Understanding Robots: Making Robots More Legible in Multi-User Interactions

Miguel Faria[1] Francisco S. Melo[1] Ana Paiva[1]

***Abstract*— In this work we explore implicit communication between humans and robots—through movement—in multi-user interactions. In particular, we investigate how can a robot move to better convey its intentions using legible movement. We consider a factor seldom considered in human-robot interaction settings which largely impacts the ability of human users to assign meaning to a robot's movement: the number of users the robot has to simultaneously interact with. We propose an extension the notion of legible motion that considers that the legibility of a movement depends on all human users involved in the interaction, and should take into consideration how each of them perceives the robot's movements from their respective points-of-view. We show, through simulation and a user study, that our proposed model of multi-user legibility leads to movements that, on average, optimize the legibility of the motion as perceived by the group of users. Our model creates movements that allow each human to more quickly and confidently understand what are the robot's intentions, thus creating safer, clearer and more efficient interactions and collaborations.**

I. Introduction

Robots are an increasingly more common sight in everyday lives in the most diverse capabilities, from toys like Anki's Vector to iRobot's Roomba vaccum cleaner to Paro the Therapeutic Robot. Thus, research in the field human-robot interaction and its possible applications has been important for a correct integration of robots and adapting them to correctly assist in tasks like healthcare [1]–[3], education [4], [5], entertainment [6], among others. However, to fully integrate robots in society, they need to be able to correctly interact and communicate with humans. The need for correct communication is essential for interactions between humans and robots to be safe [7], natural [8] and efficient [9], [10].

Humans communicate in a myriad of different ways: explicitly through speech or haptic communication; or implicitly, through body movements or by proactively executing an action or task. Implicit communication, in particular, requires interpretation by the receiver on the sender's intentions [11]. The combination of explicit and implicit communication allows humans to easily interact with each other, even in situations where no explicit signals are traded—in many collaborative tasks, all parties work together towards the same goal without always explicitly communicating their current objectives or intentions. Therefore, robots involved in human-robot interaction must also leverage explicit and implicit means of communication for those interactions to feel normal by humans [12].

In fact, leveraging both implicit and explicit communication has been shown to have positive impacts in human-robot interactions. Breazeal et al [13] showed how the use of a combination of implicit and explicit communication, in a collaborative task, caused less errors and less time to finish the task than using only explicit communication. Likewise, Liang et al [14] deployed an AI agent that used *implicature*—an implicit communication where utterances imply information beyond what the words literally convey—in a collaborative game, achieving more natural-like interactions. In [15], [16] robots used signaling and body movements to convey intentions, thus overcoming limitations caused by their embodiment.

Nonverbal communication through movement is an essential mean of implicit communication, because it is cross-cultural and most humans comprehend without being previously explicitly taught [17]. The use of body movement by a robot can lead to people considering it more socially engaging, warm, friendly and empathetic; as well as more expressive when conveying different emotions. The correct use of body movement can also improve task performance, causing less errors and better completion times [18]. Both [19], [20] use expressive movements, called *legible movements*, to allow the human partners to better understand the robot's objectives. In [15], expressive movements mimicking those of home pets improved a Sphero robot's communication capabilities, allowing humans to understand they should follow the robot to another room. Che et al [21] combine explicit communication and body movements to make a robotic mobile platform navigate through an environment shared with humans without colliding with them and reducing human effort in predicting the robot's intentions.

Legible movements, defined by Dragan et al [22], are movements that augment a robot's expressiveness: they are easier "to read" by human users, as they exaggerate the robot's trajectory to remove any ambiguity about what the robot's objective is. Legible movements have been explored to understand their true impacts and advantages in human-robot interactions. In [19], [20], legible movement was explored as a means to improve the efficiency of a team in human-robot collaboration tasks. Busch et al [23] explored how a robot can improve its legibility by learning while interacting with humans to improve their prediction capabilities of the robot's objectives. Legible motion is also powerful because it increases the transparency of a robot's state, as evidenced in [24] where a robot used legible movements to

[1] All authors are with INESC-ID and Instituto Superior Técnico, University of Lisbon, Portugal. E-mail: miguel.faria@tecnico.ulisboa.pt and {fmelo,ana.paiva}@inesc-id.pt.

express its limitations and inabilities.

Most works explore the uses and advantages of legible movements in single-user scenarios [19], [23]. However, when integrated in society, a robot must be expressive both in single and multi-user scenarios, as would happen in a surgery scenario [2], or in an educational scenario where a robot would have to interact with multiple learners. Given evidences that support the use of legible movement in multi-user scenarios for specific tasks [20], we find necessary to further explore how legible movements would impact multi-user interactions.

In this work we contribute an extension to the concept of legibility to multi-user scenarios. Our proposal reduces to the standard notion of legibility in single user scenarios while, in multi-user scenarios, extends this notion in a natural way to accommodate for the fact that the different users will observe the robot's motion from a different point-of-view and, as such, will interpret the robot's motion in a different way. Our proposal considers the legibility of a movement to be the average legibility of that same from the point-of-view of each of the human users involved. By optimizing the average value of legibility across points-of-view, the movement performed by the robot will be intelligible for all human users involved in the interaction.

The paper is organized as follows. Section II defines legible movement and discusses how to create legible movements considering a user's point-of-view (POV). Section III introduces legibility in multi-user scenarios, discussing the drawbacks of considering only the POV of a single user, and continues by establishing how to generate legible movements that take into consideration multiple users. Section IV presents the results of two different validations: Section IV-A presents simulation results that compare the legibility of the multi-user approach versus the traditional single-user approach in a multi-user interaction scenario; Section IV-B then presents the results of a user study conducted on Amazon's Mechanical Turk (M-Turk), using the same scenario as in Section IV-A.

II. Legible Motion

The concept of *legible motion* was introduced by Dragan et al [22] as a type of movement that allows a robot to be more expressive during interaction with humans, drawing on the fact that humans interpret an action's objective in a *goal-oriented* manner [25]. This means that humans, when observing an agent performing an action, consider the agent to be logical and thus will execute the action in the most efficient way possible. Leveraging on this fact, a legible movement tries to emphasize the movement's objective by bringing the robot as close as possible to the objective and at the same time as far away as possible from other possible objectives as soon as possible. To achieve this emphasis on the objective, legible movements leverage animation principles to create trajectories that are readable [26] and that encourage "anticipatory" movements [27]. Such trajectories convey the most relevant information for goal prediction in the beginning of the movement.

A legible movement is thus a movement that, when observed by a human, allows the human to infer the objective G_R as soon as possible given the observed trajectory ξ. In other words, a legible movement (towards goal G_R) should maximize the value

$$Legibility(\xi) = \frac{\int P(G_R|\xi_{S\to\xi(t)})f(t)dt}{\int f(t)dt}, \tag{1}$$

where $P(G_R|\xi_{S\to\xi(t)})$ is the likelihood of objective G_R given the partial trajectory between the start point S and the current point $\xi(t)$, denoted as $\xi_{S\to\xi(t)}$. The function f is a weighting function, that gives more weight to earlier parts of the trajectory.

Intuitively, the term $P(G_R|\xi_{S\to\xi(t)})$ in (1) seeks to capture the likelihood that a human would assign to goal G_R given the trajectory up to time step t [22], and is modeled using a max-entropy distribution

$$P(G_R|\xi_{S\to\xi(t)}) = \frac{1}{Z}\exp\left\{-C(\xi_{S\to\xi(t)})\right\}, \tag{2}$$

where Z is a normalizing constant and C is a cost function modeling how a human expects the robot to move in an efficient manner. Following [28], one possible cost function is the sum of squared velocities

$$C(\xi) = \frac{1}{2}\int \|\xi'(t)\|^2 dt$$

that encourages smooth trajectories that go straight to the goal.

Using the definition of legibility in (1), it is possible to generate legible trajectories using a gradient ascent approach that, in each iteration, improves the legibility score of the trajectory ξ:

$$\xi_{i+1} = \xi_i + \frac{1}{\eta}M^{-1}\nabla Legibility(\xi), \tag{3}$$

where M is used to measure the norm of a trajectory, $\|\xi\|_M^2 = \xi^T M \xi$.

The concept of legible movement as defined in [22] assumes humans have an omniscient view of the workspace. This assumption can lead to movements that go outside the field-of-view (FOV) of a human, go through human blind spots or through obstructed parts of the workspace—in other words, leads to trajectories that are not as legible as expected. To solve this problem, Nikoladis et al [29] propose an extension of the original notion of legibility that uses a modified cost function $\bar{C}$, defined as

$$\bar{C}(\xi) = C(T(\xi)), \tag{4}$$

where $T(\xi)$ transforms ξ from world coordinates to the human referential and then projects the 3D points describing the trajectory into 2D points in the human's FOV.

With the extension, the legibility metric becomes dependent of the point of view of the human user. The transformation T to the human's viewpoint allows the optimization procedure to improve the legibility, creating trajectories that are always within the user's FOV. This improves the human user's ability to correctly predict the robot's objective

[29]. We henceforth refer to this approach as **Single-User Legibility (SUL)**.

III. Multi-User Legible Motion

As discussed in Section II, legible motions are defined to allow robots to be more expressive, by leveraging animation principles to create movements that convey intentions as soon as possible. These motions have been tested in single-user scenarios: scenarios where a robot interacts with one human at a time.

However, there are several scenarios where a robot must interact with multiple humans at the same time. For example, in [6] a robot plays a game of cards, simultaneously interacting with three human users; [2] describe a scenario in which a robot is deployed as part of a surgical team to support the staff; [20] describe a scenario where a robot sequentially serves cups of water to different human users. In all of these scenarios, for a robot to use legible motions it must be able to generate movements that are simultaneously legible for all partners involved; otherwise, it could optimize the legibility for one partner but reduce the legibility for another one.

For legible motions to be legible for multiple users simultaneously, we propose an extension to SUL where instead of considering the point-of-view of a single user in the computation of the legibility metric, we consider the point of view of all users when evaluating the legibility of a trajectory. We thus propose the legibility metric in multi-user settings to be the average value of the legibility perceived across the different users. In other words, when we improving the legibility of a trajectory, we maximize its legibility as perceived by the group and not a by single user. Such modification allows humans to better understand the robot, leading to better cooperation within the group. We name this new approach to legibility **Multi-User Legibility (MUL)**.

A. Definition of Multi-User Legibility

In MUL, we define the legibility of a trajectory ξ in a setting comprising N users as

$$Legibility_{MUL}(\xi) = \frac{1}{N}\sum_{n=1}^{N} Legibility_n(\xi), \qquad (5)$$

where $Legibility_n(\xi)$ is the SUL of trajectory ξ as perceived by user n. With MUL the expression for the update in the optimization procedure becomes

$$\xi_{i+1} = \xi_i + \frac{1}{\eta} M^{-1} \nabla Legibility_{MUL}(\xi), \qquad (6)$$

with

$$\nabla Legibility_{MUL}(\xi) = \frac{1}{N}\sum_{n=1}^{N} T_n^{-1} \cdot \nabla Legibility_{SUL_n}(\xi), \qquad (7)$$

where T_n^-1 is the transformation from the human partner to the coordinate space where ξ was defined and

$$\nabla Legibility_{SUL_n}(\xi) = \nabla \mathcal{L}(\xi(t)) \cdot \nabla T_n(\xi(t)), \qquad (8)$$

where $\nabla T_(\xi(t))$ is the gradient for the transformation of the trajectory ξ to the perspective of the human partner in question and $\nabla \mathcal{L}(\xi(t))$ is the gradient of (1).

IV. Experimental Evaluation

We validated our proposed model in two stages: first we defined a simulation scenario where we compared the legibility of a trajectory optimized using MUL, versus the average legibility of trajectories that used SUL; second, we used the same scenario, created videos for trajectories executed by a robot using MUL and SUL and tested how the trajectories obtained differed for real humans.

A. Simulations

The scenario for the simulations was an interaction where a robot had to grab one of three objects on a table - a soda can, a rubber duck and a telephone - and three humans had to predict which object the robot was going to grab. We defined four different configurations of object positioning and human poses, under the restriction that the human orientation was such that their FOV captured both the robot and the objects on the table. Also, since we were modeling a real world scenario, the humans were oriented as if looking in a slightly downward manner as they would in real life; this downward angle would vary between 10^o and 15^o depending on the distance between the human and the objects on the table. The different configurations are showed in Figure 1 and represent configurations of objects that one could face in real world interactions with the robot.

For each configuration, we ran four optimizations on legibility - one using MUL and three using SUL applied to each of the defined points of view - starting with a straight line from the robot's starting position to the target of the optimization. This process was repeat once for each object. The optimization process stopped either when the maximum number of iterations was reached or if the difference between improvements was less than a defined threshold. After each optimization finished, we evaluated the obtained trajectory's legibility for each user.

*1) **Results and Analysis**:* Tables I, II and III present the results for the simulations. Each table presents the legibility results for one of the three targets: Table I the soda can, Table II the rubber duck and Table III the telephone; organized by the configuration and the optimization applied. When evaluating the trajectory's legibility for each user, if the trajectory went outside a user's FOV it was attributed a legibility value of 0. This attribution was based on the fact that trajectories outside the FOV can lead to unsafe and harmful interactions and an incorrect communication of intentions.

The legibility values in Tables I, II and III show that MUL generates trajectories that on average are more legible for the entire group than if we optimize using SUL applied to each human partner. Also, several optimizations using SUL created trajectories that would go outside the one or multiple of the other partners' FOV - like the case of User 2 in configuration 1 for the telephone where the trajectory had 0 legibility for both users 1 and 3. Sometimes SUL creates trajectories that are better for the team than MUL. This is possible if a user has a privileged perspective on the task - the human partner's FOV gives a very good perspective

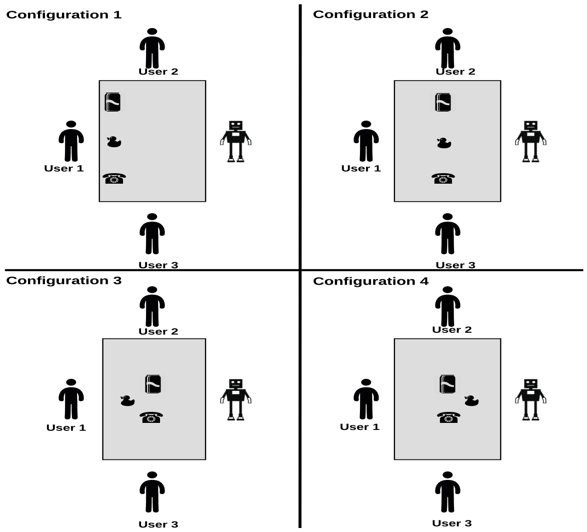

Fig. 1: The different target and user configurations. **User 1** is opposite to the robot, **User 2** to the right of the robot and **User 3** to the left of the robot. The configurations presented are numbered from configuration 1 in the top left corner to configuration 4 in the bottom right.

TABLE I: Table, for the soda can target, with the legibility values for each user and their average, organized by configuration (y-axis) and optimization used (x-axis).

		MUL	SUL User 1	SUL User 2	SUL User 3
1	User 1	0.86698	0.85394	0.48204	0
	User 2	0.93454	0	0.59744	0
	User 3	0.862486	0.84119	0.45229	0.93651
	Average	**0.888**	**0.56504**	**0.51059**	**0.31217**
2	User 1	0.46814	0.58293	0.27584	0.73607
	User 2	0.51256	0.87007	0.8644	0
	User 3	0.59155	0.67691	0.03682	0.93358
	Average	**0.52408**	**0.70997**	**0.39235**	**0.55655**
3	User 1	0.57381	0.87294	0.31076	0.55507
	User 2	0.33217	0	0.83371	0.03175
	User 3	0.8118	0.93651	0.0331	0.93651
	Average	**0.57259**	**0.60315**	**0.39252**	**0.50777**
4	User 1	0.48004	0.77565	0.43542	0.03194
	User 2	0.90923	0	0.50269	0
	User 3	0.52480	0.04127	0.40403	0.92770
	Average	**0.63803**	**0.27231**	**0.44738**	**0.31988**

TABLE II: Table, for the rubber duck target, with the legibility values for each user and their average, organized by configuration (y-axis) and optimization used (x-axis).

		MUL	SUL User 1	SUL User 2	SUL User 3
1	User 1	0.56878	0.56795	0.56315	0.56315
	User 2	0.47615	0.47969	0.48196	0.46095
	User 3	0.47615	0.47969	0.46095	0.48196
	Average	**0.50703**	**0.50911**	**0.50202**	**0.50202**
2	User 1	0.36148	0.3615	0.36147	0.36147
	User 2	0.38534	0.3853	0.38535	0.38532
	User 3	0.38534	0.3853	0.38532	0.38535
	Average	**0.37738**	**0.37737**	**0.37738**	**0.37738**
3	User 1	0.57381	0.87294	0.55507	0.31076
	User 2	0.8118	0.93651	0.9365	0.0331
	User 3	0.33217	0	0.03175	0.83371
	Average	**0.57259**	**0.60315**	**0.50777**	**0.39252**
4	User 1	0.84706	0.93648	0.5144	0.5144
	User 2	0.29202	0	0.39081	0.3879
	User 3	0.29202	0	0.3879	0.39081
	Average	**0.47704**	**0.31216**	**0.43103**	**0.43103**

TABLE III: Table, for the telephone target, with the legibility values for each user and their average, organized by configuration (y-axis) and optimization used (x-axis).

		MUL	SUL User 1	SUL User 2	SUL User 3
1	User 1	0.86698	0.85082	0	0.48324
	User 2	0.86249	0.84132	0.93651	0.47535
	User 3	0.93454	0	0	0.58933
	Average	**0.888**	**0.56405**	**0.31217**	**0.51597**
2	User 1	0.46814	0.75825	0.7569	0.26634
	User 2	0.59155	0.8486	0.93482	0.03186
	User 3	0.51257	0	0	0.93414
	Average	**0.52408**	**0.53562**	**0.56391**	**0.41078**
3	User 1	0.81781	0.65822	0.93334	0.93334
	User 2	0.36397	0.36358	0.45865	0.3339
	User 3	0.36397	0.36358	0.3339	0.45865
	Average	**0.51525**	**0.4618**	**0.5753**	**0.5753**
4	User 1	0.48004	0.77565	0.03194	0.43542
	User 2	0.52480	0.04127	0.92770	0.40403
	User 3	0.90923	0	0.00000	0.50269
	Average	**0.63803**	**0.27231**	**0.31988**	**0.44738**

on the targets and the robot - allowing SUL to take better advantage of the POV than MUL because MUL is weighted down by the other partners' POVs. We hypothesize that a possible solution to this is to give each partner a weight depending on how good the POV is, leaving this for future work.

B. User Study

*1) **Scenario and Hypotheses**:* We conducted a user study through M-Turk to validate two working hypotheses:

- ***H1:*** Participants will consider a movement optimized using MUL clearer than when optimized using SUL.
- ***H2:*** Participants will understand quicker and more confidently the robot's target, when faced with a movement optimized with MUL than with one optimized with SUL.

The study followed a between-subjects design, with each group being a different optimization approach - optimized with MUL or optimized with SUL for each POV. Each participant watched 3 sets of 3 videos, each video with different length - 6, 12 and 18 seconds. After each video the participant was asked to predict what object the robot was going to grab and rate how confident they were in their prediction. After watching the 3 sets of videos the participant would have to rate how clear the three movements were. The videos were created using the WeBots simulator, where we recreated configuration 2 from Figure 1.

*2) **Results and Analysis**:* We recruited 315 participants using M-Turk, after removing those that failed the control question, with approximately 98% from the USA and the remaining 2% distributed across Canada, Australia and the

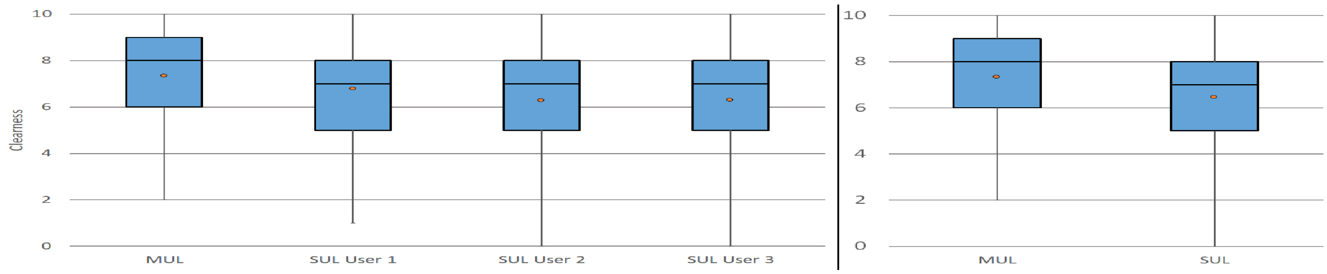

Fig. 2

UK. We restricted the participation on the questionnaire to these four countries to reduce language barrier problems. The participants' age varied from 23 to 76 years old and an average of 42 years old. An analysis of the education level of the participants shows that 60% has a higher education degree and 90% has finished at least high school. Regarding the occupation of the participants, 94% are employed and 2% are students.

The questionnaires were analysed in two steps. The first compared the results of MUL against each of the 3 instances of SUL, analysing how MUL individually fared against SUL and the second analysis combined the results of the three applications of SUL comparing them with MUL, evaluating how MUL fared against the overall performance of SUL and attenuating the impacts of privileged and bad POVs. For both steps we measured the average perceived clearness rate at which the participants evaluated the movements, the time taken to correctly predict the robot's target and the average confidence in the prediction. We considered that a wrong prediction of the target was considered as taking the full length of the video to make a decision and that for correct predictions the time taken corresponded to the earliest a participant answered correctly without wrongly predicting afterwards. To analyse the confidence in the prediction we combined the scores for each 6, 12 and 18 second videos as in [29]. We also conducted a normality test that showed us that all three measures - average perceived clearness, time taken and confidence - did not follow a normal distribution.

To verify hypothesis **H1**, we conducted a Kurskal-Wallis test comparing MUL individually with all the applications of SUL showing a significant difference between the four optimizations, $\chi^2(3) = 9.344, p = 0.025$. A follow-up analysis with a Mann-Whitney test to determine where this difference existed, showed that MUL was rated as significantly clearer than the application of SUL to User 2 ($U = 2250, p = 0.009$) and User 3 ($U = 2333.50, p = 0.008$), thus supporting hypothesis **H1** for both users. For the comparison between MUL and the aggregated SUL applications, we conducted a Mann-Whitney test that showed MUL was considered significantly clearer by users than SUL, $U = 7277.5, p = 0.006$, thus supporting hypothesis **H1**. Figure 2 shows the result distribution of perceived clearness for both analysis: the individual on the left graph and the aggregated on the right graph, with the average for each marked with a point.

For hypothesis **H2**, we analysed both the time taken to predict what target the robot was going to grab and the confidence associated with said prediction. Again, both measures did not follow a normal distribution so we applied the non-parametric Kurskal-Wallis test to find differences between the four optimizations. The Kurskal-Wallis test showed significant differences between the four optimizations for both the time ($\chi^2(3) = 31.877, p < 0.001$) and confidence measures ($\chi^2(3) = 48.821, p < 0.001$), supporting hypothesis **H2**. To further understand where both measures are most significant we applied a Mann-Whitney test, comparing the MUL with each SUL. For the time to correctly predict, we found participants paired with MUL took significantly less time than those paired with SUL for User 2, $U = 21244, p < 0.001$. For the confidence in the predictions, we found that participants were significantly more confident in their predictions with MUL than with SUL for User 2, $U = 11379, p < 0.001$, and for User 3, $U = 14367.5, p = 0.002$, as evidenced in Figure 3 left graph. For the comparison between MUL and the SUL aggregation, we conducted a Mann-Whitney test for both the time and confidence measures. The test showed people took significantly less time to correctly predict the robot's target with MUL optimization than with SUL optimization, $U = 75722, p = 0.037$. Regarding confidence in prediction, the Mann-Whitney test and the right graph in Figure 3 show participants are significantly more confident in their predictions with MUL than with SUL, $U = 43971.5, p < 0.001$. Both results support hypothesis **H2**.

V. Conclusion

In this work we tackled the use of legible movement in multi-user interactions, proposing a model - MUL - that creates movements which improve the legibility of the group instead of each user individually. Using simulations and a user study, we showed that MUL creates movements with better average legibility and allows participants to better understand a robot's intention than standard single-user approaches: creating clearer movements and allowing faster and more confident predictions of the robot's intentions.

MUL is a promising model for applications in entertainment areas such as theatre where movement is an important communication tool or healthcare areas like surgery teams where a robot needs to correctly communicate for the rest of the medical staff to perform adequately.

Finally, we should highlight that if there is just one human interacting with the robot, MUL works just like SUL.

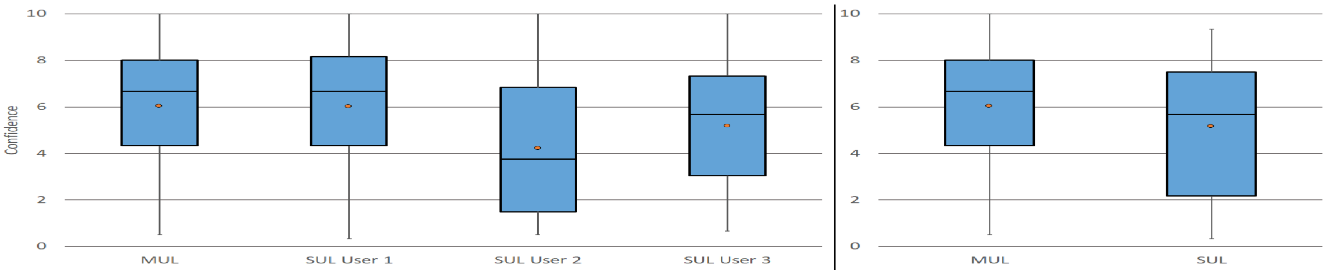

Fig. 3

In future work we intend to explore how MUL's performance is influenced by humans' points-of-view and possible solutions to balance each user's contribution given their view of the movement and workspace.

Acknowledgments

This work was supported by national funds through the Portuguese Fundação para a Ciência e a Tecnologia under project UIDB/50021/2020. Miguel Faria acknowledges the PhD grant PD/BD/143144/2019.

References

[1] F. S. Melo, A. Sardinha, D. Belo, M. Couto, M. Faria, A. Farias, H. Gambôa, C. Jesus, M. Kinarullathil, P. Lima *et al.*, "Project inside: towards autonomous semi-unstructured human–robot social interaction in autism therapy," *Artificial intelligence in medicine*, vol. 96, pp. 198–216, 2019.

[2] K. E. Kaplan, K. A. Nichols, and A. M. Okamura, "Toward human-robot collaboration in surgery: performance assessment of human and robotic agents in an inclusion segmentation task," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2016, pp. 723–729.

[3] C. A. Huijnen, M. A. Lexis, R. Jansens, and L. P. de Witte, "Mapping robots to therapy and educational objectives for children with autism spectrum disorder," *Journal of autism and developmental disorders*, vol. 46, no. 6, pp. 2100–2114, 2016.

[4] S. Chandra, P. Alves-Oliveira, S. Lemaignan, P. Sequeira, A. Paiva, and P. Dillenbourg, "Children's peer assessment and self-disclosure in the presence of an educational robot," in *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2016, pp. 539–544.

[5] E. Yadollahi, W. Johal, A. Paiva, and P. Dillenbourg, "When deictic gestures in a robot can harm child-robot collaboration," in *Proceedings of the 17th ACM Conference on Interaction Design and Children*, 2018, pp. 195–206.

[6] F. Correia, P. Alves-Oliveira, T. Ribeiro, F. S. Melo, and A. Paiva, "A social robot as a card game player," in *Thirteenth Artificial Intelligence and Interactive Digital Entertainment Conference*, 2017.

[7] G. Michalos, S. Makris, P. Tsarouchi, T. Guasch, D. Kontovrakis, and G. Chryssolouris, "Design considerations for safe human-robot collaborative workplaces," *Procedia CIrP*, vol. 37, pp. 248–253, 2015.

[8] A. Sauppé and B. Mutlu, "Design patterns for exploring and prototyping human-robot interactions," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2014, pp. 1439–1448.

[9] B. Mutlu, A. Terrell, and C.-M. Huang, "Coordination mechanisms in human-robot collaboration," in *Proceedings of the Workshop on Collaborative Manipulation, 8th ACM/IEEE International Conference on Human-Robot Interaction*. Citeseer, 2013, pp. 1–6.

[10] M. A. Goodrich and D. R. Olsen, "Seven principles of efficient human robot interaction," in *SMC'03 Conference Proceedings. 2003 IEEE International Conference on Systems, Man and Cybernetics. Conference Theme-System Security and Assurance (Cat. No. 03CH37483)*, vol. 4. IEEE, 2003, pp. 3942–3948.

[11] A. Bauer, D. Wollherr, and M. Buss, "Human–Robot Collaboration: a Survey," *International Journal of Humanoid Robotics*, vol. 05, no. 01, pp. 47–66, 2008.

[12] N. Gildert, A. G. Millard, A. Pomfret, and J. Timmis, "The need for combining implicit and explicit communication in cooperative robotic systems," *Frontiers in Robotics and AI*, vol. 5, p. 65, 2018.

[13] C. Breazeal, C. D. Kidd, A. L. Thomaz, G. Hoffman, and M. Berlin, "Effects of nonverbal communication on efficiency and robustness in human-robot teamwork," in *2005 IEEE/RSJ international conference on intelligent robots and systems*. IEEE, 2005, pp. 708–713.

[14] C. Liang, J. Proft, E. Andersen, and R. A. Knepper, "Implicit communication of actionable information in human-ai teams," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–13.

[15] M. Faria, A. Costigliola, P. Alves-Oliveira, and A. Paiva, "Follow me: Communicating intentions with a spherical robot," in *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2016, pp. 664–669.

[16] K. Baraka, A. Paiva, and M. Veloso, "Expressive lights for revealing mobile service robot state," in *Robot 2015: Second Iberian Robotics Conference*. Springer, 2016, pp. 107–119.

[17] P. Ekman and W. V. Friesen, "The repertoire of nonverbal behavior: Categories, origins, usage, and coding," *semiotica*, vol. 1, no. 1, pp. 49–98, 1969.

[18] S. Saunderson and G. Nejat, "How robots influence humans: A survey of nonverbal communication in social human–robot interaction," *International Journal of Social Robotics*, vol. 11, no. 4, pp. 575–608, 2019.

[19] A. D. Dragan, S. Bauman, J. Forlizzi, and S. S. Srinivasa, "Effects of robot motion on human-robot collaboration," in *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2015, pp. 51–58.

[20] M. Faria, R. Silva, P. Alves-Oliveira, F. S. Melo, and A. Paiva, ""me and you together" movement impact in multi-user collaboration tasks," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017, pp. 2793–2798.

[21] Y. Che, A. M. Okamura, and D. Sadigh, "Efficient and trustworthy social navigation via explicit and implicit robot–human communication," *IEEE Transactions on Robotics*, 2020.

[22] A. D. Dragan, K. C. Lee, and S. S. Srinivasa, "Legibility and predictability of robot motion," in *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2013, pp. 301–308.

[23] B. Busch, J. Grizou, M. Lopes, and F. Stulp, "Learning legible motion from human–robot interactions," *International Journal of Social Robotics*, vol. 9, no. 5, pp. 765–779, 2017.

[24] M. Kwon, S. H. Huang, and A. D. Dragan, "Expressing robot incapability," in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 2018, pp. 87–95.

[25] G. Csibra and G. Gergely, "'obsessed with goals': Functions and mechanisms of teleological interpretation of actions in humans," *Acta psychologica*, vol. 124, no. 1, pp. 60–78, 2007.

[26] L. Takayama, D. Dooley, and W. Ju, "Expressing thought: improving robot readability with animation principles," in *Proceedings of the 6th international conference on Human-robot interaction*, 2011, pp. 69–76.

[27] M. J. Gielniak, C. K. Liu, and A. L. Thomaz, "Generating human-like motion for robots," *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1275–1301, 2013.

[28] A. Dragan and S. Srinivasa, "Generating legible motion," 2013.

[29] S. Nikolaidis, A. Dragan, and S. Srinivasa, "Viewpoint-based legibility optimization," in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2016, pp. 271–278.