# Haptic Communication in Human-Human and Human-Robot Co-Manipulation

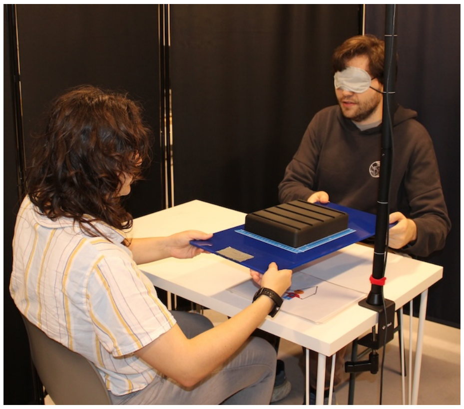
Katherine H. Allen[1], Chris Rogers, Elaine S. Short

***Abstract*— When a human dyad jointly manipulates an object, they must communicate about their intended motion plans. Some of that collaboration is achieved through the motion of the manipulated object itself, which we call "haptic communication." In this work, we captured the motion of human-human dyads moving an object together with one participant leading a motion plan about which the follower is uninformed. We then captured the same human participants manipulating the same object with a robot collaborator. By tracking the motion of the shared object using a low-cost IMU, we can directly compare human-human shared manipulation to the motion of those same participants interacting with the robot. Intra-study and post-study questionnaires provided participant feedback on the collaborations, indicating that the human-human collaborations are significantly more fluent, and analysis of the IMU data indicates that it captures objective differences in the motion profiles of the conditions. The differences in objective and subjective measures of accuracy and fluency between the human-human and human-robot trials motivate future research into improving robot assistants for physical tasks by enabling them to send and receive anthropomorphic haptic signals.**

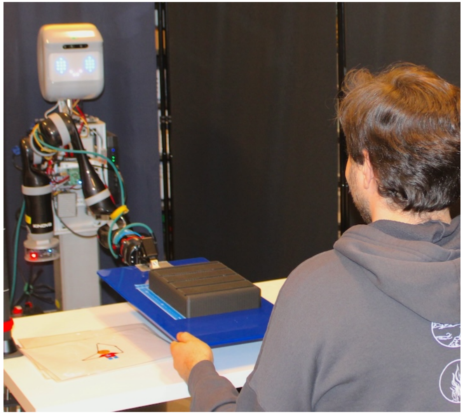
## I. INTRODUCTION

In physical collaboration tasks like carrying a couch or moving a table, haptic signals are an important channel of communication between participants to coordinate the group action. In human-human interactions, the communication and interpretation of these signals is primarily subconscious, but prior research suggests that they may enable more efficient human-robot collaboration [1]. In order for robots to participate in this haptic conversation, we need to develop a more robust understanding of how haptic communication occurs in both human-human and human-robot interaction. This knowledge can then be used to develop models for interpreting haptic intent, provide robots with comprehensible and predictable behavior, and avoid unwanted oscillations in collaborative manipulation.

In this paper, we present a study of haptic interaction, without use of visual or auditory signaling, during the collaborative manipulation of a shared object. We compare human-human and human-robot dyads to test whether there are observable differences in the subjective fluency of human-human and human-robot dyads, and whether these correlate with changes in the character of acceleration profiles of the co-manipulated objects. We additionally collect data on human perceptions of robot collaborators to identify potential co-variables in subjective fluency.



(a) Human-Human study setup



(b) Study setup for the Robot Leading Human condition.

Fig. 1: Experiment Setups

We conducted a user study with 34 participants. In the study, two agents collaborated to move a shared object, with one participant designated as the motion leader and one as a follower. Each participant acted in one these roles, and interacted with both another human participant and a mobile manipulator robot. We collected measures of subjective and objective measures of task fluency, as well as video and IMU recordings. We found that the acceleration data from an IMU mounted on the shared object changes more smoothly in human-human dyads than in human-robot dyads, with more fluent collaborations having smaller accelerations overall and smaller changes in acceleration during the task. We further find that common objective measures of collaboration fluency (e.g. task duration) do not correlate linearly with subjective fluency measures, and propose alternate measures based

---

*This work was supported by the National Science Foundation

[1] All authors are with Tufts University, Medford, MA, USA `kat.allen@tufts.edu`

on our data. This work contributes to our understanding of the differences and similarities in current human-human and human-robot haptic communication during collaborative manipulation, and provides insights that can inform future methods for autonomous haptic signaling by robots.

## II. BACKGROUND

Prior work in human-robot and human-human co-manipulation suggests the existence of some kind of difference between naïve motion planning strategies and those actually used by human collaborators. Reed et al [1] showed that a human-human dyad was able to complete a physical collaboration task more quickly than any of: the human alone, the human with a robot partner, or the human with a robot partner whom they believed to be a human partner. However, the human who mistakenly believed that they had a human partner completed the task more quickly than in the case where they knew their partner was a robot. This suggests that there is a *conversational* aspect to human-human haptic communication, beyond moving the object and waiting for the partner to catch up. "Hesitation" in the human response to a robot's "least-jerk" model lifting of a shared table-like object [2] also suggests that a follower is waiting for some sort of indication from the leader of their plan. Outside of haptic interaction, research has shown that motion patterns that are "legible", i.e., that allow the collaborator to infer an unknown goal based on the observed motion, reduced coordination and task time and improved subjective robot acceptance measures [3]. More recent work has also explored physical human-human interaction to classify collaboration styles from haptic features [4], [5], and identifying implicit communication in collaborative transportation scenarios by "via subtle signals encoded in velocities transmitted to the transported object" [6], using them as a method for probabilistically identifying the user's goals based on the joint action. Our work measures the acceleration of the co-manipulated object to identify whether these legible, conversational interactions are observable with a low-cost IMU, which could be used to train future models for robot co-manipulation control.

Research in developmental biology provides potential insight into what bio-memetic haptic communication might involve. For example, pointing and joint attention gestures develop out of reaching behaviors[7], and objects are grasped differently according to their planned use [8]. Trajectory modification is also sometimes used in joint manipulation grasping tasks to increase the probability of comprehension [9]. In our study, we explore the motion of co-manipulated objects by human dyads to identify factors relevant to these signals—if robotic agents can identify human co-manipulation signaling, they can predict the planned motion and react accordingly.

Prior research has shown that timing of inputs is critical to human comprehension and to a stable human-in-the-loop control system. In work with human-robot handoffs, a delay introduced during a handover improved compliance[10]. Other work identifies an ideal 500ms period for object handoff [11], and that control inputs for collaborative action occur at no more than 0.5 Hz[12]. In addition to meeting the human input criteria, the human-robot interaction must not induce instability into the underlying control system[13]. Our research shows that, in addition to the timing of inputs, the character of the inputs (especially their magnitude and rate of change) is important to the perceived fluency of the interaction.

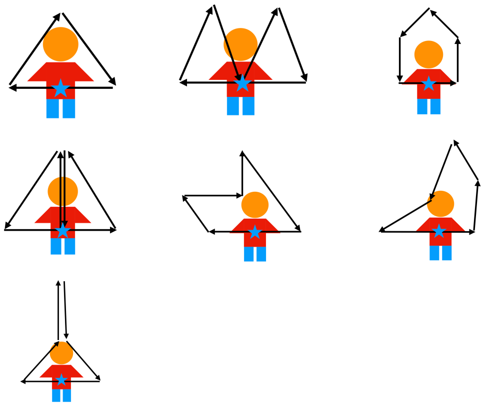


Fig. 2: Motion Plan Cards for all experiment stages. The triangle, 3-part motion plan card (top-left) is used for participants to practice the procedure before data is taken. The other 6 cards, each of which has a 6-stage motion plan, are used for data collection, with three cards in each condition (human collaborator or robot collaborator.

## III. METHODOLOGY

We identify three hypotheses concerning the fluency and characterization of co-manipulation tasks:

- **H1:** Human-robot co-manipulation is less fluent than human-human co-manipulation. Subjective collaboration fluency will be higher for human-human interactions than for human-robot interactions
- **H2**: There are "haptic communication previews" in recorded IMU acceleration data—changes in the acceleration of the co-manipulated object that occur before changes in the direction.
- **H3**: These occur more frequently in collaborations that are subjectively described as fluent.

In addition, we conducted an exploratory analysis of potential objective metrics for fluency by correlating them with the subjective fluency measurements from participant questionnaires:

- Task Completion Time
- Average acceleration magnitude and change in acceleration experienced by the shared object/IMU (in all trials)

To investigate these hypotheses, we conducted a two by two between by within study, with a between-subjects condition of role (leader vs follower) and within subjects condition of partner type (robot vs human). Each human participant was randomly assigned as either a leader or a follower, and collaborated with both a human and a robot partner. Participants were additionally randomly assigned to experience either the robot partner condition first or the human partner condition first. Table I summarizes the conditions of the study and order randomization, with participants acting as either **L**eader or **F**ollower, and with a **H**uman partner or a **R**obot Partner first.

We measured the fluency of collaboration using survey questions after each collaboration trial, to capture the participant's immediate reactions to the experience. We collected an additional subjective measurement of fluency after all three collaborations in the condition were complete.

TABLE I: The four possible sequences for a human participant as either **L**eader or **F**ollower, and with a **H**uman partner or a **R**obot Partner, and the number of participants who experienced that sequence.

| | Leader | Follower |
|---|---|---|
| Human First | 7 | 10 |
| Robot First | 10 | 7 |

### A. *Experimental Apparatus*

At the start of the session, the Leader was given 3 motion plan cards selected from a set of six similar motion plans (Figure 2). The Leader additionally was given a card with a simplified motion plan card (a triangle) for practicing the protocol. The cards used were pre-randomized as a balanced set: three were used in each of the two conditions (human partner/robot partner) such that all participants saw all the cards from the set during the experiment, in a random order.

In order to isolate the haptic communication channel from the effect of visual and aural communication, all participants were asked to remain silent during the experiment except to coordinate with the researcher. In addition, participants in the Follower role were blindfolded during the manipulation of the object to avoid accidental visual exposure to the motion plans. Participants taking the Leader role were not blindfolded in order to enable them to use visual feedback for their own motion to replicate the trajectories on the cards.

The shared object (Figure 3) used in this study is a custom-made, low-cost, instrumented board approximately 30 cm wide, 60 cm long and .5 cm thick. An Arduino RP2040 Connect running a custom web server controls the 2040's onboard Inertial Measurement Unit, giving 6 DoF data on the object's motion at its approximate center of mass. IMU data from the shared object and overhead video of the task execution provided ground truth data on the motion of the object as the participants manipulated it through the challenges.

In the human-robot interactions, we added to the experimental apparatus a social mobile manipulator robot (Figure 4). We used a social mobile manipulator robot rather than just a robot arm based on the Reed's finding [1] that users appeared to use haptic signaling only when they believed that they had a human partner: by making the robot collaborator more anthropomorphic we increase the chances of haptic communication in the human-robot collaborations. The robot took the place of one of the human participants in the collaboration, as either Leader or Follower.

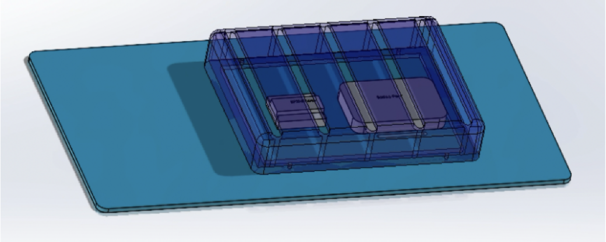


Fig. 3: A transparent CAD view of the shared object, showing the acrylic board, 3D printed electronics cage, RP2040 Connect and LiPo battery

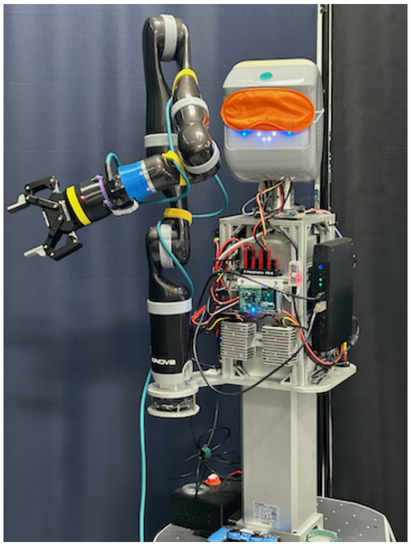


Fig. 4: The Social Mobile Manipulator Robot, blindfolded as in the Human Leading Robot condition. The blindfold covers both the LED eyes and the robot's RGBD camera to strengthen the similarity in the mind of the participant between the robot and the participant's human collaborator in the other condition.

The social mobile manipulator robot used in this study is a 1.5m tall, custom-built humanoid robot, with a Fetch robotics base, central pillar and LED expressive face, a 7DoF Kinova Jaco 2 arm, a Bota SensOne force-torque sensor with IMU, and a Robotiq two-fingered 85mm gripper. It uses AWS Polly for speech synthesis of phrases in English to greet and say goodbye to participants using pre-generated phrases.

In the Follower role, the robot used a custom, hand-tuned proportional and derivative controller over the Cartesian forces detected at the force-torque sensor. The sensor is located just above the gripper, to capture off-axis forces that are not reflected in joint torques due to system compliance. The controller used the same proportional and derivative gains ($K_P = -0.4$, $K_D = -5.0$) applied in all three cardinal directions (x,y,z) and included a deadband of ±2 Newtons for noise reduction.

In the Leader role, the robot used waypoints collected to match the trajectories from the motion planning cards and then processed into trajectories for robot arm movement using the OMPL planner for the Jaco2 arm.

### B. Study Procedure

We recruited 34 participants in 17 pairs from the university and the nearby community to collaborate in manipulating a shared object through a series of predetermined motion plans. Participants ranged in age from 18 to 54 years old, with 14 women, 18 men, and 2 non-binary people. The protocol of this study was approved by the University's Institutional Review Board (IRB) as Study 00002967.

In each condition (human-human, robot leading human, human leading robot), the agent acting as the leader was given three of the motion plan cards, each with a plan for how to move the shared object, according to one of six trajectory-following motion plans (Figure 2). To control against the follower pre-generating their own motion plans (which might agree or conflict with those of the leader in confounding ways), the follower in each condition did not have any prior knowledge of the motion plans and was blindfolded for the duration of the task.

We evaluated the collaborations both qualitatively and quantitatively, using annotated video recordings of participants and annotated motion data from the RP2040 Connect IMU on the shared object. IMU data allows the capture of subtle motion characteristics which might be obscured in a purely visual observation.

*1) Human-Human Collaborative Manipulation:* In the human-human collaborations, two humans participate in a physical collaboration task. We record their motion via the acceleration data from the RP2040 Connect IMU. Participants were asked to introduce themselves to their collaborator prior to the experiment, and to disclose their level of familiarity with their assigned collaborator, which is considered as a potential co-variable. Each participant was assigned either the "Leader" role or the "Follower" role for all their interactions.

*2) Human-Robot Collaborative Manipulation:* In the human-robot collaborations, a human and an anthropomorphic robot participate in a physical collaboration task. At the beginning of each participant's interaction with the social mobile manipulator robot of Figure 4, they were introduced to the robot and given a short description of its capabilities.

Each of the 34 participants experienced the same condition (Leader or Follower) with the robot collaborator that they experienced with their human collaborator, but with a different subset of 3 cards to ensure motion plan novelty. Half of the participants collaborated with the robot before collaborating with a human partner, and half experienced the collaboration with a human partner first, as per Table I.

Time sequences for the robot collaboration were captured both with the IMU data from the shared object and with force data and joint positions captured via ROSbag from the robot.

### C. Measures

The same subjective measures are used to evaluate the fluency of task collaboration for both the human-human and the human-robot interactions.

*1) Subjective measures of Task Fluency:* After each task, the participants were was asked to rate the fluency of the collaboration on a 1-5 scale. These comprise the "Task Rating" of fluency. In addition, after each set of three tasks with a single collaborator, each participant was asked to give an overall rating for fluency with this collaborator, the "Overall Rating" of fluency.

*2) Task Duration Video Analysis:* Task duration is captured via video annotation and cross-checked with IMU and ROSBag datastamps. In addition to task duration, the duration of each of the six legs of the motion plan (henceforth "subtask durations" ) were collected from visual observation of the video data. Video data was not captured for five interactions (four human-human and one robot-leader interaction), which were excluded from this annotation and subsequent analyses that rely on task or subtask durations.

*3) Motion Characterization:* The IMU in the RP2040 captures acceleration in the X, Y, and Z directions and rotational acceleration around the X, Y, and Z axes, at a variable data rate between 2 and 5 Hz. The IMU start time is written to the IMU data file at the beginning of data collection, and task start time is captured via video data. In the version of the IMU data collection used in the study, the clock time at the end of the IMU recording was not recorded in the IMU data file, so the video-captured task end clock time is used as the end time for both.

To account for the differences in data rate, task completion times, and wait time (the time between the start of IMU data recording and the video-annotated task start time), we scale the collected IMU data in two ways depending on the needs of our analysis:

**Scaling Method 1: IMU data scaling to clock time** The wait time is added to the task completion times (from video data), and then IMU datapoints are scaled to cover this clock time. In this visualization, the variance in total task duration is visible between trials.

**Scaling Method 2: IMU data scaling by task subsegment** To more accurately compare each task subsegments and identify common features at the subsegment level, after the data was correlated to clock times, the subtask durations (identified by video annotation) were used to identify the breakpoints in each IMU file correlating to the point where the interaction leader changed their primary motion direction to follow the new leg of the motion plan.

These were used to re-scale the data again, such that the duration of each subtask was equal to 1 unit. This allows different participants' performances of the same motion plan card to be directly compared, even if they did not perform the plans at the same speed or at consistent speeds. In this visualization, all breakpoints occur at the same time on the graph.

**Acceleration Data Calibration** In addition to the following three scaling methods for the x-axis, we have normalized

TABLE II: Overall and Per-Trial Subjective Fluency Ratings by Role and Interaction Type (**H**uman **H**uman **I**nteraction or **H**uman **R**obot **I**nteraction)

| Role, Interaction | Mean | Median | Mode |
|---|---|---|---|
| Leader, HHI | 4.294 | 5 | 5 |
| Follower, HHI | 3.706 | 4 | 4 |
| Follower, HRI | 3.313 | 3 | 4 |
| Leader, HRI | 2.627 | 3 | 2 |

the y-axis values on each plot by taking the average of the first half of the wait time data for each and subtracting that from the full y-axis set, making the y-axis not acceleration but change from baseline acceleration. This allows us to account for miscalibrations in the IMU that would make comparing trials more difficult, and re-zeros the z-axis acceleration to account for gravity.

**Mean Time Series Acceleration by Card Type** To identify changes in acceleration data that are most relevant to the motion plans associated with each card, we average the time series data by interaction type (human-human, robot-leading-human, or human-leading-robot). For this analysis, an example of which is included in Figure 10, we first scale by segment (Scaling Method 2), and then combine data into a single dataframe by timepoint. For each time, we fill forward the acceleration data from the previous true measurement, which can then be used to average across the trials despite the data rate for each trial being inconsistent due to the time-scaling process.

## IV. RESULTS

### A. *Fluency by Collaborator*

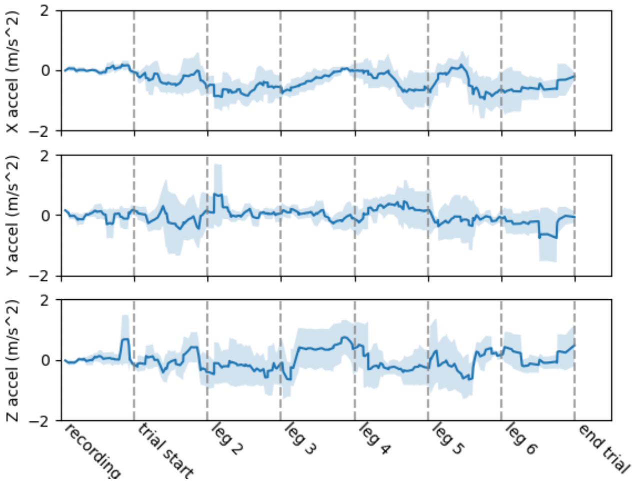
We measure subjective fluency with the overall rating given by each participant to their collaborator at the conclusion of the three trials, as well as with the rating given after each trial.

Figure 6 and Table II show the ratings given to human and robot collaborators. Human-human interactions are rated as the most fluent overall, supporting hypothesis **H1**. Leaders are more likely to rate the collaboration as fluent than followers.

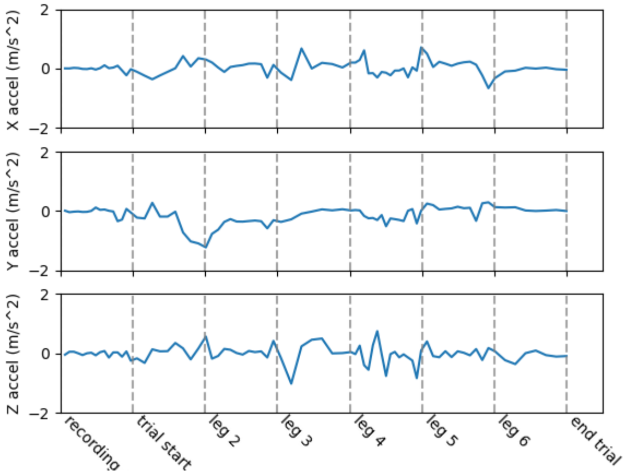
### B. *Haptic Collaboration Previews*

Of the 153 trials (17 pairs of participants each with three human-human, three human-leading-robot, and three robot-leading-human trials), 13 were excluded from our analysis due to missing/corrupted video recording (9) or because the leader did not follow the prescribed motion plan correctly (4). From the remainder, we plot the linear acceleration by segment divided by various task factors.

We observe visual differences in the rates of change of acceleration of the three types of interactions, including acceleration prior to motion breakpoints (grey vertical lines separating segments). Figure 5a has rounded "hills" and "valleys" of acceleration changes, while the robot-led, pre-programmed trajectory of 5b has sharp transitions and short bursts of acceleration that correspond with sudden initiation and cessation of robot movements. While there are some acceleration profiles that seem to preview the motion of the following segment in the human-human collaborations, they are not clearly correlated or uniformly present before each segment boundary, even in the most fluent human-human collaborations.



(a) Human-human Trials of the DT card rated Fluency 5

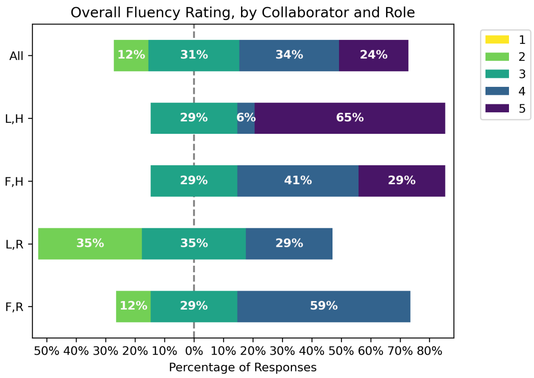


(b) Robot Leading Human Trials of the DT card rated Fluency 5

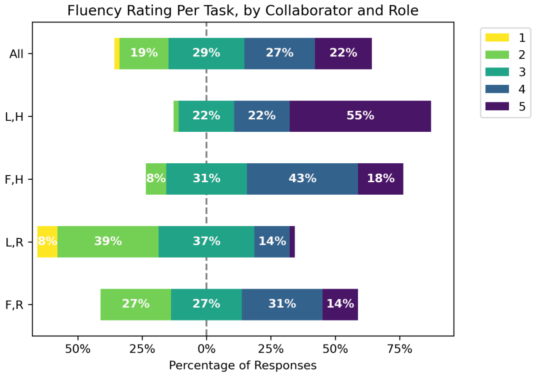
Fig. 5: Linear Acceleration Data in the x (left and right), y (toward or away from the participants), and z (up and down) directions, averaged over multiple trials

### C. *Exploratory Analysis*

*1) Task Completion Time:* Task completion time is frequently used as a proxy for collaboration fluency in human-robot interactions. However, in human-human interactions in our task, we found that task completion time did not correlate with subjective evaluations of fluency. The distribution of task times for tasks rated as more or less fluent is shown in Figure 8. Pearson's product-moment correlation between task-specific rating and task time was found to be not significant (p=.39, correlation factor of -0.086) indicating that there is no correlation between task time and fluency for our human-human tasks.



(a) Overall Fluency (3 trials)

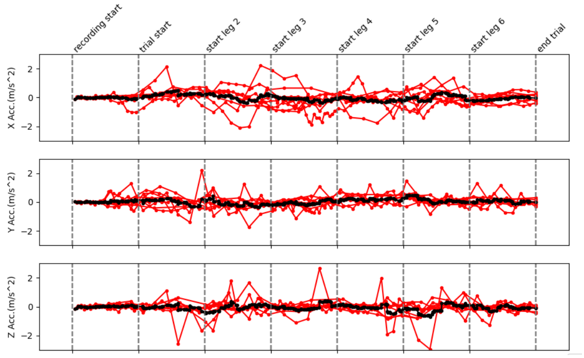


(b) Task Fluency

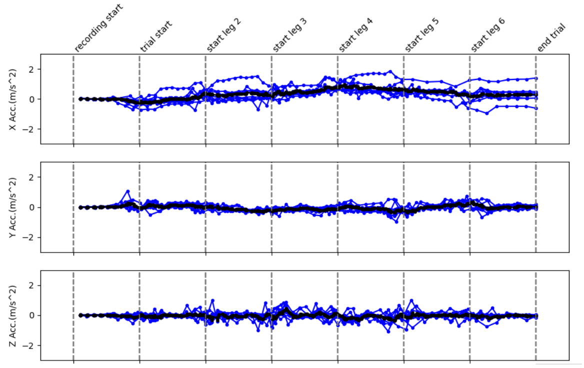
Fig. 6: Fluency (1-5) by role (leader, L or follower, F) and collaborator (human, H or robot, R)

For our human-robot tasks with a human leader[1], the correlation of task time to fluency rating is again not straightforward. Figure 8 shows the average time for human-led human-robot co-manipulations (yellow) and human-human co-manipulations (purple and blue). There is no significant correlation between time and fluency for human-human tasks (Spearman correlation -0.1, p=0.28) or human-robot tasks (Spearman correlation -0.04, p=0.66). While the lowest ratings (1) and the longest times (40 seconds) were both observed in the human leading robot condition, they did not occur in the same trial and the correlation is not statistically significant (Spearman correlation -0.24, p=0.08).

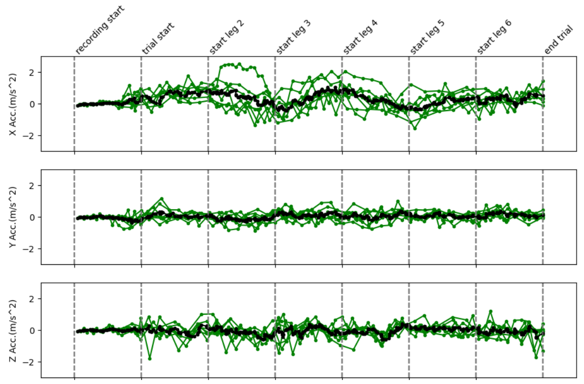
*2) Force and Jerk:* During the experiment, we observed that some participants leading the robot appeared to have more difficulty, while others were able to lead the robot effectively. The most visible feature of these difficult collaborations were oscillatory behaviors as the participant started a motion, did not observe the robot responding, and exerted more force resulting in an overshoot of the planned motion and the necessity for a corrective action. (Figure 9a shows an example of this behavior). This oscillatory behavior is a known and undesirable feature of many human-robot shared control systems.



(a) Human Human, M card, 9 groups and mean



(b) Robot Leading Human, M card, 8 groups and mean



(c) Human Leading Robot, M card, 7 groups and mean

Fig. 7: IMU data in three axes for all trials of a single card

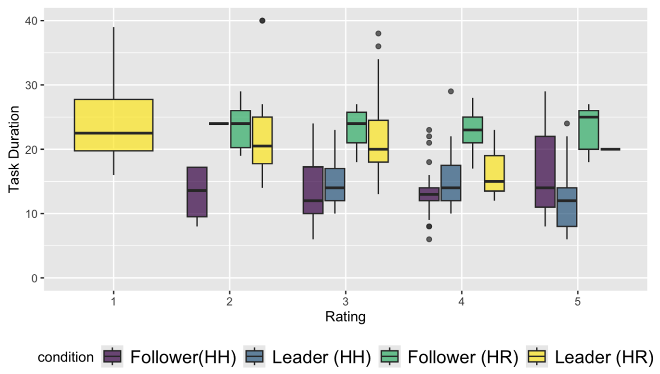
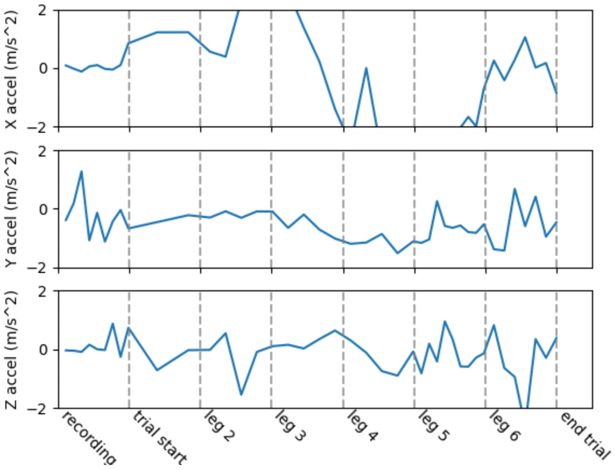


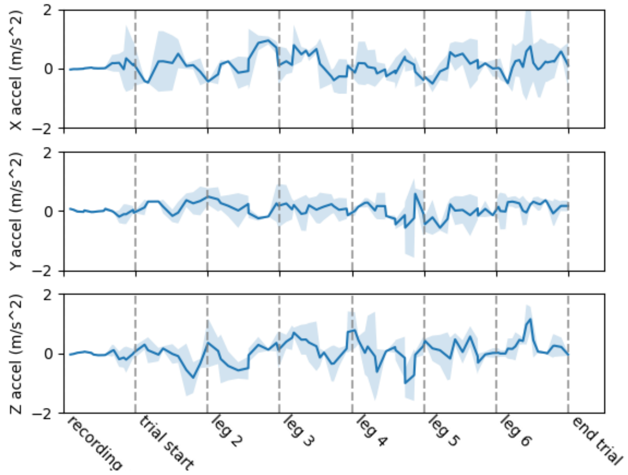
Fig. 8: Task Time by Fluency Rating

---

[1] We exclude robot-led tasks from this analysis as all the robot-led tasks were pre-planned and therefore have the same task completion time.



(a) Human Leading Robot, Fluency 1



(b) Human Leading Robot, Fluency 4

Fig. 9: IMU data from interactions grouped by subjective fluency ratings

Because our subjective fluency ratings are only at the task level (we do not have instantaneous fluency ratings at intermediate points during the task), we compare the average IMU acceleration and average IMU acceleration change per timestep. To allow for analysis of the human-human interactions as well as the human-robot interactions, we focus on just the IMU data and not the robot end-effector data.

Fluency does have a correlation with average single-trial acceleration (Figure 10) and with average *change* in acceleration between adjacent IMU datapoints (Figure 11). Average acceleration over the duration of a single trial is more than three times as high for the lowest-fluency trials as for the highest-fluency trials, with a Spearman correlation of -0.1 (p=0.039). The maximum acceleration over the trial is also correlated with fluency, with a Spearman correlation of -0.1 (p=0.037). The change in acceleration over the trial (point-by-point derivative) is correlated with fluency in all three axes with a Spearman correlation of -0.133 (p=0.006), but the *maximum* change in acceleration is not significantly correlated with fluency (Spearman correlation of -0.049, p=0.309).

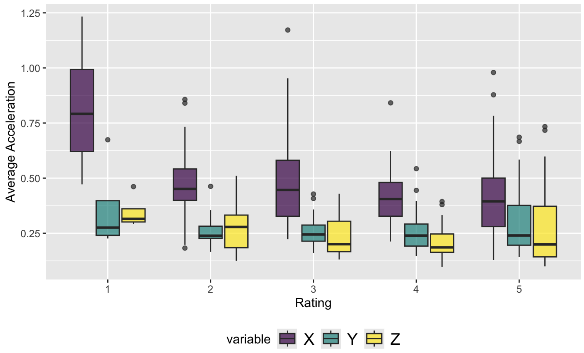


Fig. 10: Mean Acceleration by Fluency Rating

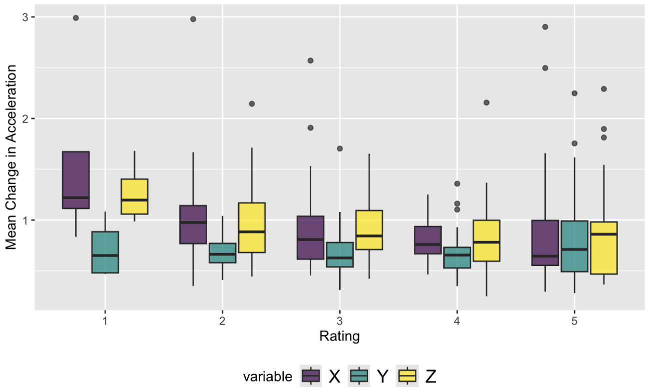


Fig. 11: Mean Change in Acceleration by Fluency Rating

## V. DISCUSSION

### A. *Haptic Communication Features*

While the IMU acceleration data does show characteristic differences between the motion of human-human, human leading robot, and robot leading human motion trials, it is not clear from this dataset that this characteristic motion includes reliably identifiable motion previews distinguishable from noise, failing to prove our hypothesis **H2**.

However, the IMU data does show observable differences between human-human and human-robot acceleration profiles, and between profiles rated more or less fluent. For example, in Figure 12, the very smooth profile of Group 7 was rated as much more fluent than the noisy profile of Group 8, suggesting that smooth changes in acceleration might be a feature of fluent co-manipulation. These smooth motion profiles do appear more frequently in collaborations identified as fluent (partly supporting our hypothesis **H3**), as we observed in our exploratory analysis of technical metrics for fluency—visually smooth acceleration changes will result in numerically low mean changes in acceleration, one of our proposed technical metrics for fluency.

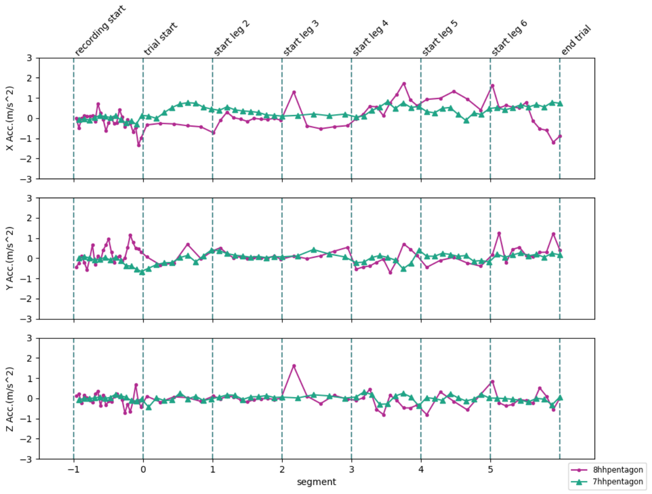


Fig. 12: Two "pentagon" motion plans, both executed by human-human dyads. Group 8 (rated as a fluency of 3 by both participants, in magenta with circle point markers) has a path with many apparently noisy features, while Group 7s path is smooth (rated a fluency of 5 by both participants, in teal with triangle point markers)

### B. *Fluency Metrics*

In this experiment, we identified no correlation between fluency and time, which is commonly used as a proxy for subjective fluency. This suggests that task duration may not be an appropriate metric for fluency in this type of task.

We propose as alternatives average acceleration or average change in acceleration (Figures 10 and 11), both of which correlate with our measures of subjective fluency.

### C. *Limitations and Future Work*

Our analysis is limited by several technical parameters of the experiment. First, while the IMU data is essential to capture human-human interactions, the limitations of the low-cost IMU in both data rate and timestamping capability limit the data on which to base conclusions about signaling in various interactions, and make correlating precise clock-time events to the IMU data challenging. A higher-fidelity sensor may be able to better distinguish between accidental and intentional motions to identify true haptic signaling, and timestamped data collection will allow for more precise correlation with video data. Additional sensors on the wrists or arms of the participants may be necessary to identify and distinguish Leader-initiated motion from Follower-initiated motion of the co-manipulated object.

## VI. CONCLUSION

Haptic communication is used intuitively in physical collaboration tasks to communicate between human agents about safety, readiness, and to coordinate the group action. For robotic participants to join physical collaborations as full participants, they require explicit training or programming to be able to interpret and identify these signals, and to generate them to coordinate with other agents.

Our work identifies a preliminary metric (IMU data) for identifying haptic communication in a physical collaboration, and demonstrates that haptic communication between collaborators allows them to more effectively and efficiently complete their tasks, with higher satisfaction among the collaborators.

## REFERENCES

[1] K. Reed and M. Peshkin, "Physical Collaboration of Human-Human and Human-Robot Teams," *Haptics, IEEE Transactions on*, vol. 1, pp. 108–120, Jul. 2008.

[2] C. Parker and E. Croft, "Design & Personalization of a Cooperative Carrying Robot Controller," *2012 IEEE International Conference on Robotics and Automation*, 2012.

[3] 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*, ser. HRI '15. New York, NY, USA: Association for Computing Machinery, Mar. 2015, pp. 51–58. [Online]. Available: http://doi.org/10.1145/2696454.2696473

[4] Z. Al-Saadi, D. Sirintuna, A. Kucukyilmaz, and C. Basdogan, "A Novel Haptic Feature Set for the Classification of Interactive Motor Behaviors in Collaborative Object Transfer," *IEEE Transactions on Haptics*, vol. 14, no. 2, pp. 384–395, Apr. 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9241412/

[5] Z. Rysbek, K. H. Oh, and M. Zefran, "Recognizing Intent in Collaborative Manipulation," in *INTERNATIONAL CONFERENCE ON MULTIMODAL INTERACTION*, Oct. 2023, pp. 498–506, arXiv:2308.09177 [cs]. [Online]. Available: http://arxiv.org/abs/2308.09177

[6] E. Yang and C. Mavrogiannis, "Implicit Communication in Human-Robot Collaborative Transport," in *Proceedings of the 2025 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '25. Melbourne, Australia: IEEE Press, Mar. 2025, pp. 23–33.

[7] J. I. M. Carpendale and A. B. Carpendale, "The Development of Pointing: From Personal Directedness to Interpersonal Direction," *Human Development*, vol. 53, no. 3, pp. 110–126, 2010, publisher: Karger Publishers. [Online]. Available: https://www.karger.com/Article/FullText/315168

[8] D. A. Rosenbaum, K. M. Chapman, M. Weigelt, D. J. Weiss, and R. van der Wel, "Cognition, action, and object manipulation," *Psychological Bulletin*, vol. 138, no. 5, pp. 924–946, 2012, place: US Publisher: American Psychological Association.

[9] G. Pezzulo, F. Donnarumma, and H. Dindo, "Human Sensorimotor Communication: A Theory of Signaling in Online Social Interactions," *PLOS ONE*, vol. 8, no. 11, p. e79876, Nov. 2013, publisher: Public Library of Science. [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0079876

[10] H. Admoni, A. Dragan, S. S. Srinivasa, and B. Scassellati, "Deliberate delays during robot-to-human handovers improve compliance with gaze communication," in *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, ser. HRI '14. New York, NY, USA: Association for Computing Machinery, Mar. 2014, pp. 49–56. [Online]. Available: https://doi.org/10.1145/2559636.2559682

[11] W. P. Chan, C. A. Parker, H. M. Van der Loos, and E. A. Croft, "Grip forces and load forces in handovers: implications for designing human-robot handover controllers," in *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, ser. HRI '12. New York, NY, USA: Association for Computing Machinery, Mar. 2012, pp. 9–16. [Online]. Available: https://doi.org/10.1145/2157689.2157692

[12] C. A. C. Parker and E. A. Croft, "Experimental investigation of human-robot cooperative carrying," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sep. 2011, pp. 3361–3366, iSSN: 2153-0866.

[13] Y. Aydin, O. Tokatli, V. Patoglu, and C. Basdogan, "A Computational Multicriteria Optimization Approach to Controller Design for Physical Human-Robot Interaction," *IEEE Transactions on Robotics*, vol. 36, no. 6, pp. 1791–1804, Dec. 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9162045