2024 33rd IEEE International Conference on
Robot and Human Interactive Communication (RO-MAN)
August 26-30, 2024. Pasadena, California, USA.

2024 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN) | 979-8-3503-7502-2/24/$31.00 ©2024 IEEE | DOI: 10.1109/RO-MAN60168.2024.10731388

Investigating Non-Verbal Cues in Cluttered Environments: Insights Into Legible Motion From Interpersonal Interaction

Melanie Schmidt-Wolf[1], Tyler Becker[2], Denielle Oliva[3], Monica Nicolescu[4], and David Feil-Seifer[5]

***Abstract*— In human-robot collaboration, legible intent of the robot is critical to success as it enables the human to more effectively work with and around the robot. Environments where humans and robots collaborate are widely varied and in the real world are most often cluttered. However, prior work in legible motion utilizes primarily environments which are uncluttered. Success in these environments does not necessarily guarantee success in more cluttered environments. Furthermore, the prior work has been primarily performed based on results from robot-human studies and the problem has not been studied from the prospective of what people do to express intent to each other. Therefore, this work addresses a gap in current research into legible robot arm motion in the following ways: first we perform a human-human study in order to establish the factors which humans use to express their intent through body language, and second we perform the study in a cluttered and varied environment. Through the study we showed that the primary factors which people considered are: timing, kinematic parameters, hand gestures, object proximity, etc. The results also showed that legibility is correlated with perceived safety, perceived social intelligence, the collaborator's contribution, and trust which further speaks to the importance of legible motion. Future work will utilize the pose data extracted from the study's video recordings to develop a model for legible motion.**

I. Introduction

In collaborative human-robot tasks, people and robots work together to complete a shared objective. They communicate both explicitly and implicitly to show their intention and avoid collisions. An effective communication system is even more imperative when a collaborative space is cluttered which can produce overlapping path trajectories for collaborators. This interaction can elicit repetitive patterns in how humans find optimal solutions when reaching for cluttered objects with the goal of avoiding contact with their partners.

In this research, we focus on legible motion, which is a form of implicit communication. Legible motion is intent-expressive motion that allows the human collaborator to infer the correct target object quickly and confidently and increases the understanding of a human collaborator [1]. Collaborative robots (cobots) face challenges since humans have high expectations for the cobot's performance.

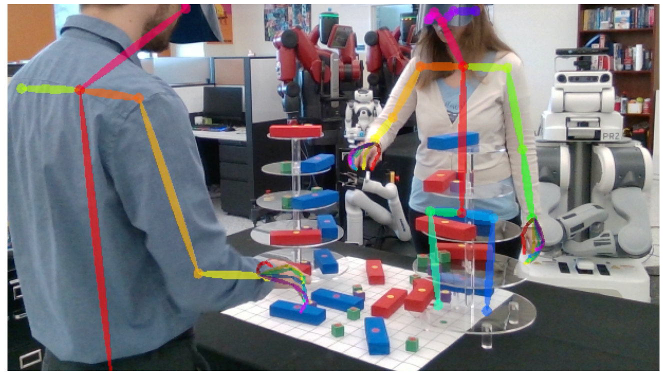


Fig. 1: Experiment setup example with a two-person team organizing a cluttered space to collect data for developing generalizable legible motion. Superimposed is a skeleton-based person computer vision detector called Openpose [2] to detect the participants' arm joints.

In cluttered environments, the effectiveness of both explicit and implicit communication decreases due to the increased number of objects and their proximity. Collision avoidance between the collaborating human and the robot is especially challenging in cluttered environments. Previous research investigating legible motion focused on uncluttered environments. Although studying legible motion in uncluttered environments for initial studies can be useful, they are not frequently encountered in human-robot collaboration scenarios.

In our previous work [3], we showed the need to explore legible motion in cluttered environments further to develop an adequate solution to generalizable legible motion planning. Therefore, since it is unknown which factors influence the legibility of robot motion in cluttered environments, in this work we focus on identifying those factors. We are exploring the behavior of humans when collaborating in cluttered spaces only using nonverbal cues, see Figure 1. Exploring the patterns of human planning and avoidance behaviors in naturally occurring arm motion trajectories when reaching for objects is essential for understanding how to model legible robot arm motion. The results from this study provide information that can be used to develop legible robot arm motion for a robot collaborator in a cluttered environment. Further, modeling arm trajectories after human practices can improve understanding in a human-

[1]Melanie Schmidt-Wolf, [2]Tyler Becker, [3]Denielle Oliva, [4]Monica Nicolescu, and [5]David Feil-Seifer are with the Department of Computer Science and Engineering, University of Nevada, Reno, 1664 N. Virginia Street, Reno, NV 89557-0171, USA *mschmidtwolf@unr.edu*, *tbecker@unr.edu*, *denielleo@unr.edu*, *monica@cse.unr.edu* and *dave@cse.unr.edu*

This work was supported by the National Science Foundation (IIS 2150394). This material is based upon work supported under the AI Research Institutes program by National Science Foundation and the Institute of Education Sciences, U.S. Department of Education through Award # 2229873 - AI Institute for Transforming Education for Children with Speech and Language Processing Challenges. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation, the Institute of Education Sciences, or the U.S. Department of Education.

robot interaction when working in a collaborative cluttered environment.

The major contributions of this paper are as follows:

- identifying factors that influence the legibility of robot motion in cluttered environments;
- determining if legible motion, the collaborator's contribution, perceived level of safety, perceived social intelligence and trust are correlated; and
- collecting a dataset of human arm motions in a collaborative task setting via video recordings.

We present insights into the design and implementation of legible motion for human-robot collaboration in cluttered environments.

The remainder of this paper is organized as follows: Section II covers the related work, Section III describes the experiment design for the study, Section IV reports on the results of the study, Section V discusses our thoughts based on our experience with the participants we found, and Section VI concludes the paper.

II. Related Work

A. Legible robot arm motion

Legible motion is defined as motion that allows the human collaborator to infer the correct target object quickly and confidently [1]. It implicitly expresses the robot's intent to be easily understood by a human collaborator. Pan et al. [4] conducted a study on the impact of the initial position of the robot arm, grasp type, and retraction speed on robot social attributes of human-to-robot handovers which can also inform the factors which are important for legible motion.

Prior research investigating legible robot arm motion has focused on developing motion planners that express the robot's intent to be more comprehensible to humans evaluated in uncluttered environments. Dragan et al. [1] introduced mathematical models to differentiate predictability and legibility, assessed in an uncluttered environment with two objects using recorded videos. They measured the legibility of a trajectory by showing the participants a video of the trajectory and asking them to stop the video as soon as they determined the goal [1]. Bied and Chetouani [5] proposed a reinforcement learning-based approach to maximize legibility metrics, evaluated in an abstracted graphical environment without a user study. Their evaluation focused on a single object in an uncluttered environment with a variable number of observers, and the observer tries to infer the goal as fast as possible [5].

Faria, et al., [6] presented a solution for multiple humans observing a motion by optimizing for the best collective value instead of a single person. Similarly, they maximized the likelihood of reaching the objective with the observed trajectory. Wallkötter, et al., [7] employed supervised learning to generate legible motion, leveraging data evaluated and labeled with established legibility measures. Their evaluation was conducted through accuracy scoring and obtaining legibility ratings from human feedback rather than user studies, utilizing an environment with seven unevenly spaced objects. Bronars, et al., [8] utilized conditional generative models guided by established legibility measures [1] to generate legible motion, evaluated against other planners with legibility measures in an uncluttered environment with two evenly spaced objects.

A survey of ten legibility frameworks was conducted by Wallkötter, et al., [9], revealing that the legibility framework proposed by Bodden, et al., was most legible [10]. Bodden, et al., [10] investigated the parameters of point position, pointing, and velocity in the context of legible motion in uncluttered environments. An underlying assumption in prior work is that those proposed solutions are extendable to more challenging, cluttered environments by studying a simplified, uncluttered environment. However, the results we obtained in our previous work [3] show that this assumption is incorrect due to complexities in cluttered environments, such as the proximity of objects to each other and the number of choices of objects.

Since it is unknown which factors influence the legibility of robot motion in cluttered environments, we therefore focus on identifying those factors in this work by studying the factors which people identify that they are focusing on when collaborating with another person.

B. Human Behavior Modeling

Since it is unknown which factors influence generalizable legible motion, we collect human-human data to be able to establish and transfer a robot motion model.

Human-human data is frequently used to derive a model that is valid for robot movements [11]–[14]. For example, human-human data has been used to model and evaluate handover behaviors in human-robot interactions [11], [12]. Human-human data can subsequently be used to derive a model for the robot through learning-based methods with supervised learning or reinforcement learning approaches [15].

To be able to create such a learning-based robot motion model in future work, it is necessary to collect human-human data and to identify factors and patterns that need to be considered when developing generalizable legible motion for human-robot collaboration.

C. Correlations to Legible Motion

In addition to identifying what parameters need to be considered when developing generalizable legible motion for human-robot collaboration, we study the correlation of legible motion to other aspects in human-robot collaboration.

Dragan et al. [16] showed a significant correlation between the perception of predictability and legibility. Lichtenthäler and Kirsch [17] concluded from a literature review that safety, comfort, surprise, efficiency, and the perceived value of a robot are correlated with legibility. Further, in a virtual human-robot path crossing task Lichtenthäler et al. [18] found that legibility is correlated with the perceived safety of the robot's behavior.

In this work, we focus on the correlation of legible motion to perceived safety, perceived social intelligence, contribution and trust. Salek Shahrezaie et al. [19] related homophily to trust. In this work, we are studying human behavioral strategies in order to learn more about what factors affect the legibility of a motion.

Perceived safety is defined as the human collaborator's perception of the level of danger and the level of comfort during the interaction [20]. *Perceived social intelligence* refers to the social intelligence observed by a human bystander. Social intelligence is defined as the ability to interact effectively in social settings to accomplish relevant objectives [21], [22] and refers to the robot's mental capability [23]. *Contribution* refers to the team members relative contribution to the task [24]. *Trust* can be defined as a trustor's belief that the trustee will act to reduce the trustor's risk when the trustor's outcomes are at risk in a given situation [25]. We aim to emphasize the crucial role of legible motion in successful collaboration by quantifying these correlations.

III. Method

In this section we describe our human-human user study design to identify legible motion patterns in cluttered spaces.

A. Research Questions

The purpose of this human-human interaction study is to evaluate naturally occurring patterns in collision avoidance when two people are working in a cluttered collaborative space. Data from this study will be used to develop legible robot arm motion for human-robot interaction in cluttered environments. As part of the user study, we are investigating the following research questions:

- **RQ1**: What parameters need to be considered when developing generalizable legible motion for human-robot collaboration?

We will investigate RQ1 through the open-ended questions about the collaboration partner's movement, see Section III-B.2.

- **RQ2**: Are legible motion, the collaborator's contribution, perceived level of safety, perceived social intelligence and trust correlated?

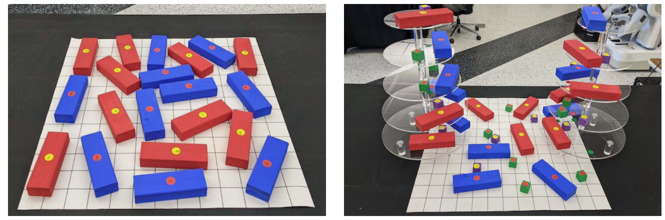
We analyzed RQ2 by calculating Spearman's ρ, which is a non-parametric correlation coefficient. A significant correlation would indicate a relationship between a motion's legibility and perceived trust, perceived social intelligence, and trust.

B. Study Design

Participants were asked to work in pairs to organize a cluttered space legibly. When collaborating with a robot, its intent should be evident to any human collaborator to prevent collisions. This study aims to find path trajectories that will be used to develop legible robot arm motion for human-robot collaboration. The participants were asked to answer questions based on the interaction. The survey concluded with demographic questions.

44 participants were recruited via flyers and in person advertisements to participate in the IRB-approved[1] study.

[1]IRBNet ID: 2133435-1



(a) 2D setup (b) 3D setup

Fig. 2: In this human-human study we used the following experiment setups: (a) A 2D setup in a cluttered space with blocks with 20 objects in total as a baseline for comparison with our previous study, see [3] (b) A 3D setup in a cluttered space with cubes and blocks with 40 objects in total. The participants were asked to organize a cluttered space with another person.

The study took place in the Robotics Research Lab at the University of Nevada, Reno (UNR). The user study had a duration of about 30 minutes.

In this human-human study we used the following experiment setups:

- A 2D setup in a cluttered space with blocks with 20 objects in total as a baseline for comparison with our previous study [3], see Figure 2a.
- A 3D setup in a cluttered space with cubes and blocks with 40 objects in total, see Figure 2b.

Each experimental setup was arranged in a table-top environment with two participants situated opposite to each other. The 3D setup included two five-tier display stands, see Figure 2b. We used two RGB-D Intel RealSense cameras to record the top and the side view of the experiment. For the organization task, the participants were provided with a number referring to the next object they were to pick up. They repeated this process for ten of the color-coded objects for each experiment setup.

1) Questionnaire: After each experiment setup, we asked the participant to complete a questionnaire via a tablet device. The questionnaire included the following items that measured responses related to RQ2 on a five-point Likert scale:

- Dragan's et al. [16] *Robot Contribution* and *Legibility* questions.
- Barchard's et al. [22] Social Information Processing Items regarding behavior and cognition from the *Perceived Social Intelligence* (PSI) Short Form Scale.
- Schaefer's "*Trust* Perception Scale-HRI" [26] 14 item sub-scale.
- Akalin's et al. [27] *Perceived safety* questionnaire.

We added randomization to the question order to reduce bias.

2) Interview: At the end of the study, we asked the participant open-ended questions to answer RQ1. These questions included:

- What would make your partner's movement easier to understand?

- What factors influence the legibility of their motion?
- What did the collaboration partner (or you) do that worked well to correctly infer the collaboration partner's (or your) target object quickly and confidently?
- How legible was the movement of (or for) your collaboration partner? Please explain.
- How did you know what the collaboration partner was about to do, and which object the collaboration partner was reaching for?
- Could you correctly infer the collaboration partner's target object quickly and confidently?

IV. Results

A. Legible Motion Patterns

In this section, we present the results for RQ1 regarding how humans create legible motion and identify the parameters to consider when developing generalizable legible motion for human-robot collaboration. Table I provides a summary of these parameters, as well as the percentage of participants who mentioned them and the most representative quotes for each parameter.

Timing: Timing was a crucial factor that affected participants' collaboration while moving towards their goal object. They reported that they hesitated, waited, and divided their roles into an active and a passive part. The active part went first, while the passive part moved when it was clear that there would be no collision with the other person. Furthermore, the participants identified hesitation and confidence as two distinct roles in their collaboration. Participants who exhibited hesitation were more cautious and were more likely to wait for the active partner to move before taking any action. In contrast, participants with greater confidence were more likely to take the initiative.

Direction: Participants also indicated the arm and hand movement direction as a crucial factor for legible arm motion. They maneuvered from different directions to prevent collision.

Avoidance Behavior: One participant reported that their collaboration partner moved around the target object instead of directly moving above it. Similarly, another participant mentioned a curving path.

Consistency: Participants emphasized the importance of movement smoothness and consistency, indicating that sudden direction or velocity changes should be avoided. Additionally, the participants reported taking a direct path and following a straight line. One participant mentioned that the movement should be smooth, similar to avoiding collision with skateboarders by following a curved path. Participants reported that both straight lines and curves are compatible with legible motion. This finding can be understood as an indication of a tendency to move in the most direct manner possible while also avoiding obstacles and maintaining a consistent level of smoothness and consistency.

Angles: One participant mentioned body inclination, which shows when the arm movement is initiated. Another participant suggested that having more information about the collaborative partner's behavioral pattern, such as the angle and side from which they initiate their movement, would make it easier to understand their movements.

Position: Participants also reported that the collaboration partner positioned the body in front of the target object to grasp the the target object easier. They also reported that moving the whole upper body half sideways to reach in and grasp objects facilitated collaborating in a legible manner. Participants also reported that the hand position influenced the legibility of the motion.

Another factor that was mentioned was the height of the hand and the arm. In the 3D setup it was easier to predict the collaborator's movement if the arm was at a similar height as the target object. Further, the collaborator would go for their own part of the board when having a low elevation and go to the other participants part of the board when having a high elevation.

Speed: An additional aspect are the position derivatives, namely speed, velocity and acceleration. One participant suggested that slower movements could improve legibility, while another participant preferred medium-fast movements. That means the movements should be fast enough to notice the arm's direction but slow enough to process the information without making the trajectory too vague. This would enable the participants to react accordingly. Moreover, the participants mentioned that they varied their reaching speed based on the collaboration partner's movement and the target object. In this regard, participants also mentioned acceleration.

Upper Body Movement: Further, participants reported that upper body movement would make the collaboration partner's movement easier to understand.

Hand Gestures: Several participants suggested hand gestures, such as pointing at the target object before reaching for it or if they moved in the same direction.

Object Proximity: Participants also mentioned that the positions of objects and the proximity to other objects influences the legibility.

Training Effect: Participants reported a training effect. Once they understood how their collaboration partner would move, avoiding collisions became easier. One participant reported that non-verbal communication was easier when one knows how their partner will move in a collaborative setting. Another participant reported that it was easy to understand how the collaboration partner was reaching for the object, especially after repeating it a few times.

This participant feedback provided valuable insights on how to develop generalizable legible robot motion. In our previous study [3], we had focused on avoiding moving towards other objects in the shared workspace to increase the trajectories' legibility. However, participants recommended smoother motions and found the sharp turns of the motion planner to be confusing. Previous work on legible motion in uncluttered environments used only some of the mentioned parameters, see Section II-A, but often implicitly. Bodden et al. [10] investigated the parameters of point position, pointing, and velocity in the context of legible motion in uncluttered environments. We did not find related work in

Parameter	Percentage	Most Representative Quotes
Timing	72%	"the collaboration partner would sometimes reach for their object quickly which while initially worrying meant that since I was slightly slower I could reach with confidence since they usually be out of my way by that time" "Hovering over an object before picking it up" "Partner would pause or wait for me to make direction clear prior to movement" "I intentionally hesitated to avoid collision"
Direction	54%	"When I saw the direction my partners arm was going to it was pretty easy to know which block or area he was reaching for"
Avoidance Behavior	22%	"If you're kind of clearly going around a thing that kind of helps with that I guess that happened a few times. Because instead of going over the top, she kind of went like like she's heading towards the thing clearly going around so okay, it's not the one closest or whatever"
Consistency	18%	"the smoothness of motion thing, because this kind of this kind of remind me a little bit of avoiding getting hit by skateboarders. Because in general, if someone's going downhill, they're following like a curving path. But I can still tell what they're doing and where they're going to go. And but if they change their direction, a lot, so I just stay put, because I can't tell what they're going to do. And I'm confident in their ability to not hit me"
Angles	16%	"Having more information about their behavioral pattern like how they make their moves, from which side or what angle" "the angling of the partners body and arm prior to reach helped with predictions. Wrist angle and article direction/angle of hand also assisted"
Position	18%	"I positioned my body so I can go directly for it." "In general height, especially if she started moving before I started moving, because if she would go low, I can kind of assume she is going for her half of the board. If she is going with high height soon, she is going for my half of the board" "For height differences, I would try to flatten my arm so more of my arm is close in height to my object"
Speed	45%	"How fast she was going because super fast, she was already there before I went. She didn't go super fast, but if she did, I wouldn't have been able to tell where she was going until she got there. Slower. It was kind of more, it was very vague. ... I think I mentioned this earlier also, but medium range speed because too slow it's a little bit like yeah, you could be going anywhere"
Upper Body Movement	22%	"If they moved their upper body more it would have been easier" "moving their whole shoulder or body to indicate height" "It was easy to understand where my parts was reaching, because they exaggerated their movement toward the object with their shoulders and turning their torso"
Hand Gestures	20%	"Or perhaps if we coordinated and pointed at our objects first" "Having a consistent gesture for types of movement"
Object Proximity	20%	"the location of our objects. If their object was not in close proximity to mine or behind mine as in closer to my side of the table it was harder to go for my object and look at where they were going" "It was difficult I found when objects were in closer range of each other"
Training Effect	43%	"I thought it was easy to understand, especially after doing it a few times, I understood how she was reaching for it, and that she was kind of waiting for me to go first. So especially after a few times easy" "Telegraphed his movements"

TABLE I: Summary of the key parameters identified in the human study as important for creating legible motion. The study found that timing, direction, avoidance behavior, consistency, angles, position, speed, upper body movement, hand gestures, object proximity, and training effect are all crucial parameters to consider. The table includes the parameters listed with the percentage of participants who mentioned them and the most representative quotes for each parameter.

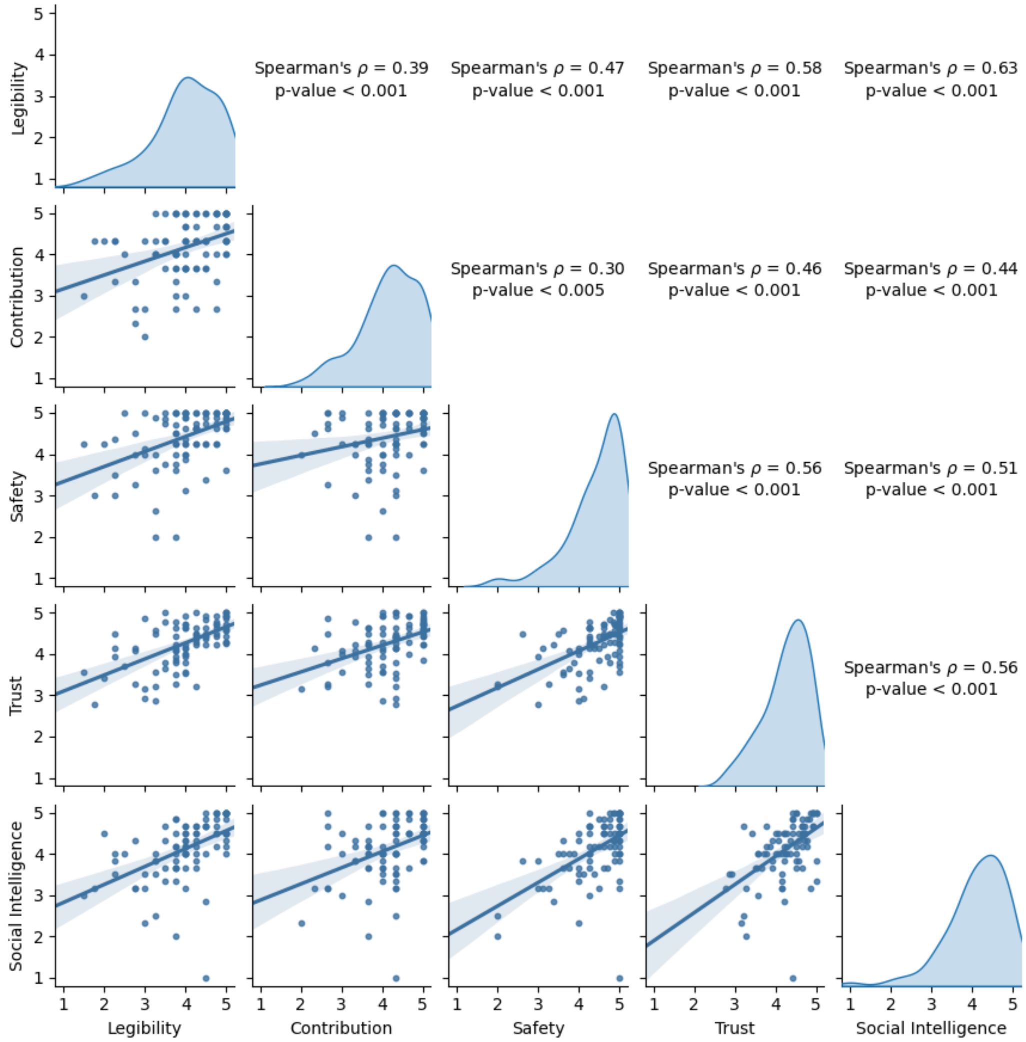


Fig. 3: The correlation matrix for legibility, contribution, safety, trust and PSI obtained from the questionnaire data. The lower triangle of the plot displays scatterplots with fitted regression lines. The diagonal plots show the marginal distribution of each variable. The upper triangle displays the Spearman's correlation coefficients and p-values. Therefore, legibility is correlated with contribution, safety, trust and PSI.

legible motion that studied upper body movement. While out of scope when creating legible motion for a robot arm, we believe that this is an interesting finding for future exploration. In summary, we have found that timing, direction, avoidance behavior, consistency, angles, position, speed, upper body movement, hand gestures, object proximity, and training effect are crucial parameters to consider when creating legible motion.

B. Correlations

In order to answer RQ2, we calculated the Spearman's correlation coefficients between legibility, contribution, safety, trust and PSI obtained from the questionnaire responses, see the correlation matrix in Figure 3.

The results of the Spearman correlation between legibility and the collaboration partner's contribution indicate a positive weak correlation (Spearman's $\rho = 0.39$, p-value < 0.001). This suggests that when arm motions are legible, they tend to facilitate higher levels of contribution from the collaboration partner. That means, the more legible the arm motions are, the more likely they are to result in improved collaboration.

There is a positive moderate correlation between legible arm motion and perceived safety (Spearman's $\rho = 0.47$, p-value < 0.001). This implies that when arm motions are legible, they are perceived as being safer by the observers. This is an important finding as it suggests that legibility can play a significant role in improving the overall safety of a collaboration task.

The results reveal a positive moderate correlation between legible arm motion and trust (Spearman's $\rho = 0.58$, p-value < 0.001). This suggests that when arm motions are legible in a collaboration task, they tend to increase the trust between the collaborators. This is a crucial element in any collaborative effort since trust enables people to work together effectively.

The results also show a positive moderate correlation between legible arm motion and perceived social intelligence (Spearman's $\rho = 0.63$, p-value < 0.001). This implies that when arm motions are legible, they are perceived as being more socially intelligent by the observers. This means that legibility can contribute to the overall impression of the collaborators as being socially intelligent, which is a valuable quality in any collaborative effort.

These data suggest that legibility is an important factor that can significantly impact the success of a collaborative effort. The positive correlations between legibility and contribution, safety, trust, and PSI indicate that legibility plays a crucial role in improving these aspects of collaboration.

V. Discussion

The results presented in this paper show multiple factors to be considered when creating models for legible motion as well as significant evidence to show the importance of legible motion with regard to trust, perceived social intelligence, etc. The most interesting things we learned from people that participated in the study were the emergent behaviors and strategies that people were able to agree upon without the need for explicit communication. The most prominent example of this was mentioned by many participants and has been grouped under the umbrella term *timing*. This factor aligns with Moon et al.'s [28] findings on hesitation behavior for nonverbal negotiation of conflicts between humans and robots. Many participants, rather than trying to avoid the collaborator's hand while grasping, simply chose to either wait for the other person to finish or to quickly grasp the object before the other person had a chance to reach it. Quite often the participants would be unofficially assigned as either the passive or active participant and then the order would be decided for the rest of the experiment. It seemed to us that the role which was chosen for each person was done based on a willingness to reach first, which we think could be related to both personality type and temperament. Furthermore, the fact that these roles did not change after they were decided implied that the participants were more willing to find some simple strategy to complete the study effectively without needing to perform the harder task, which would involve reaching at the same time as the collaborator. Reaching and causing potential collisions may have caused mild discomfort with participants and may have been the reason that this strategy was mentioned most often, along with its simplicity.

For participants who did not establish a timing with which they could avoid collisions, there were multiple methods they used to avoid collisions, which also utilize the other terms that we mentioned in the results sections. Many participants would, for example, reach towards an object at the same time, stop when they came close and then take turns. More rarely, participants were able to reach around the hand of the collaborator and would even grasp objects that were directly beneath their arm. This implies a higher level of comfort between the collaborators or a smaller perceived personal space, at least with respect to the collaborator. This could be due to a higher level of familiarity between the collaborators, which was not measured through the questionnaire. However, this could give us further insight into the strategy that was chosen as it relates to their relationship with the other person. Then, further work could be done by applying a familiarity level of the robot and the collaborator and how it applies to the required legibility of the robot from the person.

Prior work in the area of legible motion has shown the effect of factors like velocity, pointing, and position [10] on legible motion as perceived by an observer. That model considered only a few of our discovered factors, and therefore an optimization-based model that considers all of our factors is expected to be more legible. It is vague what the desired capability of the robot is when it comes to legible motion. People still struggle with this problem, as we observed through the study, and in general, the solutions that people use are reactive in nature and rely on being able to observe the partner, implicitly communicate personal boundaries, and the speed at which they will reach for the object. Therefore, if we want robots to work as well as people do in this area, then it is important to create legible models that are also reactive in the sense that they are also measuring the position of the person's hand as well. In other words, it will not be enough to generate a legible motion; the robot will also have to react to the movements of the person which can of course be unexpected. Furthermore, people adapt to the preference of their collaborator in order to establish these strategies, so it is important that the robot can do the same. This also relates to Theory of Mind [23], social intelligence and trust. This capability would require the robot to be able adapt to the movements of the person which would also require a way to classify the preferred strategy of the person as well as update it over time as familiarity increases, confidence increases, and the strategy changes.

Additionally, multiple participants stated that although they could not consistently identify the target object's exact number, they could quickly and accurately identify a cluster of objects and the intended direction toward the target. They also stated that they utilized this information to help avoid collisions along their path to their target object. The three primary factors mentioned by participants in identifying these clusters were *direction*, *speed*, and *position*. Participants were able to quickly rule out non-target objects based on the initial *direction* of the grasp. Participants could infer the intended distance from the collaborator's *speed*. Further, participants stated that in the 3D environments, the initial height of the grasp (*position*) allowed them to quickly identify the level at which the intended object was located. While prior work in *uncluttered* environments focused on the ability of the participant to identify the specific target object, this may not extend to *cluttered* environments. The responses from these participants imply that a more accurate metric for *cluttered* environments would be the identification of a target cluster, referring to a group of objects containing the target.

Legible motion is a task that people do not utilize one single strategy for and from what we saw in this study people instead agree upon some chosen general strategy by communicating with their body language. Therefore, there is no obvious way to create a motion that will always be legible to any given observer. However, if we have models that can utilize these different strategies for legible motion, then we can adapt to the collaborator. For the acceptance of robots in collaborative task robotics, this flexibility will help to increase perceived social intelligence, trust, etc.

VI. Conclusion and Future Work

In this paper, we focused on identifying parameters that influence legible robot motion in cluttered environments based on the behavior of humans in a collaboration task. Previous work showed that avoiding ambiguity with other objects is important for legible motion generation [3]. As part of the user study, we aimed to identify additional key parameters that need to be considered when developing generalizable legible motion.

The study found that timing, direction, avoidance behavior, consistency, angles, position, speed, upper body movement, hand gestures, object proximity, and training effect are important for legibility. Furthermore, the user study results show that legibility is correlated with perceived safety, perceived social intelligence, contribution and trust. Hence, we provide further evidence on the importance of legible motion.

The video data collected in this study will be used to develop a model based on human behavior, which will help develop generalizable legible robot motion that can be applied in real-world environments.

References

[1] 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.

[2] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2d pose estimation using part affinity fields," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7291–7299.

[3] M. Schmidt-Wolf, T. Becker, D. Oliva, M. Nicolescu, and D. Feil-Seifer, "Through the clutter: Exploring the impact of complex environments on the legibility of robot motion," *arXiv preprint arXiv:2406.00119*, 2024.

[4] M. K. Pan, E. A. Croft, and G. Niemeyer, "Evaluating social perception of human-to-robot handovers using the robot social attributes scale (rosas)," in *Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction*, 2018, pp. 443–451.

[5] M. Bied and M. Chetouani, "Integrating an observer in interactive reinforcement learning to learn legible trajectories," in *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 2020, pp. 760–767.

[6] M. Faria, F. S. Melo, and A. Paiva, "Understanding robots: Making robots more legible in multi-party interactions," in *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*, 2021, pp. 1031–1036.

[7] S. Wallkötter, M. Chetouani, and G. Castellano, "Slot-v: Supervised learning of observer models for legible robot motion planning in manipulation," in *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 2022, pp. 1421–1428.

[8] M. Bronars and D. Xu, "Legible robot motion from conditional generative models," 2023.

[9] S. Wallkotter, M. Chetouani, and G. Castellano, "A new approach to evaluating legibility: Comparing legibility frameworks using framework-independent robot motion trajectories," *arXiv preprint arXiv:2201.05765*, 2022.

[10] C. Bodden, D. Rakita, B. Mutlu, and M. Gleicher, "A flexible optimization-based method for synthesizing intent-expressive robot arm motion," *The International Journal of Robotics Research*, vol. 37, no. 11, pp. 1376–1394, 2018.

[11] V. Ortenzi, A. Cosgun, T. Pardi, W. P. Chan, E. Croft, and D. Kulić, "Object handovers: a review for robotics," *IEEE Transactions on Robotics*, vol. 37, no. 6, pp. 1855–1873, 2021.

[12] K. Strabala, M. K. Lee, A. Dragan, J. Forlizzi, S. S. Srinivasa, M. Cakmak, and V. Micelli, "Toward seamless human-robot handovers," *Journal of Human-Robot Interaction*, vol. 2, no. 1, pp. 112–132, 2013.

[13] D. Mukherjee, K. Gupta, L. H. Chang, and H. Najjaran, "A survey of robot learning strategies for human-robot collaboration in industrial settings," *Robotics and Computer-Integrated Manufacturing*, vol. 73, p. 102231, 2022.

[14] D. Vogt, S. Stepputtis, R. Weinhold, B. Jung, and H. B. Amor, "Learning human-robot interactions from human-human demonstrations (with applications in lego rocket assembly)," in *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*. IEEE, 2016, pp. 142–143.

[15] H. Duan, Y. Yang, D. Li, and P. Wang, "Human-robot object handover: Recent progress and future direction," *Biomimetic Intelligence and Robotics*, p. 100145, 2024.

[16] A. D. Dragan, S. Bauman, J. Forlizzi, and S. S. Srinivasa, "Effects of robot motion on human-robot collaboration," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, 2015, pp. 51–58.

[17] C. Lichtenthäler and A. Kirsch, "Legibility of robot behavior: A literature review," 2016.

[18] C. Lichtenthäler, T. Lorenzy, and A. Kirsch, "Influence of legibility on perceived safety in a virtual human-robot path crossing task," in *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 2012, pp. 676–681.

[19] R. Salek Shahrezaie, B. A. Anima, and D. Feil-Seifer, "Birds of a feather flock together: a study of status homophily in hri," in *International Conference on Social Robotics*. Springer, 2021, pp. 281–291.

[20] C. Bartneck, D. Kulić, E. Croft, and S. Zoghbi, "Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots," *International journal of social robotics*, vol. 1, pp. 71–81, 2009.

[21] M. E. Ford and M. S. Tisak, "A further search for social intelligence." *Journal of Educational Psychology*, vol. 75, no. 2, p. 196, 1983.

[22] K. A. Barchard, L. Lapping-Carr, R. S. Westfall, A. Fink-Armold, S. B. Banisetty, and D. Feil-Seifer, "Measuring the perceived social intelligence of robots," *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 9, no. 4, pp. 1–29, 2020.

[23] S. Sturgeon, A. Palmer, J. Blankenburg, and D. Feil-Seifer, "Perception of social intelligence in robots performing false-belief tasks," in *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2019, pp. 1–7.

[24] G. Hoffman, "Evaluating fluency in human–robot collaboration," *IEEE Transactions on Human-Machine Systems*, vol. 49, no. 3, pp. 209–218, 2019.

[25] A. R. Wagner and R. C. Arkin, "Recognizing situations that demand trust," in *2011 RO-MAN*. IEEE, 2011, pp. 7–14.

[26] K. E. Schaefer, "Measuring trust in human robot interactions: Development of the "trust perception scale-hri"," in *Robust intelligence and trust in autonomous systems*. Springer, 2016, pp. 191–218.

[27] N. Akalin, A. Kristoffersson, and A. Loutfi, "Evaluating the sense of safety and security in human–robot interaction with older people," *Social robots: Technological, societal and ethical aspects of human-robot interaction*, pp. 237–264, 2019.

[28] A. Moon, M. Hashmi, H. M. V. D. Loos, E. A. Croft, and A. Billard, "Design of hesitation gestures for nonverbal human-robot negotiation of conflicts," *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 10, no. 3, pp. 1–25, 2021.