Late Breaking Work: An Investigation into Autonomy Levels Across Telerobotic Platforms With and Without Manipulators

Akash Vidyadharan, Jennifer L. Tennison, and Jenna L. Gorlewicz[1]

I. Introduction and Background

Telerobots - mobile robot platforms that allow one's presence to extend to remote settings - arose out of a need to maintain a sense of presence without being physically located in a given environment [1]. In telerobotics, there are two primary users: 1) the pilot (remote) user who operates the telerobot from a distant location and 2) the local user who interacts with the telerobot in its local environment [2]. As the use of telerobots expands into areas like elderly care, medical settings, and remote education, there is growing interest in their ability to facilitate social interactions [3]. Yet, challenges persist in achieving social presence, as telerobots often struggle with conveying non-verbal cues such as facial expressions and gestures, which are key components of rich social interaction [4]. Further, telerobot systems often lack haptic feedback which also reduces the sense of physical presence and connection across users [5]. To promote enhanced social connection in telerobotic interactions, this research examines two dimensions of inquiry: the effect of 1) control autonomy and 2) the robot form factor (namely, the presence of manipulators) on performance metrics and pilot user experience in social contexts. This study is part of a larger research program aimed at understanding and enhancing social connection in telerobotic interactions [6].

II. Hardware and Software Systems

In this work, we used two telerobotic platforms: Temi and Quori, each chosen for their unique capabilities. Temi represents a commercial robotic solution with built-in navigation and control systems, but no manipulators. In contrast, Quori is a platform designed specifically for advancing research in Human-Robot Interaction (HRI), featuring a humanoid torso, rear-projected face, and dual manipulators. The differences in these platforms, particularly the presence of manipulators, make them ideal for exploring the affordances and limitations of the varying form factors and the possible social interaction experiences enabled. Both telerobots were configured to operate under three control modes: manual, semi-autonomous (shared), and autonomous control. The graphical user interfaces (GUIs) for controls were standardized across both telerobotic platforms to ensure consistency in user experience. Navigation occurred via the same keyboard inputs for both platforms. For social interactions, both platforms also had live video and audio streaming between the pilot user and the confederate local users. The only difference in the GUI design was for Quori, where a computer vision-based virtual joystick was available to control the onboard manipulators to support the social interaction. The virtual joystick consists of an overlay of two sets of 4 directional arrow buttons corresponding to the left and right arm of Quori, where users can move their finger over the arrows to control the movement of the Quori's arms or gesture to initiate an interaction.

In manual mode, the pilot user directly controls the telerobot's movements. For example, the pilot user manually navigates the telerobot to a specific location using the keyboard and then initiates one of three social interactions with the local user: 1) a wave, 2) a point, or 3) a query. For Temi, these interactions are initiated via the live video/audio stream. For Quori, these interactions are initiated via the live video/audio stream and the use of the virtual joystick. For waving, the pilot user would move their finger from left to right over the virtual arrows to mimic a wave. For pointing, the user would move their finger on the up arrow icon on the virtual joystick until the arm reaches the desired pitch. For querying, the pilot user would move their finger over the virtual arrows to raise the arm to its maximum pitch and wave right to left once to mimic a call/query gesture. In shared control mode, human oversight is balanced with autonomous operation. For example, the telerobot autonomously navigates to one of the three interaction points and waits for the pilot user to initiate a specific interaction. For Temi and Quori, these interactions are again initiated via the live video/audio stream. For Quori, these interactions are also initiated using the virtual joystick. In the shared control mode, the pilot user presents a hand gesture in front of the laptop's webcam with one finger to initiate Quori's manipulator to point, two fingers for query, and all five fingers open to perform a handwave. In autonomous mode, the pilot user has a passive role, enabling the telerobot to navigate and perform tasks independently. In this mode, the user initiates the sequence of tasks by pressing the "Start" button on the GUI. The telerobot would autonomously travel to designated waypoints and perform the prescribed social interactions without human intervention, unless prompted by the pilot user to stop the sequence.

III. User Study Evaluation

To evaluate the different control schemes, we conducted a within subjects human user study (N=30) across the 2 telerobot platforms. The participant pool included 24 males and 6 females (ages 18-36), with educational qualifications ranging from high school graduates to PhDs. A lab environment was arranged to simulate an office space, with desks and local users positioned at three specific locations. One of the local users in the room was a confederate taking note of collisions, deviations, stalls, and time-stamping each navigation and interaction tasks for further analysis. With each robot platform, the pilot user was required to navigate to each of the three locations, stop at the location and position the telerobot toward the local user, and engage in a

[1]Research supported by the Quori Award Program and previous research conducted under NSF 1618926.

All authors are with Saint Louis University, St. Louis, MO 63103 USA in the Department of Aerospace and Mechanical Engineering; (corresponding author email: jenna.gorlewicz@slu.edu).

prescribed social interaction task. Training sessions were conducted to familiarize participants with the telerobots and their control modes. During the training session, participants practiced navigating the telerobots to specific locations and performed social interaction tasks such as waving, pointing, and querying on all three autonomy levels. Participants successfully completed navigation and interaction tasks under observation before proceeding to the actual experiment.

For the experimental trials, participants navigated the telerobot to three designated waypoints and completed one of three social interactions: greeting, pointing, and a query. Specifically, participants first navigated the telerobot to a waypoint to greet a local user by saying "Hello, how are you?" and (on Quori) performing a hand wave motion. Next, they moved to the second waypoint to reference where a bookshelf was located, which (on Quori) elicited pointing while stating "The bookshelf is over there." Finally, they navigated to a third waypoint to query local users, asking "Can I help you?" while (on Quori) gesturing a call/query by raising the arm in the air, before returning to the home position. The confederate local users responded to these interactions to simulate a natural conversation and maintain consistency across trials. Each participant completed the 3 navigation and interaction tasks under all 3 control modes for both telerobots, resulting in a total of 18 trials per participant. The entire study was IRB approved, and each session lasted approximately 90 minutes.

For each task, performance metrics were collected, including task completion time (the total time for tasks from start to finish), deviations (instances where the telerobot strayed from its intended path), stalls (moments when the telerobot had to stop and recalibrate its course), and accident counts (the number of times the telerobot collided with any object during the experiment). This data was collected to assess the impact of the different control modes and the different form factors on task execution and user experience. Qualitative feedback from the pilot user was gathered through NASA Task Load Index (TLX), System Usability Scale (SUS), and User Experience Questionnaire (UEQ), alongside post-study structured interviews. The NASA TLX assesses perceived workload, indicating how effortlessly users can interact with local users through the telerobot. The SUS is used to measure the robot's usability, while the UEQ measures the user experience of the system and the overall naturalness of the interaction. This data was used in assessing perceived workload, usability, user satisfaction, interaction quality, and social presence.

IV. Results

We conducted a within-subjects ANOVA investigating the effects of autonomy within and across the two telerobots. Initial data analysis suggests the following main takeaways:

The autonomous conditions resulted in significantly improved navigation performance with respect to task completion time, deviations, stalls, and accidents compared to semi-autonomous and manual control for both platforms. Significant differences were also observed between all autonomy conditions within each robot condition, indicating that increased autonomy leads to more efficient navigation regardless of robot form factor. Performance metrics were significantly better on all navigation and social interaction tasks in Temi compared to Quori. We hypothesize this may be due to the increased complexity of having to operate the manipulators, but further research is needed to understand this finding.

The NASA TLX ratings showed that Temi had significantly lower task load scores compared to Quori in all three control modes. This supports our current hypothesis that the complexities of Quori's form factor, particularly the manipulator control, may be important for further development. Similar trends were observed in usability ratings on the SUS and UEQ, which were significantly higher for Temi compared to Quori in all three control modes. Surprisingly, Temi was also perceived as more socially engaging and satisfying than Quori, which is an unexpected result given the social interactions possible with Quori. Analysis is ongoing to further investigate this rich corpus of data to understand the affordances and limitations of varied levels of control and capability onboard telerobots and how these impact usability and user experience.

V. Conclusion & Future Work

In conclusion, this research highlights a new line of inquiry investigating the impacts of autonomy levels and form factors of telerobots, in social interactions. Such investigations are critical to understanding if and how telerobots could move beyond "video conferencing on wheels" providing a more immersive and embodied social experience that is currently difficult to capture. Future work will focus on how these affordances and limitations generalize across pilot and local users and will investigate and identify the optimal control architecture to support enhanced capability such as controlling a manipulator to increase social presence and connection in telerobot-mediated interactions.

References

[1] Goodrich, M. T. and Schultz, A. C. (2007). Human-robot interaction: a survey. Foundations and Trends® in Human-Computer Interaction, 1(3), 203-275. https://doi.org/10.1561/1100000005

[2] Okamura, A. M., Matarić, M. J., \& Christensen, H. I. (2010). Medical and health-care robotics. IEEE Robotics \&Amp; Automation Magazine, 17(3), 26-37. https://doi.org/10.1109/mra.2010.937861

[3] Zhou, J., Ding, X., Sun, L., Yang, Y., Wang, Y., \& Liu, W. (2022). A remote monitoring and maintenance system for industrial robots. Journal of Physics: Conference Series, 2366(1), 012036. https://doi.org/10.1088/1742-6596/2366/1/012036

[4] Chen, N., Liu, X., Zhai, Y. et al. Development and validation of a robot social presence measurement dimension scale. Sci Rep 13, 2911 (2023). https://doi.org/10.1038/s41598-023-28817-4

[5] Ahumada-Newhart, V., Warschauer, M. (2023). Mobile Computing: The Role of Autonomous Features in Robot-Mediated Virtual Learning. In: Duffy, V.G., Ziefle, M., Rau, PL.P., Tseng, M.M. (eds) Human-Automation Interaction. Automation, Collaboration, & E-Services, vol 12. Springer, Cham. https://doi.org/10.1007/978-3-031-10788-7_17

[6] William N. Benson, Zachary Anderson, Evan Capelle, Maya F. Dunlap, Blake Dorris, Jenna L. Gorlewicz, Mitsuru Shimizu, and Jerry B. Weinberg. 2023. Embodied Expressive Gestures in Telerobots: A Tale of Two Users. J. Hum.-Robot Interact. 12, 2, Article 24 (June 2023), 20 pages. https://doi.org/10.1145/3570908