2021 I-RIM Conference
October 8-10, Rome, Italy
ISBN: 9788894580525
DOI: 10.5281/zenodo.5900619

A pilot study towards the implementation of perceptual and motor adaptation in robots

Matilde Antonj[*, 1, 2], Joshua Zonca[*, 2], Linda Lastrico[*, 1, 2], Maura Casadio[*,1, 2], Alessandra Sciutti[*, 2]

***Abstract*—The aim of our study is to understand the perceptual and motor mechanisms of adaptation underlying human-robot interaction. Our long-term goal is to develop novel models of adaptation that could be implemented in robots to enhance human-robot collaboration. Realizing adaptive robots would be fundamental not only in the biomedical field for assistance and rehabilitation, but also in industrial settings to improve human-robot cooperation. In the current paper, we present a pilot experiment aimed at exploring perceptual and motor strategies adopted by participants who try to adapt their perception to that of a robot with different prior sensory experience.**

Keywords—Human-robot interaction, adaptation, shared perception, human kinematics

I. INTRODUCTION

Humans are continuously exposed to sensory stimuli coming from the surrounding environment. In order to cope with neural noise and sensory uncertainty, human perception does not rely only on the current sensory information, but takes into account previous experience [1]. For instance, when a set of stimuli (e.g., spatial lengths or temporal durations) is presented to a subject, the perception of the current stimulus is affected by the mean magnitude of previous stimuli. This perceptual mechanism can be referred to as central tendency. Central tendency can be modeled through Bayesian techniques [2, 3], by treating previous experience as priors [4]. Through the mechanism of central tendency, we can improve our perception by decreasing the magnitude of the total error and, therefore, increasing the reliability of what is perceived [5]. On the contrary, in interactive contexts, central tendency might not be the best strategy to coordinate with others. When we aim at coordinating with individuals that have different priors than ours, one hypothesis is that our perceptual processes would benefit from relying less on priors and more on the sensory information coming from the environment or from our partners [6]. In line with this hypothesis, a recent study has shown that, if a robot presents a series of stimuli during a task, participants perceive those stimuli relying less on their prior experience and this effect is stronger when the behavior of the robot is human-like [7].

To collaborate efficiently with others, humans have also to identify their partner's motor cues: thanks to this kinematic feedback, spatial and temporal adaptations between agents are facilitated [8, 9]. Therefore, understanding the interplay between perceptual mechanisms and motor strategies may be crucial to characterize adaptation during interaction, improving the development of adaptive robots, able to collaborate with humans in a human-like way [10].

In this pilot study, we designed an experimental task aimed at investigating how humans spontaneously adapt their perception and their kinematics to those of a robot in a joint temporal perception task. Participants had to reproduce the length of temporal stimuli by hitting a target with a wooden stick. Then they received a feedback on the estimate of a robot, which was programmed to have a different prior experience than participants' one and, therefore, a different perceptual model. In this case, participants had the goal to match the robot's estimate. We investigated how participants adapted their perception and their kinematics to align with the perceptual model of the robot.

In future experiments, the perceptual and motor strategies of adaptation identified in the current study will be implemented in robots in order to investigate the interplay between perception and kinematics in more complex and interactive human-robot contexts.

II. MATERIALS AND METHODS

A. Experimental Setup

The participant is positioned in front of a square-shaped table (78x78 cm). At the same time, two strips of leds are located horizontally on the table, respectively at a distance of 24 cm and 32 cm from the edge near the participant. A piezoelectric sensor is enveloped in a rubber pad, to register the event of the participant's hit. This represents the target for the participant's movement. To acquire the kinematic data, markers are attached to the shoulder, the elbow and three points on the metacarpal bones of the hand. The Optotrak system is used to register the positions of the markers, with a frequency of 100 Hz. (For a complete view of the experimental setup, see Fig. 1.)

B. Experimental Paradigm

Two consecutive flashes are presented to participants through a strip of white leds. The two flashes are separated by a temporal duration, chosen randomly in the range of 1.0-4.0 s with a step of 0.375 s. This temporal duration represents the target stimulus duration. The participant's task varies across three different conditions:

a) Individual (IND): The participant's goal is to estimate the temporal duration between the presented flashes in the most accurate way.

b) Social with Feedback (SWF): Differently from the Individual condition, participants perform the task with a robot, which is supposed to produce a perceptual estimate of the current temporal stimulus. The participant's goal is to align their response with the robot's one. The temporal estimate of the robot is showed through the lighting of a strip of green leds, whereas the robot is not physically present in the room during the experiment.

c) Social No Feedback (SNF): As in SWF, the participant's goal is to align their response with the robot's one. Differently from SWF, the robot's response is not showed and participants have to use the strategy they have learnt during SWF.

The participants' estimated temporal duration (which should be accurate in the Individual condition and match the one of the robot in the social conditions) is defined as the temporal distance from the second flash and the hit on the target. To replicate the desired temporal interval, participants have to move a wooden stick from the resting position and hit a target on the same horizontal plane.

We gratefully acknowledge the financial support of the European Financial Council (ERC Starting Grant 804388, wHiSPER).

[1] DIBRIS, Università degli Studi di Genova, Genoa, Italy
[2] Istituto Italiano di Tecnologia, Genoa, Italy
[*] {matilde.antonj, joshua.zonca, linda.lastrico, alessandra.sciutti}@iit.it, maura.casadio@unige.it

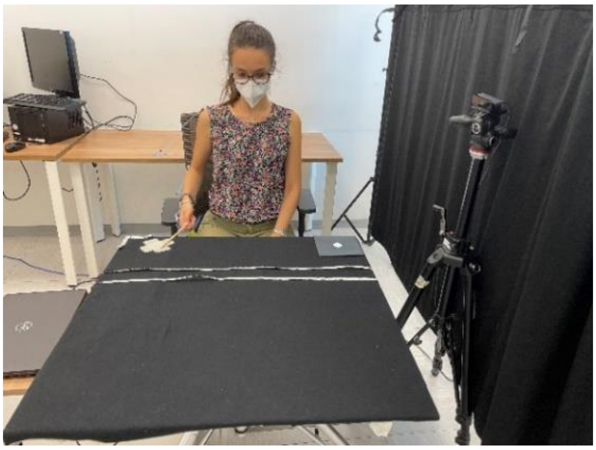

Fig. 1. Experimental setup. The participant is seated in front of a square-shaped table with two strips of leds, located horizontally. The participant, with motion capture markers on the right arm, assumes the resting position. The target point is positioned on the opposite corner of the table, aligned to the resting position.

Ten volunteers took part in the experiment and each of them addressed the three sessions in the described order. IND and SWF conditions are characterized by 54 trials where 9 different stimuli are presented six times in a random order. In SNF, each stimulus is repeated three times, for a total of 27 trials.

C. *Perceptual Behavior of the robot*

To realize the robot behavior, we ran a modified version of the Individual condition of the experimental paradigm with other 10 participants. In this version, we presented temporal durations in the range 4.0-7.0 s with a step of 0.375 s, resulting in 9 different stimuli, the mean of the participants' temporal estimates was computed for each stimulus and linear regression was used in order to simulate the participants' mean responses in the range of shorter stimuli (1.0-4.0 s): therefore, these simulated values were used as the means of nine different normal distributions with the corresponding standard deviation, derived from the analyzed data. Following the previous approach, the standard deviation was realized from behavioral data: we considered the mean of the standard deviation of the reproduced temporal durations in the range of 4.0-7.0 s. Subsequently, applying linear regression, mean standard deviations were computed in the range of shorter stimuli (1.0-4.0 s).

In this way, the behavior of the robot was realized so that every time the robot has to give its response, a value was picked from the normal distribution, associated to the corresponding stimulus. Therefore, the model of the robot was designed in order to produce perceptual responses in the range of shorter stimuli (1.0-4.0 s), taking into account a prior experience in longer temporal durations (4.0-7.0 s). Thanks to this approach, it was possible to investigate human adaptation to an agent, characterized by a different sensory experience (i.e., prior, in Bayesian terms).

III. Results

A. *Perceptual Results*

First, we highlight that participants' responses were shorter than that of the robot (mean participants' estimated duration in IND: 2244 ms; mean robot's estimated duration in SWF: 3211 ms). This is the result of the manipulation of the robot's perceptual behavior, which was programmed to converge to a different (i.e., longer) prior (Fig. 2 and 3). To investigate between-condition differences in participants' perceptual responses, we ran a mixed-effects model with trial-by-trial reproduced duration (in ms) as dependent variable, condition as categorical factor and random effect at the subject level. The random effect was applied to the intercept to adjust for the individual differences in baseline reproduced duration levels and model intra-subject correlation of repeated measurements. Results reveal that participants' reproduced durations were longer in the social conditions (SWF and SNF) than in the Individual one (SWF – IND: B = 444.47, z = 7.93, p < 0.001; SNF – IND: B = 993.23, z = 14.47 p < 0.001). Moreover, reproduced temporal durations were longer in SNF compared to SWF (SNF – SWF: B = 548.77, z = 7.99, p < 0.001). In addition, we compared participants' and robot's estimation means in IND and SNF conditions, to understand whether participants shifted from their original response distribution in IND to that of the robot in the SNF condition. Results show that in the IND condition participants and the robot have a significantly different mean of perceptual responses, while in the SNF participants' mean of reproduced temporal durations is aligned with the robot's one (Wilcoxon-rank sum test, IND, ROBOT-HUMAN: z = 3.02, p = 0.002; SNF, ROBOT– HUMAN: z = -0.53, p =0.597). This result demonstrates that participants' responses in the SNF condition are converging to a different prior, which coincides with that of the robot (Fig. 2 and 3).

B. *Kinematic Results*

In each trial, the onsets of the movement were computed as the 20% of the peak of the velocity profile (Fig. 4). Results show that average onsets were significantly different across the three conditions: we ran a mixed-effects model with trial-by-trial onset as dependent variable, condition as categorical factor and random effect at the subject level. Results reveal that participants' movements started later in the social conditions (SWF and SNF) than in the Individual condition (SWF – IND: B = 0.50, z = 4.26, p < 0.001; SNF – IND: B = 1.04, z = 7.30, p < 0.001). Moreover, movements started later in SNF compared to SWF (SNF – SWF: B = 0.54, z = 3.78, p < 0.001). Results on participants' kinematics mirror our behavioral results, revealing that the perceptual adaptation observed in SWF and SNF (i.e., increase in the reproduced stimulus duration) is implemented through a delay in the onset of the action (Fig. 5).

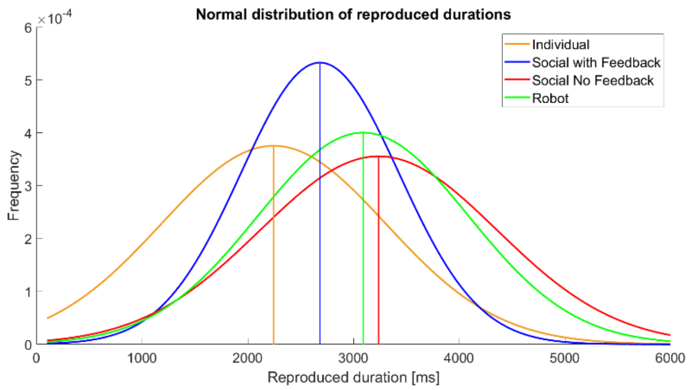

Fig. 2. Normal distribution of reproduced durations. We plot the frequency of each reproduced duration: reproduced durations are represented on the x-axis, while the y-axis refers to their frequencies. Comparing these three distributions, it is notable that there is a shift in the prior (distribution mean) among conditions.

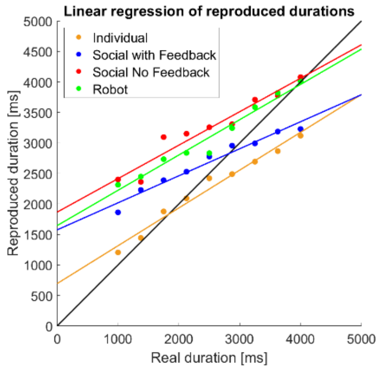

Fig. 3. Linear regression of reproduced durations. Real durations are represented on the x-axis, while the y-axis refers to the reproduced durations of each values. Each colored line represents the linear regression of the means of subjects' reproduced durations in the different conditions. The green line represents the set of responses of the robot. The black one corresponds to the identity line: it represents the line of the correct responses. We can see a shift in the mean of reproduced durations among conditions (vertical shift between IND and SNF). Moreover, we can observe that participants' estimate gravitate towards these means due to the phenomenon of central tendency (the slope of the regression lines in IND and SNF is lower than that of the identity line).

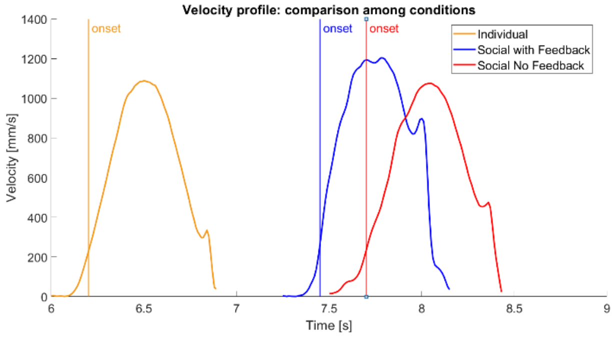

Fig. 4. Velocity profile: comparison among conditions. This figure shows only one subject's velocity profile in three trials with the same stimulus duration, assigned to different conditions. Time is represented on the x-axis, while the y-axis refers to the velocity of the hand until the participant's hit on the rubber pad. The three vertical lines identify the onset of the movement to reproduce the temporal duration equal to 2.5 s in the three different conditions. It is clear that the onset of the movement is delayed in the social conditions (SWF and SNF).

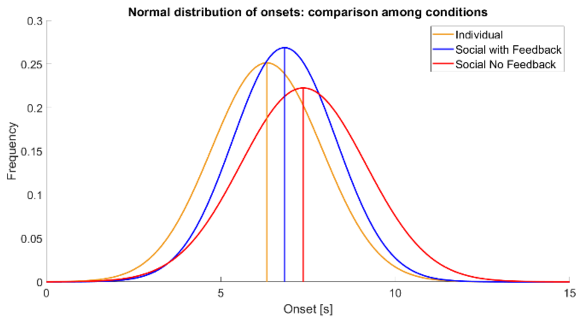

Fig. 5. Normal distribution of onsets: comparison among conditions. The values of frequency of each onset are plotted in this figure: onsets are represented on the x-axis, while the y-axis refers to the frequency of each value. Comparing these three distributions, it is notable that there is a shift in the start of the movement among conditions, mirroring the behavioral analysis (Fig. 2).

IV. CONCLUSION

Results of this pilot study reveal that participants can learn and align with the perceptual model of a robot during a joint perceptual task. Moreover, through the analysis of kinematic data, we identified a relationship between perceptual and motor mechanisms of adaptation. In particular, participants use the robot's feedback to shift the prior of their response distribution to align it with the one of the robot. From a motor perspective, this adaptation is implemented through a shift in the onset of movement, suggesting the emergence of top-down processes of action modulation, guided by perceptual adaptation.

Future studies will implement a model of the perceptual mechanisms and the motor strategies observed in the current pilot study in a social robot. Indeed, participants will physically see the robot and will have an embodied interaction with it. In this way, the movements of the robot will be visible: therefore, participants' kinematics will be studied in order to investigate whether and how motor feedbacks can help the participant to align with the robot.

In conclusion, the current study offers novel insights on how humans adapt their perceptual and motor strategies to align with another agent during interaction. In the future, the implementation of human-inspired perceptual and kinematic models will enable robot to be adaptive so that they can assist elderly people and be used for rehabilitation. In the former scenario, robots will manage to predict humans' perceptual, motor and behavioral peculiarities in order to tailor their actions or warn them in case of inaccuracy. In the context of rehabilitation, patients have to repeat the same actions multiple times, inducing to regress to the mean: thanks to this research, it may be possible to understand and predict patients' actions and guide them towards an improvement [11].

REFERENCES

[1] Knill, D. C. & Pouget, A. "The Bayesian brain: the role of uncertainty in neural coding and computation." TRENDS in Neurosciences 27.12 (2004): 712-719.

[2] Kersten, D. & Yuille, A. "Bayesian models of object perception." Curr. Opin. Neurob. 13.2 (2003): 150-158.

[3] A. Sciutti, D. Burr, A. Saracco, G. Sandini, and M. Gori, "Development of context dependency in human space perception," Exp. Brain Res., vol. 232, no. 12, pp. 3965–3976, 2014.

[4] Jazayeri, M. & Shadlen, M. N. "Temporal context calibrates interval timing." Nat. Neurosci. 13.8 (2010): 1020.

[5] Cicchini, G. M., Arrighi, R., Cecchetti, L., Giusti, M., & Burr, D. C. (2012). "Optimal encoding of interval timing in expert percussionists." Journal of Neuroscience, 32(3), 1056-1060.

[6] J. Zonca, A. Folsø and A. Sciutti, "If you trust me, I will trust you: the role of reciprocity in human-robot trust," submitted for publication, arXiv preprint: http://arxiv.org/abs/2106.14832, unpublished.

[7] Mazzola, C., Aroyo, A. M., Rea, F., & Sciutti, A. (2020, March). Interacting with a social robot affects visual perception of space. In Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (pp. 549-557).

[8] Torres-Oviedo, G., & Bastian, A. J. (2010). "Seeing is believing: effects of visual contextual cues on learning and transfer of locomotor adaptation." Journal of Neuroscience, 30(50), 17015-17022.

[9] Sing, Gary C., and Maurice A. Smith. "Reduction in learning rates associated with anterograde interference results from interactions between different timescales in motor adaptation." PLoS computational biology 6.8 (2010): e1000893.

[10] Hogan, N., & Sternad, D. (2012). Dynamic primitives of motor behavior. Biological cybernetics, 106(11), 727-739.

[11] Krebs, H. I., & Hogan, N. (2006). Therapeutic robotics: A technology push. Proceedings of the IEEE, 94(9), 1727-1738.