

Research Article

Open Access

Fabio Vannucci*, Alessandra Sciutti, Hagen Lehman, Giulio Sandini, Yukie Nagai, and Francesco Rea

Cultural differences in speed adaptation in human-robot interaction tasks

<https://doi.org/10.1515/pjbr-2019-0022>

Received December 13, 2018; accepted June 25, 2019

Abstract: In social interactions, human movement is a rich source of information for all those who take part in the collaboration. In fact, a variety of intuitive messages are communicated through motion and continuously inform the partners about the future unfolding of the actions. A similar exchange of implicit information could support movement coordination in the context of Human-Robot Interaction. In this work, we investigate how implicit signaling in an interaction with a humanoid robot can lead to emergent coordination in the form of automatic speed adaptation. In particular, we assess whether different cultures – specifically Japanese and Italian – have a different impact on motor resonance and synchronization in HRI. Japanese people show a higher general acceptance toward robots when compared with Western cultures. Since acceptance, or better affiliation, is tightly connected to imitation and mimicry, we hypothesize a higher degree of speed imitation for Japanese participants when compared to Italians. In the experimental studies undertaken both in Japan and Italy, we observe that cultural differences do not impact on the natural predisposition of subjects to adapt to the robot.

Keywords: human-robot interaction, cultural differences, mutual adaptation, joint task

1 Introduction

The commercialization of social robots is spreading fast. Some good examples of this trend are NAO [1] and Pepper [2] that can welcome and play with visitors in museums [3] or give directions in shopping malls [4]. Since social robots are designed to interact with people, they must be skilled in the different aspects of communication. One crucial element is the selection of behaviors that maximize intuitive understanding and acceptance in the human partners [5]. In this respect, one feature to which humans are particularly sensitive to is motion: a lot of information is transmitted with our “way of moving” [6]. In particular, motion is a key feature when we have to adapt to each other. For example we unconsciously adapt to the speed of our partner when walking together. Robots should be able to exploit the same mechanism [7] and recent evidences show that human-inspired movements performed by humanoid robots are effective in facilitating intuitive temporal coordination in HRI contexts [8].

However, it is important to take into account that the specific needs of the final users might vary as a function of the environment in which the robot will operate, potentially due to the cultural differences between the countries where it will be deployed. One example is body language, where very clear differences exist between Eastern and Western cultures. For instance, Italians tend to use many different gestures, especially while talking, whereas Japanese have a very specific set of rules for body movements in social interactions [9], such as bows. Even considering more implicit signals which are not subjected to direct voluntary control as eye movements [10] substantial differences exist [11], such as how often direct eye contact is established during a social exchange [12]. The same robot movement might be very effective in interaction within a certain cultural context and highly disruptive in a different one. In this work, we want to assess the relative impact of Eastern and Western cultures, namely Japanese and Italian, on movement-based implicit communication, and in particular emergent temporal coordination through automatic imitation [13].

***Corresponding Author: Fabio Vannucci:** DIBRIS, Università di Genova, Italy; Robotics, Brain and Cognitive Sciences, Istituto Italiano di Tecnologia, Italy; E-mail: fabio.vannucci@iit.it

Alessandra Sciutti: COgNiTive Architecture for Collaborative Technologies, Istituto Italiano di Tecnologia, Italy

Hagen Lehman: Università degli Studi di Macerata, Italy

Giulio Sandini: Robotics, Brain and Cognitive Sciences, Istituto Italiano di Tecnologia, Italy

Yukie Nagai: International Research Center for Neurointelligence, The University of Tokyo, Japan

Francesco Rea: Robotics, Brain and Cognitive Sciences, Istituto Italiano di Tecnologia, Italy

1.1 Cultural differences and robots

Cultural differences have been studied in many aspects in the field of social robotics [14–16]. In 2004 Kaplan [17] analyzed the problem of understanding the possible cultural issues involved in the acceptance of humanoid robots. Differences in the perception of robots between the eastern and western parts of the world, in particular between Japan and the West, are affected by literature, philosophy, and history among many other factors. Kaplan, in his preliminary conclusion about this topic, affirms that in the West there is a continuous debate on what does or does not distinguish humans from machines and that the “*possible convergence of humans and machines is both fascinating and frightening*”. On the other hand, in Japan, even though technology is strongly present in daily life, “*a distance is always maintained between the human body and technological prosthesis*” [17]. This is probably the reason why robots, and more in general technological artifacts, usually do not cause suspicion or uneasiness. These cultural differences in robot perception have clear implications in various fields of application of robotic platforms like teaching, entertainment, security [15].

1.2 Motion in human-human and human-robot cooperation

During a cooperative joint task between two humans, both the actors tend to adapt their speed to each other in order to reach the best balance and to efficiently achieve the shared goal. These behavioral phenomena have been deeply investigated and are known as “emergent coordination” [18]. In particular, humans are exceptionally skilled when it comes to perform a joint action without the need for verbal interaction. The reason is that perception and action are very tightly linked. The so-called “mirror-neuron system” is involved in perception, understanding and anticipation of others and, at the same time, in the execution of actions [19, 20]. The neural areas elicited during the execution of a goal-oriented motion in the human, are also activated when observing someone else performing an action with the same goal. This phenomenon is known as “motor resonance” and appears already in the early stages of human infancy [21, 22]. It has important implications in facilitating interaction, for instance naturally leading two partners to move at the same pace [23] or even promoting synchronized improvisation between expert musicians and actors [24]. Motor resonance in the form of automatic imitation (or the “*tendency to reproduce observed actions involuntarily*”) is not limited to human-

human interaction, but is also triggered when interacting with an agent that moves in a biological fashion, matching our subconscious expectations [25]. This leads for instance the human partner to automatically adapt to the robot speed [13] or to change the movement style to match that of the robot [26]. This implies that interactions with robots – in particular humanoids – can leverage on similar mechanisms of motor resonance as in human-human collaborations [27].

1.3 Motor resonance and culture

In this paper, we assess whether low-level mechanisms, which support interaction – as motor resonance in the form of automatic imitation – are influenced by cultural differences. Indeed, converging evidence shows that top down modulation of the activation of the neural substrates possibly supports emergent coordination. Amoruso et al. [28, 29] have for instance demonstrated that motor resonance is not an entirely automatic process, but it can be modulated by high-level contextual representations. Also social aspects might affect motor resonance. Recent neuroimaging studies show that mirror system activation is modulated by social group membership, with higher activation during action observation when the action is performed by an in-group rather than an out-group member [30, 31]. However, it is not yet known whether cultural differences might have an influence on emergent coordination, leading to different patterns of behavior across countries, also when the joint action is performed with a partner of the same culture.

1.4 Objective of the study

In this study, we aim at quantifying adaptation in the form of automatic imitation in a human-humanoid interaction in Italy and Japan. One participant and the humanoid iCub robot sat at a table face to face. They had the same task: receiving a Lego block from an experimenter and dropping it into the box prepared on the table. To assess the degree of adaptation we adopted a well-established paradigm for automatic imitation measures [8]. We set different speeds for the movement of the robot and assessed whether participants automatically imitated the robot velocity while performing the task. We also compared the rate of adaptation between Italian and Japanese participants. Moreover, we manipulated the goal of the task in order to investigate the effect of space sharing during the task in the different cultures. In one condition the robot and the participant

had to put the Lego block into one single big shared box, in another condition the participant and the robot deployed Lego blocks into two distinct boxes.

If cultural differences impacted emergent coordination, we would expect to find differences in the degree of adaptation between the Italian and Japanese population. Japanese people showed a higher general acceptance toward robots when compared with Western cultures [17]. Since acceptance, or better affiliation, is tightly connected to imitation and mimicry [32], we hypothesized a higher degree of speed imitation for the Japanese sample, when compared to the Italian one. Moreover, given the difference in the perception of personal space between the two cultures [33], we also hypothesized a stronger impact of sharing the target space in the “shared box” condition in the Japanese group, when compared to the Italian one.

2 Methods

2.1 Experimental design

We designed an experiment in which our participants had to perform a joint task with the humanoid robot iCub [34] (see Figure 1). Human and robot were sitting face to face and their goal was to fill a box with Lego blocks. Each trial started when an experimenter (different in Italy and Japan) put simultaneously one block in the open hand of the robot and another in the open hand of the participant. The following blocks were passed only when both had dropped the previous one into the box, and after they both put their hand in the initial position. Thanks to this, the robot and human always started at the same moment. The experimenter explained, at the beginning of the experiment, that the robot and the participant both had the same goal of filling the big box or the one that was closer to them with the Lego blocks that he would have provided them. He indicated that they could only get a new block after both had dropped the previous one. No instructions were provided regarding the synchronization with the robot. The robot was preprogrammed to transport the block on its open palm, drop it into the box, and then immediately go back to the initial position. Each participant performed 6 sessions of 10 repetitions, i.e. they transported 10 blocks into the box for each session. During different sessions two factors were manipulated: Robot Speed, which could be *slow*, *medium* or *fast* and Box Number, which could be *one* – corresponding to a shared target space, or *two* – corresponding to individual space. The order of conditions was randomized and unique for every subject.

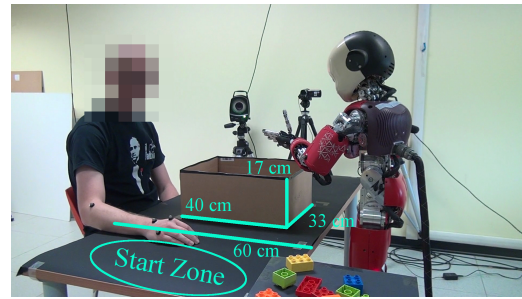


Figure 1: Setup of the experiment. Blocks are passed by an experimenter in the start zone to both the participant and the robot at the same time. Their goal is to drop these blocks in the box. The picture is a snapshot of the “one box” condition.

We set as “fast” a speed that we considered reasonable to complete the task after a few pilot tests, then we selected the “mid” and “slow” speeds by trying to maximize the differences among conditions, without making the “slow” motion seem unnatural to the participant. The average speeds in all conditions are reported in Table 1. It is slightly different between the two countries. The reason for this discrepancy is that two different iCub robots were used and the movement speed is influenced by different factors, out of which some were not under our control, including the age of the robot and low-level settings of the electronics. However the values are comparable between the two experiments and the relative variation in robot’s speed between the fast and slow conditions was very similar (54% velocity increase vs. 57% velocity increase). The motions of the robot were inspired by biological human-like movements, as detailed in the next section.

2.2 Biological movement implementation

The Two-Thirds Power Law [35, 36] is a well-known feature of human motion, that relates the speed and curvature of an elliptical movement. We designed a module that enables the robot to execute curve movements compliant with this law [37], leveraging on the existing Cartesian controller of the iCub [38]. Given a specific trajectory in 3D space, the module can convert it to a smooth human-like movement, which is then executed by the robot through the original Cartesian controller. Particular care was devoted to the generation of biologically plausible motion for the robot since, as it has been previously demonstrated, such motion is crucial in eliciting motor resonance and automatic imitation in human-robot interaction [8, 39].

Table 1: Robot hand speed in the different experimental conditions: average and standard deviation.

	Slow speed	Mid speed	Fast speed
Italy	0.100 ± 0.008 m/s	0.125 ± 0.007 m/s	0.154 ± 0.009 m/s
Japan	0.080 ± 0.008 m/s	0.108 ± 0.004 m/s	0.126 ± 0.006 m/s

2.3 Subjects

The experiment was performed in Italy (17 subjects) and in Japan (9 subjects). We excluded 3 participants (2 from Italy, 1 from Japan) for technical problems with data acquisition. Participants in both countries had different working backgrounds, from university students, to lab technicians or administrative staff. The local institutional ethics committees approved the protocol and all subjects gave informed consent before participating.

- Italy: Mean age 30 years \pm 5 SD, 6 males, 9 females, 1 Left handed, 14 right handed.
- Japan: Mean age 29 years \pm 9 SD, 3 males, 5 females, all right-handed.

2.4 Data

For each subject, we acquired video recordings and kinematic data using a motion capture system. Videos were recorded from two different points of view in order to monitor subjects' behavior in detail. Motion capture data, including 3D trajectory and speed of the hand and arm for each time frame (100 Hz), were gathered using four markers: three on the hand as shown in Figure 2, and one on the elbow. The motion capture systems are a VICON System of infrared cameras, capturing at a 100 Hz rate in Italy and a Motion Analysis MAC3D capturing at a 200 Hz rate in Japan. Data about the motion of the robot were gathered in two different ways for the two countries. In Italy, it was collected with a motion capture system using 4 markers (one on the elbow) as shown in Figure 2. In Japan data was collected through the use of “yarpdatadumper”, a module created to record and save in files different kind of information from the robot [40]. We took advantage of this to record the values of the joints of the robot arm at specific time instants. We then transformed the joints data into 3-dimensional trajectories and speeds with the specific tools supplied by the designers of the robot [41].

Our analysis focused mainly on the speed of the two agents. From the kinematic data, we extracted the starting and ending moment of each action. We considered the start as the last minimum of the speed when the hand is in the initial zone, while the end is the first minimum of the speed when the hand is in the box zone. Using these two

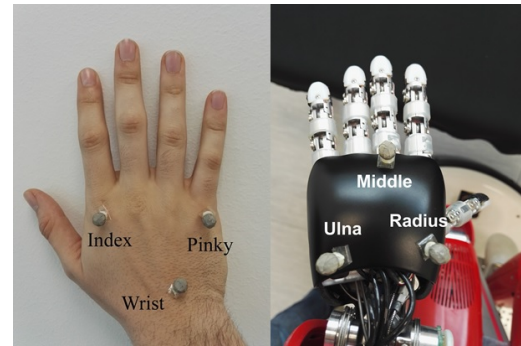


Figure 2: Motion capture markers. Positioning is different between subject and robot in order to distinguish them better during the data analysis phase.

time landmarks, we analyzed the most relevant part of the action and calculated the average speed for each repetition.

3 Results

The aim of this paper is the analysis of adaptation to a humanoid robot in a collaborative joint task when both agents were tasked to put Lego blocks into a box. In Figure 3 there is an example of the trajectories of both the robot (right) and a representative subject (left) in the “Robot Fast” and “Robot Slow” conditions. The path followed by the robot is very precise and only few variations can be noticed, while the trajectory of the subject is much more heterogeneous (see Figure 3).

From visual inspection of the participants transport trajectories in both the Italian and the Japanese sample, it emerged that during box reaching they were quite similar across different conditions. During the return to the starting zone, a lot of variability emerged within and between subjects since they did not receive any instruction regarding this phase, to keep the experiment as natural as possible. Participants chose either to stop near the box or at a different position before getting back to the start zone for the next trial (see blue lines in Figure 3 for an example).

The “reach-the-box” phase is the only part of the movement which presented a constraint: the robot and the participant started their action at the same moment. We

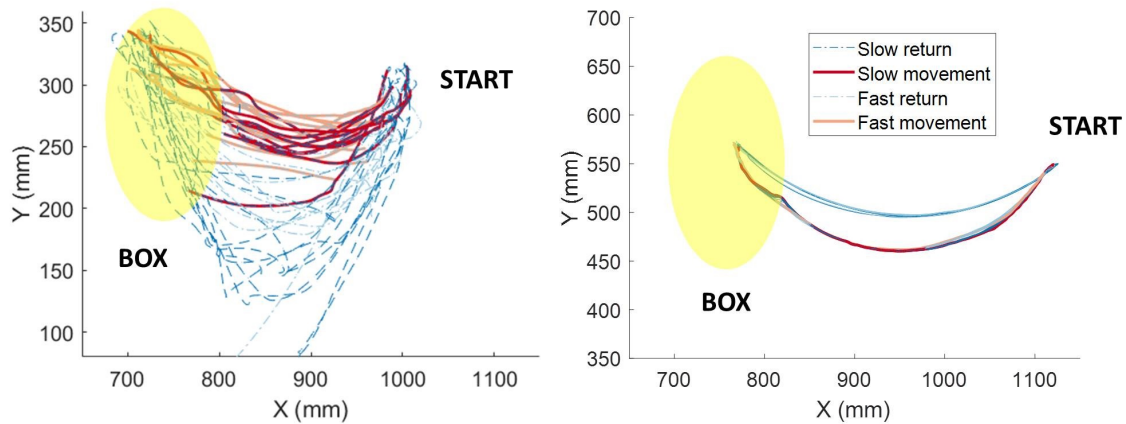


Figure 3: Trajectories of one representative subject (left) and the robot (right) during the “Robot Fast” and “Robot Slow” conditions. The plot refers to an Italian subject but is also representative for the behavior of Japanese participants. X and Y axis are a projection of the 3D space in the frame of reference of the motion capture system.

consider this as the most interesting part on which to focus our subsequent speed analysis. In the following sections we first introduce the results extracted from the analysis of the Italian experiment, then we compare them with the data acquired in Japan.

3.1 Human-robot interaction in Italy

Figure 4 shows the mean speed of all participants in Italy, for each of the different conditions. It is clear that, on average, subjects’ motion was significantly faster than the motion of the robot (two sample t-tests between subjects’ and robot velocity in the corresponding “robot speed” conditions, all p ’s < 0.05). From this chart also a form of adaptation can be noticed: even if subjects were faster than the robot, their movement speed varied according to the three different speeds of the robot. Conversely, performing the task with one shared box or two different ones does not seem to trigger a different behavior in the subjects. A two-way repeated measures ANOVA with Greenhouse-Geisser correction, on subjects’ speed with factor “Robot speed” (three levels: Slow, Medium, Fast) and factor “Number of Boxes” (two levels: 1 shared box, 2 separate boxes) followed by Tukey post hoc tests shows a significant change in participants’ movement speed as a function of robot velocity ($F(1.95, 27.25)=17.43$, $p < 0.01$), in particular between the “Robot Fast” and “Robot Slow” conditions (post hoc Tukey test: $t(28)=5.79$, $p < 0.01$), and between the “Robot Mid” and “Robot Slow” conditions (post hoc Tukey test: $t(28)=3.6$, $p < 0.05$). No significant effect is evidenced as a function of the number of boxes $F(1,14)=0.23$, $p=0.6$).

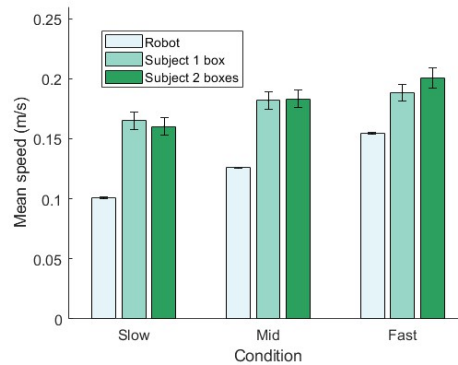


Figure 4: Mean speed, and standard error (error bars) of the robot and subjects for all conditions in Italy.

Further evidence of an automatic adaptation to the robot’s velocity comes from the analysis of individual subjects’ behaviors. The graphs in Figure 5 show that almost all participants lie above the dashed identity line, meaning that their speed in the “Robot Fast” and “Robot Mid” conditions was higher than their speed in the “Robot Slow” condition.

After the analysis of the mean from a general point of view, we looked for a possible effect of adaptation during the multiple repetitions. The panels of Figure 6 represent the mean across all Italian subjects, for each of the ten actions. We could not distinguish any form of adaptation with the progression of the repetitions, not even any particular trend. To statistically verify this observation, we ran two two-way repeated measures ANOVA with Greenhouse-Geisser correction, with factors “Robot Speed” (3 levels) and “Repetition number” (10 levels) on the “One box” and “Two boxes” conditions respectively. The analysis showed

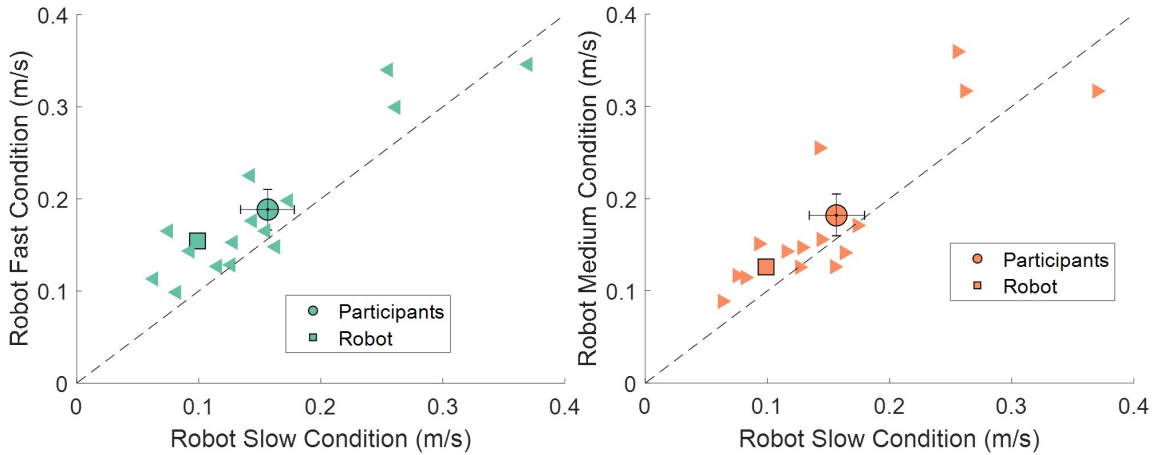


Figure 5: Individual speed of subjects in the “Robot Fast” (left) and ”Robot Mid” conditions (right), in relation to their speed in the “Robot Slow” condition in the “One Box” sessions. The bigger circles with error bars represent the sample mean and standard error. If a subject maintained the exact same speed, for example, in the “Robot Fast” and ”Robot Slow” condition, the corresponding marker would be on the dashed line. Similar results derive from the analysis of the “Two Boxes” condition (Italy).

that in neither case there was a significant effect of the interaction ($F(5.32)=1.17, p=0.33$ and $F(6.46)=1.04, p=0.41$). Moreover, there was no significant difference among the repetitions for the “One box” case ($F(3.71)=2.16, p=0.91$), whereas in the “Two boxes” case a slight reduction in speed was observed between the beginning and the end of the task ($F(4.16)=3.9, p<0.01$). In particular a post hoc Tukey test highlighted a significant difference only between the first and four last trials (all $p<0.05$).

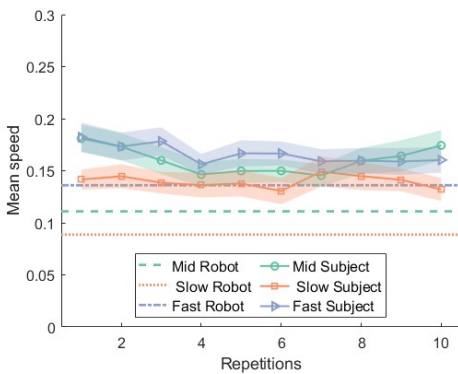


Figure 6: Mean subjects’ speed for each of the ten repetitions (“One Box”), with shaded areas representing standard error of the mean. The dashed lines represent the speeds of the robot, equal across all the ten repetitions (Italy).

3.2 Japan

As mentioned in the previous sections, one of the focuses of this work is to compare the behavior of Italian and Japanese subjects and to investigate whether the automatic imitation of the robot is affected by cultural differences between the two countries. After completing the analysis of the data gathered in Italy, we performed the same analysis with the Japanese data.

From the bar chart in Figure 7, we can see that subjects in Japan are always faster than the robot and that they tend to adapt to the different robot speeds. Concerning the ‘shared space’ condition there seem to be no differences across the “number of boxes” variable, for all the three speeds of the robot. These observations are confirmed by a two-way repeated measures ANOVA with Greenhouse-Geisser correction, on subjects speed with factors “Robot speed” and “Number of boxes”, which shows a significant change in participants’ speed as a function of robot movement velocity ($F(1.68, 11.74)=7.74, p<0.01$) and no change as a function of number of boxes ($F(1,7)=2.73, p=0.14$). Similarly to Italian subjects, there is a significant difference for Japanese subjects between the “Robot Fast” and “Robot Slow” conditions (post hoc Tukey test: $t(14)=4.99, p<0.01$), and between the “Robot Mid” and “Robot Slow” conditions (post hoc Tukey test: $t(14)=4.55, p<0.05$).

Figure 8 illustrates another validation of the aforementioned results. Here each marker represents a single subject and since the majority of them are above the dashed identity line, this means that they changed their speed according to the change of the robot velocity, adapting to it.

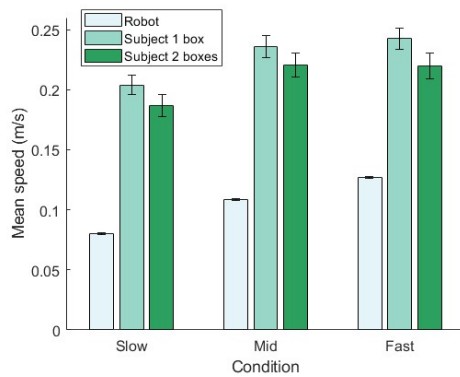


Figure 7: Mean speed of the robot and subjects in all conditions in Japan. Error bars represent standard errors.

Finally, as shown in Figure 9 there is no clear effect of adaptation with the progress of repetitions. Two two-way repeated measures ANOVA with Greenhouse-Geisser correction, with “Robot Speed” and “Repetition number” as factors, did not highlight any significant difference among repetitions for both the “One-box” ($F(2.62)=2.7$, $p=0.08$) and the “Two-boxes” ($F(3.83)=1.39$, $p=0.26$) conditions, nor any interaction ($F(5.05)=1.05$, $p=0.4$ and $F(4.42)=1.09$, $p=0.38$).

3.3 Cross-cultural comparison

From the previous sections, it appears that there is a high similarity in the behaviors of the participants from Italy and Japan during the interaction with the robot: both groups were influenced by robot’s speed, with no modulation of their behavior due to the presence of a shared box or two individual target spaces.

To directly compare the level of adaptation between the two countries, we plotted the individual speeds of all participants in the “Fast Robot” and “Slow Robot” conditions on the same graph (Figure 10). In the figure, it can be noticed that, on average, the Italian subjects tend to be slower than the Japanese, even though the robot speed was slightly faster in the Italian experiment than in the Japanese one (see squares in Figure 10 and Table 1). Besides this effect, both the Italians and the Japanese tend to change their speed similarly, adapting to the changes of the robot velocity, as shown by the similar relative position of the average markers with respect to the identity line

To quantify the degree of adaptation, and directly compare it between the two countries, we computed for each subject the relative variation of her speed as a function of the variation in the robot speed across conditions. To do so we regressed each participants’ average speeds in

the three robot speed conditions with respect to the corresponding robot’s velocities and extracted the slope of the resulting line. A number close to one would correspond to a relative change in subjects’ speed comparable with that exhibited by the robot across conditions, implying a high level of adaptation. The stem plots in Figure 11 represent the computed individual slopes, which are similar between the two groups. A mixed two-way repeated measures ANOVA with Greenhouse-Geisser correction, on subjects slope with “Number of Boxes” as within factor and “Nationality” as between factor: shows that the adaptation was not significantly different between the two groups tested ($F(1,21)=0.267$, $p=0.61$), nor was affected by the presence of a shared target space ($F(1,21)=0.005$, $p=0.94$).

3.4 Impact of common start

A possible confounding variable in the experiment was that by design we established the restriction to start each repetition together with the robot. Participants might, therefore, have been induced to slow down when the robot slowed down, in order to synchronize their arrival to the start zone with that of iCub. In other words, they might have planned their actions with the covert aim of coordinating with the starting action of the robot. If this assumption is correct, we might expect that people who showed a stronger adaptation to the robot speed (slope closer to one) were also arriving at the starting zone in synchrony with the robot. We therefore computed the “mean return difference” as the difference in timing between the instants in which the robot and the subject returned their hand to the starting position to get a new block. A correlation between high slopes (close to one) and small mean return difference would mean that they slowed down intentionally. The computation of slope and mean return difference showed no connection between the two measures, as can be seen in Figure 11. The distribution of the stems, representing the adaptation (slope), is not clearly related to the bars showing the difference in return timing. It seems therefore unlikely that the common start of repetition timing during the trials influenced the adaptation to the robot.

4 Discussion

Since mutual adaptation is a well-known effect of human-human collaboration, we wanted to investigate whether we could find a similar reaction when the counterpart is a humanoid robot and whether it can be influenced by

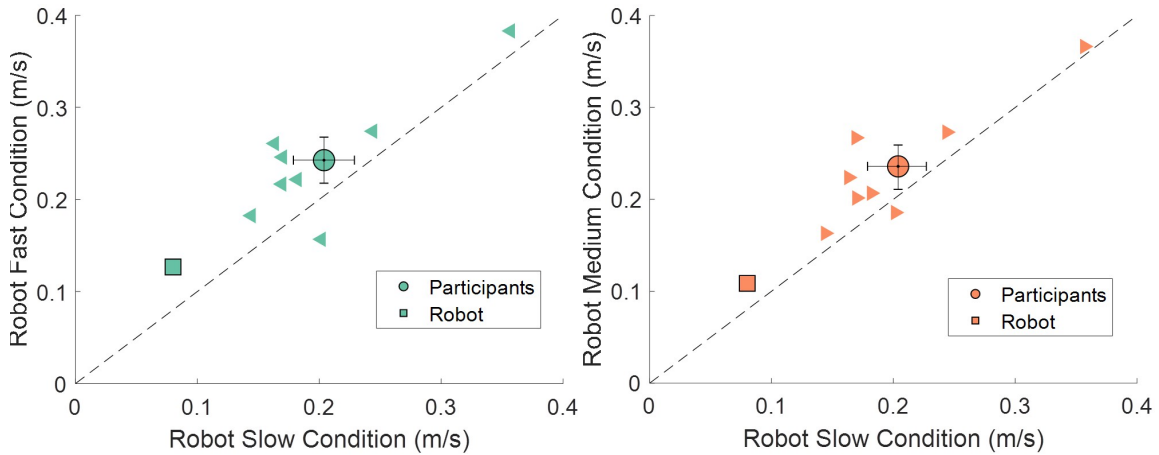


Figure 8: Individual speed of subjects in the “Robot Fast” and “Robot Mid” speed condition, in relation to their speed in the “Robot Slow” speed condition (“One Box”). Same conventions as Figure 5. (Japan)

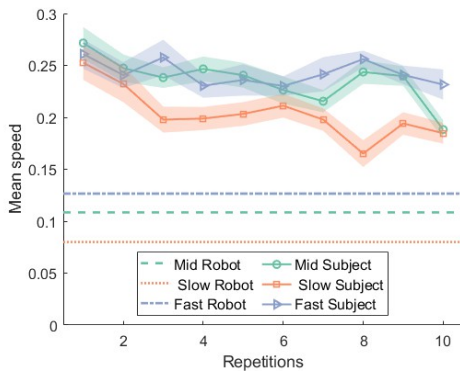


Figure 9: Mean subject’s speeds for each of the ten repetitions (“OneBox”), with shaded areas representing error bars. The dashed lines represent the speeds of the robot, equal across all ten repetitions (Japan).

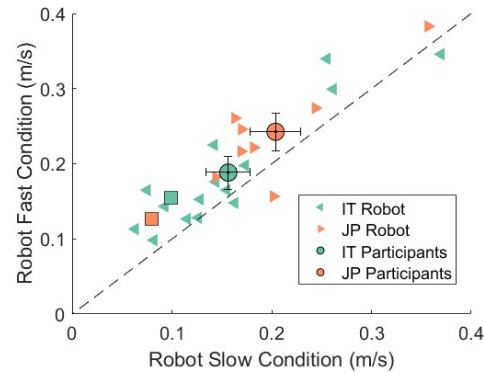


Figure 10: Individual subjects’ speed in the “Robot Fast” speed condition, displayed in relation to the corresponding speed in the “Robot Slow” speed condition (in the “One Box” condition) for both countries. The bigger circles with error bars represent the mean and standard error. If a subject had the exact same speed in the two conditions, the respective marker would be on the dashed line.

the cultural context in which the interaction takes place. Japanese and Italian cultures differ substantially in their acceptance toward robots. Considering that a strong link exists between acceptance and other mechanisms such as imitation [32] we hypothesized a higher degree of speed adaptation in the Japanese sample, when compared to the Italian one. Our results highlighted two main behaviors. The first and most important one is that people changed their movement speed according to the speed of the robot. Not only were they faster when the robot was fast, and slower when the robot was slow, but also their intermediate speed was in the middle of the other two. This is in line with results demonstrating that humans tend to exhibit speed adaptation when interacting with an embodied robotic partner, [8, 42] showing that the phenomenon generalizes to simultaneous actions in a semi-shared space.

Beyond a general difference in average speed between participants in Italy and Japan, with the Japanese moving on average faster during the experiment, the degree of adaptation to robot speed was comparable between the two groups. According to these results, we found no evidence for a cultural modulation of the degree of adaptation in the task at hand. The second result is that sharing the target space for an action did not affect the behavior of the participants, neither for Italians nor for Japanese.

These findings suggest that even though a top-down modulation of the activation of the neural substrates of motor resonance has been proven for social group membership and contextual representations [28-31], the phenomenon of automatic imitation during collaboration is preserved similarly within the two different cultural envi-

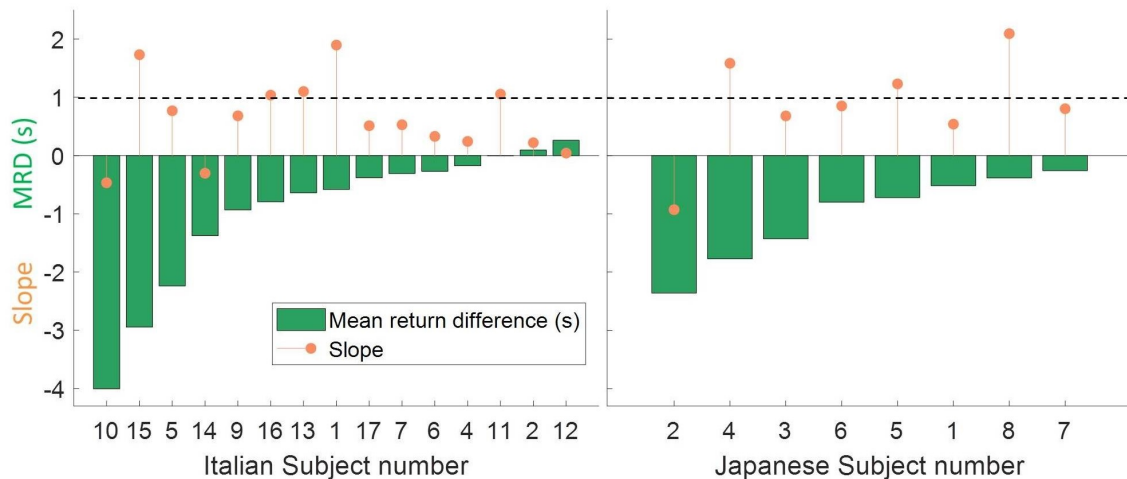


Figure 11: Stem plots represent the slope corresponding to the amount of adaptation of subjects. These slopes have been computed for the one-box condition, but those computed for the two-box condition are very similar. Bars represent the “Mean return difference” (MRD), that is the mean difference in timing between the instants in which the robot and the subject returned their hand to the starting position to get a new block. A negative value means that the subject arrived first.

ronments tested. This might be due to the high pro-social value of such low-level mechanism – which could represent a building block for the establishment of efficient collaboration on top of which different cultures develop different social constructs.

It is worth noting that we did not directly assess automatic imitation in human-human interactions but rather addressed human-robot interaction. Although there is evidence of a strong similarity in automatic velocity imitation in these two contexts [8], it is yet possible that a potential cultural top-down effect would have been present in the context of a person-to-person interaction, but that does not generalize to person-to-artificial agent collaborations. Though we cannot exclude this possibility, we would expect the opposite phenomenon. Indeed, Japanese and Italian cultures differ substantially in the way technology and in particular robots are perceived [17]. Hence, we would have predicted a larger impact of culture when the interaction involves such novel artificial agents. Future research will be needed to explore the potential different impact of culture on basic implicit mechanisms of interaction in human-human and human-machine settings.

Considering the absence of any difference between the conditions with one or two boxes, we believe that this could be due to the fact that, although the single box was a common space, it did not lead to any conflict of resources use during the execution of the task. Even if the single big box was shared by the subject and the robot, they could drop the blocks on their side without having a spatial encounter with the other.

The task described has been designed to be an ecological interaction with the iCub, considering that many previous similar experiments of cooperation involved only robotic arms or industrial scenarios [43, 44].

4.1 Limitations

The presence of a third agent with the active role of passing blocks could have influenced the interaction, even though the experimenters were trained to release the objects in the hand of the participant and robot in a stereotyped manner which was constant across all conditions. Moreover participants witnessed only the release of blocks into their hands since the experimenter picked up the blocks while subjects were completing the previous trial looking at the robot or at the box. For this reason we believe it is unlikely that the speed of the experimenter had influenced the participants. Future experiments will avoid the presence of another person that could bias the results. Another problem to discuss is the number of participants. Availability in Japan was limited thus we could not have the same amount of subjects in the two countries. Overall, testing a higher amount of people in the future would lead to more reliable results and allow a more detailed data analysis, possibly strengthening our findings and improving the comparison between the two countries. It is important to note that the selected task might not be considered a proper joint action, such that the two agents overtly collaborate to achieve a shared goal. However the two have to perform a simultaneous action in a quasi-shared space and this entails

a certain degree of coordination. It could be relevant in the future to assess whether a more explicitly collaborative task could be more affected by cultural difference. Last, in this experiment we only measured the implicit adaptation and we did not enquire about the subjective perception of the participants. Such information could help interpreting more in depth the behavioral results.

5 Conclusion

Humans are very skilled when it comes to cooperating during a joint task. A mix of implicit emergent coordination and explicit signals allows for a good collaboration. The focus of this study was to investigate whether the emergent adaptation could be triggered during a joint task with a humanoid robot in two culturally different countries. From the analysis of how humans' speed changed as a function of robot's speed, both Italian and Japanese subjects exhibited the same degree of adaptation. This suggests that this form of coordination, managed by low-level processes in the brain, is not influenced by cultural background. Future works are needed to explore the role of robot behavior in interaction with different cultures, with particular reference to the differential impact of the nature of the agent's arm motion (biological inspired or not) and the concurrent bodily behavior, such as gaze movements.

The implementation of naturalistic social behaviors such as mutual gaze on a robot can potentially enhance the quality of the interaction with this robot. When using such behaviors it is important to take the specificities of the cultural context in which the robot is used into consideration. The same behavioral pattern that might facilitate human-robot interaction in one cultural environment could have an inhibiting effect in another.

Acknowledgements: This research has been conducted in the framework of the European Project CODEFROR (FP7-PIRSES-2013-612555), ROBOCOM++ (FLAG-ERA Joint Transnational Call (JTC) 2016), and JST CREST Cognitive Mirroring (Grant number: JPMJCR16E2). We are grateful to all study participants for their contributions and to Dr. Kostantinos Theofilis for the support during the experiment.

References

- [1] SoftBank Robotics, NAO [Online], Available: <https://www.ald.softbankrobotics.com/en>
- [2] SoftBank Robotics, Pepper [Online], Available: <https://www.ald.softbankrobotics.com/en>
- [3] D. Mondou, A. Prigent, A. Revel, A dynamic scenario by remote supervision: a serious game in the museum with a Nao robot, In: *Advances in Computer Entertainment Technology (ACE2017)*, London, UK, 2017, 103–116
- [4] M. Niemelä, A. Arvola, I. Aaltonen, Monitoring the acceptance of a social service robot in a shopping mall: first results, In: *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction (HRI '17)*, 2017, 225–226
- [5] A. Sciutti, M. Mara, V. Tagliasco, G. Sandini, Humanizing human-robot interaction: On the importance of mutual understanding, *IEEE Technology and Society Magazine*, 2018, 37(1), 22–29
- [6] R. Blake, M. Shiffrar, Perception of human motion, *Annual Review of Psychology*, 2007, 58, 47–73
- [7] G. Sandini, A. Sciutti, F. Rea, Movement-based communication for humanoid-human interaction, In: A. Goswami, P. Vadakkepat (Eds.), *Humanoid Robotics: A Reference*, Springer, Dordrecht, 2018, 2169–2197
- [8] A. Bisio et al., Motor contagion during human-human and human-robot interaction, *PLoS One*, 2014, 9(8)
- [9] H. Lehmann, Y. Nagai, G. Metta, The Question of Cultural Sensitive Gesture Libraries in HRI - An Italian - Japanese Comparison, In: *Proceedings of the ICDL-EpiRob 2016 – Workshop on Vision and the Development of Social Cognition*, 2016
- [10] S. Martinez-Conde, S. L. Macknik, D. H. Hubel, The role of fixational eye movements in visual perception, *Nature Reviews Neuroscience*, 2004, 5, 229–240
- [11] H. Lehmann, I. Keller, R. Ahmadzadeh, F. Broz, Naturalistic Conversational Gaze Control for Humanoid Robots – A First Step, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2017
- [12] M. Argyle, J. Dean, Eye-Contact, Distance and affiliation, *Sociometry*, 1965, 28(3), 289–304
- [13] A. Bisio, N. Stucchi, M. Jacono, L. Fadiga, T. Pozzo, Automatic versus voluntary motor imitation: Effect of visual context and stimulus velocity, *PLoS One*, 2010, 5(10), e13506
- [14] C. Bartneck, T. Nomura, T. Kanda, T. Suzuki, K. Kato, Cultural differences in attitudes towards robots, In: *Proceedings of the AISB Symposium on Robot Companions, Hard Problems and Open Challenges in Human-Robot Interaction*, Hatfield, 2005, 1–4
- [15] D. Li, P. P. Rau, Y. Li, A cross-cultural study: Effect of robot appearance and task, *International Journal of Social Robotics*, 2010, 2(2), 175–186
- [16] S. Šabanović, C. C. Bennett, H. R. Lee, Towards culturally robust robots: A critical social perspective on robotics and culture, In: *Proceedings of HRI Workshop on Culture-Aware Robotics*, Bielefeld, 2014
- [17] F. Kaplan, Who is afraid of the humanoid? Investigating cultural differences in the acceptance of robots, *International Journal of Humanoid Robotics*, 2004, 1(03), 465–480
- [18] G. Knoblich, S. Butterfill, N. Sebanz, *Psychological research on joint action: theory and data*, Psychology of Learning and Motivation, Academic Press, 2011, 54, 59–101
- [19] G. Rizzolatti, L. Craighero, The mirror-neuron system, *Annual Review of Neuroscience*, 2004, 27(1), 169–192
- [20] L. Craighero, G. Metta, G. Sandini, L. Fadiga, The mirror-neurons system: data and models, *Progress in Brain Research*, 2007,

- 164, 39–59
- [21] Y. Kanakogi, S. Itakura, Developmental correspondence between action prediction and motor ability in early infancy, *Nature Communications*, 2011, 2, 341
- [22] T. Falck-Ytter, G. Gredebäck, C. von Hofsten, Infants predict other people's action goals, *Nature Neuroscience*, 2006, 9(7), 878–879
- [23] M. Bove, A. Tacchino, E. Pelosin, C. Moisello, G. Abbruzzese, M. F. Ghilardi, Spontaneous movement tempo is influenced by observation of rhythmical actions, *Brain Research Bulletin*, 2009, 80(3), 122–127
- [24] L. Noy, E. Dekel, U. Alon, The mirror game as a paradigm for studying the dynamics of two people improvising motion together, In: *Proceedings of the National Academy of Sciences*, 2011, 108(52), 20947–20952
- [25] A. Sciutti, A. Bisio, F. Nori, G. Metta, L. Fadiga, T. Pozzo, G. Sandini, Measuring human-robot interaction through motor resonance, *International Journal of Social Robotics*, 2012, 4(3), 223–234
- [26] S. Kashi, S. Levy-Tzedek, Smooth leader or sharp follower? Playing the mirror game with a robot, *Restorative Neurology and Neuroscience*, 2018, 36(2), 147–159
- [27] A. Sciutti, G. Sandini, Interacting with robots to investigate the bases of social interaction, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2017, 25(12), 2295–2304
- [28] L. Amoruso, C. Urgesi, Contextual modulation of motor resonance during the observation of everyday actions, *Neuroimage*, 2016, 134, 74–84
- [29] L. Amoruso, A. Finisguerra, C. Urgesi, Tracking the time course of top-down contextual effects on motor responses during action comprehension, *Journal of Neuroscience*, 2016, 36(46), 11590–11600
- [30] B. Rauchbauer, J. Majdandžić, A. Hummer, C. Windischberger, C. Lamm, Distinct neural processes are engaged in the modulation of mimicry by social group-membership and emotional expressions, *Cortex*, 2015, 70, 49–67
- [31] P. Molenberghs, V. Halász, J. B. Mattingley, E. J. Vanman, R. Cunnington, Seeing is believing: Neural mechanisms of action-perception are biased by team membership, *Human Brain Mapping*, 2013, 34(9), 2055–2068
- [32] T. L. Chartrand, J. L. Lakin, The antecedents and consequences of human behavioral mimicry, *Annual Review of Psychology*, 2013, 64, 285–308
- [33] M. Baldassarre, S. Feller, Cultural variations in personal space: theory, methods, and evidence, *Ethos*, 1975, 3(4), 481–503
- [34] G. Metta, L. Natale, F. Nori, G. Sandini, The iCub project: An open source platform for research in embodied cognition, In: *Proceedings of IEEE Workshop on Advanced Robotics and its Social Impacts, ARSO*, 2011, 24–26
- [35] P. Viviani, T. Flash, Minimum-jerk, two-thirds power law, and isochrony: converging approaches to movement planning, *Journal of Experimental Psychology: Human Perception and Performance*, 1995, 21(1), 32–53
- [36] G. Catavittello, Y. P. Ivanenko, F. Lacquaniti, P. Viviani, Drawing ellipses in water: evidence for dynamic constraints in the relation between velocity and path curvature, *Experimental Brain Research*, 2016, 234(6), 1649–1657
- [37] N. Noceti, F. Rea, A. Sciutti, F. Odone, G. Sandini, View-invariant robot adaptation to human action timing, In: *Proceedings of SAI Intelligent Systems Conference*, Springer, Cham, 2018, 804–821
- [38] U. Pattacini, F. Nori, L. Natale, G. Metta, G. Sandini, An experimental evaluation of a novel minimum-jerk Cartesian controller for humanoid robots, In: *Proceedings of IEEE/RSJ 2010 International Conference on Intelligent Robots and Systems, IROS*, 2010, 1668–1674
- [39] T. Chaminade, D. W. Franklin, E. Oztop, G. Cheng, Motor interference between humans and humanoid robots: Effect of biological and artificial motion, In: *Proceedings of 2005 4th IEEE International Conference on Development and Learning*, 2005, 96101
- [40] G. Metta, P. Fitzpatrick, L. Natale, YARP – Yet Another Robot Platform, version 2.3.20, *International Journal of Advanced Robotic Systems*, 2006, 3(1), 8
- [41] ICubForwardKinematics, 2014, <http://wiki.icub.org/wiki/ICubForwardKinematics>
- [42] D. Eizicovits, Y. Edan, I. Tabak, S. Levy-Tzedek, Robotic gaming prototype for upper limb exercise: Effects of age and embodiment on user preferences and movement, *Restorative Neurology and Neuroscience*, 2018, 36(2), 261–274
- [43] A. Cherubini, R. Passama, A. Meline, A. Crosnier, P. Fraise, Multimodal control for human-robot cooperation, In: *Proceedings of IEEE International Conference on Intelligent Robots and Systems*, 2013, 2202–2207
- [44] B. Nemeč, N. Likar, A. Gams, A. Ude, Bimanual human robot cooperation with adaptive stiffness control, In: *Proceedings of IEEE-RAS International Conference on Humanoid Robots*, 2016, 607–613