in Human-Robot Interaction, proving that robotics arms can also leverage it to convey appropriate ways of handling fragile objects.

## 5 Conclusion

In this study, we showed how a generative approach could be used to generate meaningful and communicative robot actions that a human partner can successfully interpret to infer some properties of the involved objects. This modulation on the robot side also led to a motor contagion in how the human performed its actions and synchronized with the pace of the robot; through motion alone, it was possible to open a channel of communication between the two agents, with measurable effects on the interaction.

Finally, it should be noted that we obtained these results by modulating the movements of a 7 degrees of freedom robotic manipulator, not a humanoid robot. Nevertheless, even though its kinematics was far from the one of a human arm, it was possible to achieve the desired communication intent by simply modulating

the end-effector control. This proves the power of the proposed approach and its potential scalability in other contexts and with other robots, also industrial ones, where implicit communication through motion could improve the efficiency and safety of a joint collaborative task.

In future works, we plan to exploit the same controller and have the robot actually manipulate full and empty cups to assess how the movement's modulation affects trust, perceived competency, and efficiency in a dyadic interaction, while facing a challenging task.

## References

1. Bingham, G.P.: Kinematic Form and Scaling: Further Investigations on the Visual Perception of Lifted Weight. *Journal of Experimental Psychology: Human Perception and Performance* **13**(2), 155-177 (1987)
2. Bisio, A., Sciutti, A., Nori, F., Metta, G., Fadiga, L., Sandini, G., Pozzo, T.: Motor contagion during human-human and human-robot interaction. *PLOS ONE* **9**(8), 1-10 (08 2014)
3. Duarte, N.F., Chatzilygeroudis, K., Santos-Victor, J., Billard, A.: From human action understanding to robot action execution: how the physical properties of handled objects modulate non-verbal cues. In: 2020 Joint IEEE 10th International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob). pp. 1-6. IEEE, Valparaiso, Chile (Oct 2020)
4. Duarte, N.F., Rakovic, M., Tasevski, J., Coco, M.I., Billard, A., Santos-Victor, J.: Action Anticipation: Reading the Intentions of Humans and Robots. *IEEE Robotics and Automation Letters* **3**(4), 4132-4139 (Oct 2018)
5. Fulton, M., Edge, C., Sattar, J.: Robot Communication Via Motion: A Study on Modalities for Robot-to-Human Communication in the Field. *ACM Transactions on Human-Robot Interaction* **11**(2), 1-40 (Jun 2022)
6. Garelo, L., Lastrico, L., Rea, F., Mastrogiovanni, F., Noceti, N., Sciutti, A.: Property-Aware Robot Object Manipulation: a Generative Approach. In: 2021 IEEE International Conference on Development and Learning (ICDL). pp. 1-7. IEEE, Beijing, China (Aug 2021)
7. Hamilton, A.F.C., Joyce, D.W., Flanagan, J.R., Frith, C.D., Wolpert, D.M.: Kinematic cues in perceptual weight judgement and their origins in box lifting. *Psychological Research* **71**(1), 13-21 (2007)
8. Hassanin, M., Khan, S., Tahtali, M.: Visual Affordance and Function Understanding: A Survey. *ACM Computing Surveys* **54**(3), 1-35 (Jun 2021)
9. Jamone, L., Ugur, E., Cangelosi, A., Fadiga, L., Bernardino, A., Piater, J., Santos-Victor, J.: Affordances in Psychology, Neuroscience, and Robotics: A Survey. *IEEE Transactions on Cognitive and Developmental Systems* **10**(1), 4-25 (Mar 2018)
10. Lastrico, L., Carfi, A., Rea, F., Sciutti, A., Mastrogiovanni, F.: From Movement Kinematics to Object Properties: Online Recognition of Human Carefulness. In: International Conference on Social Robotics. vol. 13086, pp. 61-72 (2021), arXiv: 2109.00460
11. Lastrico, L., Carfi, A., Vignolo, A., Sciutti, A., Mastrogiovanni, F., Rea, F.: Careful with That! Observation of Human Movements to Estimate Objects Properties. In: Saveriano, M., Renaudo, E., Rodríguez-Sánchez, A., Piater, J. (eds.) *Human-Friendly Robotics 2020*. pp. 127-141. Springer International Publishing, Cham (2021)

12. Lastrico, L., Garello, L., Rea, F., Noceti, N., Mastrogiovanni, F., Sciutti, A., Carfi, A.: Robots with Different Embodiments Can Express and Influence Carefulness in Object Manipulation (Aug 2022), arXiv:2208.02058 [cs]
13. Macciò, S., Carfi, A., Mastrogiovanni, F.: Mixed reality as communication medium for human-robot collaboration. In: 2022 International Conference on Robotics and Automation (ICRA). pp. 2796–2802 (2022)
14. Mottaghi, R., Schenck, C., Fox, D., Farhadi, A.: See the Glass Half Full: Reasoning About Liquid Containers, Their Volume and Content. In: 2017 IEEE International Conference on Computer Vision (ICCV). pp. 1889–1898. IEEE, Venice, Italy (Oct 2017)
15. Rosen, E., Whitney, D., Phillips, E., Chien, G., Tompkin, J., Konidaris, G., Tellex, S.: Communicating and controlling robot arm motion intent through mixed-reality head-mounted displays. *The International Journal of Robotics Research* **38**(12-13), 1513–1526 (Oct 2019)
16. Rosen, J., Perry, J.C., Manning, N., Burns, S., Hannaford, B.: The Human Arm Kinematics and Dynamics During Daily Activities – Toward a 7 DOF Upper Limb Powered Exoskeleton. 12th International Conference on Advanced Robotics p. 8 (2005)
17. Rosenbaum, D.A., Chapman, K.M., Weigelt, M., Weiss, D.J., van der Wel, R.: Cognition, action, and object manipulation. *Psychological Bulletin* **138**(5), 924–946 (2012)
18. Saponaro, G., Jamone, L., Bernardino, A., Salvi, G.: Interactive Robot Learning of Gestures, Language and Affordances. In: GLU 2017 International Workshop on Grounding Language Understanding. pp. 83–87. ISCA (Aug 2017)
19. Sciutti, A., Patanè, L., Sandini, G.: Development of visual perception of others' actions: Children's judgment of lifted weight. *PLoS ONE* **14**(11), 1–15 (2019)
20. Vasalya, A., Ganesh, G., Kheddar, A.: More than just co-workers: Presence of humanoid robot co-worker influences human performance. *PLOS ONE* **13**(11), 1–19 (11 2018)
21. Yamani, Y., Ariga, A., Yamada, Y.: Object Affordances Potentiate Responses but Do Not Guide Attentional Prioritization. *Frontiers in Integrative Neuroscience* **9** (Jan 2016)
22. Yoon, J., Jarrett, D., van der Schaar, M.: Time-series Generative Adversarial Networks. In: Wallach, H., Larochelle, H., Beygelzimer, A., Alché-Buc, F.d., Fox, E., Garnett, R. (eds.) *Advances in Neural Information Processing Systems*. vol. 32. Curran Associates, Inc. (2019)
23. Yu, L.F., Duncan, N., Yeung, S.K.: Fill and Transfer: A Simple Physics-Based Approach for Containability Reasoning. In: 2015 IEEE International Conference on Computer Vision (ICCV). pp. 711–719. IEEE, Santiago, Chile (Dec 2015)