A Study of Bidirectionally Telepresent Tele-action During Robot-Mediated Handover

Jianqiao Li[1], Zhi Li[2], and Kris Hauser[1]

***Abstract*— The addition of manipulation capabilities to telepresence robots holds the promise of enabling remote humans to perform tele-labor, hands-on training, and collaborative manipulation, but the use of a robot as a mediator to human-human physical interaction is not yet well understood. This paper studies the impact of telepresence modalities in the context of robot-mediated object handover.A teleoperation system was developed involving a bimanual mobile manipulator with telepresence head and sensing capabilities, and a user study was conducted with n=10 pairs of subjects under a variety of audio and visual telepresence conditions. Results show that telepresence does not significantly affect objective handover fluency, but both audio and video telepresence do significantly improve user experience on subjective measures including intimacy and perceived fluency.**

I. Introduction

Telepresent tele-action (TPTA) is a recently proposed extension for the next generation of telepresence robots, in which an operator can not only communicate remotely via bidirectional audio / video through a mobile robot, but can also manipulate the environment using one or more mechanical arms (e.g., Tanaka et al [1], Indiegogo project ORIGIBOT [2]). Such devices have the potential to provide a much richer set of interaction capabilities to the remote operator than simple mobile devices. Mobility may be improved because the robot could open doors, call elevators, or move obstacles out of the way. The robot could also be used to operate equipment, demonstrate tasks, perform gestures, and interact more intimately with humans in the environment. Potential applications for TPTA devices include telecommuting, remote training, elder care, and telemedicine.

A classical notion of telepresent tele-action is *unidirectional* in which video and audio are transmitted from the robot to the remote operator such that the operator feels present in the environment. We contrast the undirectional case (UTPTA) with the bidirectional case (BTPTA) in which other humans in the robot's local environment can see and hear the remote operator via a display and speakers. Bidirectional communication is therefore likely to provide a rich channel for operators to convey information, intent, social context, instructions, and emotion, all of which would be challenging to express using the robot's body alone.

[1]Jianqiao Li and Kris Hauser are with the Electrical and Computer Engineering Department, Duke University, Durham, NC 27708, USA {jianqiao.li, kris.hauser}@duke.edu

[2]Zhi Li is with the Department of Mechanical Engineering, Worcester Polytechnic Institute, Worcester, MA 01609, USA zli11@wpi.edu

*This work is supported by NSF RAPID #IIS-1513221 and NSF CAREER #1253553 grants.

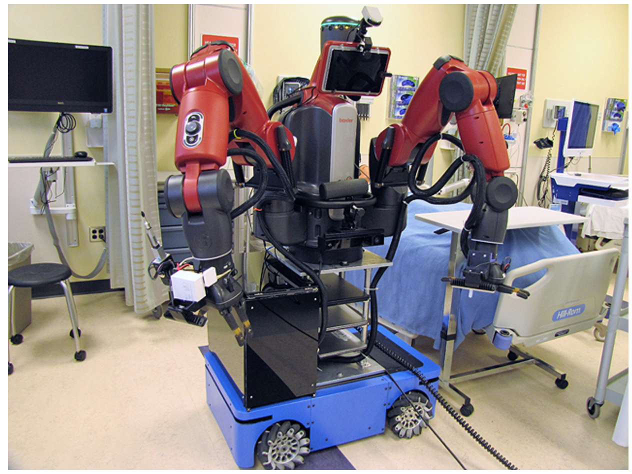


Fig. 1: The TRINA robot used in this study is a bimanual mobile manipulator developed for tele-nursing, shown here in a hospital simulation lab used for nurse training. Video telepresence is provided through the robot's face tablet.

Our study compares unidirectional and bidirectional telepresent tele-action in a cooperative robot-mediated handover task. Our work is grounded in the context of a tele-nursing platform, which is a mobile robot with two arms, hands, and a suite of visual sensors (Fig 1). Its purpose is to provide nurses and other caregivers a way to provide care to highly infectious patients while minimizing exposure to dangerous pathogens. Robot-mediated handover is a common task in such a scenario due to the regular need to deliver food, medicine, and other objects to the patient, and is an ideal task for evaluating telepresence because it requires joint physical coordination between patient and caregiver as well as coordination of handover location and timing. In human-human handovers, both the giver and receivers play active roles by observing each other and potentially rendering expectations of when and where the object will be transferred [3]. Our hypothesis is that bidirectional telepresence improves the fluency, acceptance, and user experience of robot-mediated physical interaction.

A pool of $n = 10$ subject pairs were recruited from a nursing school, trained to use the robot, and evaluated on handovers of 4 different objects under different telepresence channels (no telepresence, audio, video, and audio+video). Altogether 480 handover trials were recorded. The data suggest that bidirectional telepresence does indeed significantly improve intimacy and workload in handover, and that audio is preferrable to video. Surprisingly, telepresence did not affect objective measures of fluency, but users curiously perceived themselves as being more fluent. Overall, this study suggests that bidirectional telepresence is an effective aid to robot-mediated physical interaction in teleoperated robots.

II. Background and Related Work

The breadth of available technologies for streaming live bidirectional video and video conferencing [4]–[7] has led to the proliferation of commercially available mobile telepresence robots. A review of current platforms is given in [8]. Several platforms are also available specifically for the tele-health setting for facilitating communication between specialists and patients as well as for clinician training [9]–[11]. Existing commercial platforms provide communication and navigation capabilities, which are mutually beneficial: communication allows a robot operator to ask for help or to negotiate passage, while movement affects social interactions [12]. However, these devices have extremely limited ability to affect the robot's environment. The concept of BTPTA is extremely recent, with few published studies [1] and a single crowdsourced project with a BTPTA system with 3+1-DOF gripper [2].

Telepresent teleaction has been used mainly in the context of UTPTA, which intends to improve the experience for the operator [13], [14]. Indeed, a major challenge in developing telerobotic systems is usability. Medical robotic systems are generally difficult to control for complex manipulation tasks, even with training and high-fidelity robots and input devices [15]. However, unidirectional telepresence neglects the experience of humans physically interacting in the same space as the robot, which is important for telemedicine, telecommuting, and remote training.

Handover is an active topic of study in human-robot interaction [3], [16], [17]. Fluent handover is achieved via an intricate spatio-temporal coordination between giver and receiver. Verbal, eye gaze, gestural, postural, and tactile cues can signal intent and negotiate transitions between the many phases of handover, including readiness to give, readiness to receive, etc. Several authors have studied performance and human preferences for human-robot handover [18]–[20]. There is experimental evidence that temporal accuracy is more important than spatial accuracy to provide a satisfying user experience [21]. However, this paper is to our knowledge the first paper that studies handover in the context of robot-mediated human-human interaction.

III. Hypotheses

Our experiment is designed to test these hypotheses:

- **H1.** Bidirectional telepresence improves objective task performance.
- **H2.** Bidirectional telepresence reduces subjective workload for both operator and partner.
- **H3.** Bidirectional telepresence improves user experience, in terms of subjective rating of task performance and intimacy between the partner and the robot/operator.
- **H4.** A bidirectional audio channel is more important than video in improving user experience and task performance.

These hypotheses originated with pilot testing on the robot. H1 is suspected because bidirectional communication should assist in the intricate spatial and temporal coordination involved in handover. H2 is suspected to hold because richer communication is expected to build joint understanding, assist in social bonding, and reduce frustration. Similarly in H3, participants are unlikely to feel comfortable without telepresence because a robot is far less familiar than a human voice and face. As for H4, we felt that video may be less important for several reasons. First, the participants need to shift gaze from the robot hand to the telepresence screen to see each other. Second, although body gesture and facial expression can be used as communication cues, they are often less direct and more ambiguous than verbal communication.

IV. Experimental Platform

The experimental platform consists of a bimanual mobile manipulator with a rich sensor suite, bidirectional telepresence capabilities, as well as an operator console and devices to enable teleoperation of the robot's base and arms. Here we describe the system components and capabilities in detail.

A. Robot Hardware

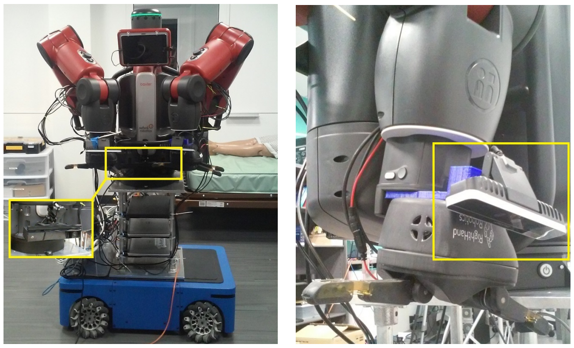
The TRINA robot [22] is a mobile manipulator system standing 175 cm tall, shown in Fig. 2. The body consists of a Rethink Robotics Baxter Research Robot mounted on a HStar Technologies AMP-I omnidirectional mobile base. The Baxter has two 7-DOF arms, and upon each arm is mounted a Righthand Robotics ReFlex three-fingered gripper.

A few items should be noted about the system in regards to robot-mediated interaction. Although Baxter's arms are 7-DOF like a human, they are non-anthropomorphic in two respects. First, the elbow points "up" rather than "down" in its natural pose, and joint limits prevent it from reaching human-like elbow-down postures. Second, the wrist is approximately 27 cm away from the palm of the gripper, so that wrist flexion produces a large movement at the gripper. Likewise, rotation about the palm or opening of the fingers causes a large movement of the wrist and elbow. The grippers are also compliant and under-actuated, with only one motor per phalanx, and can be operated only in pinch and power grasp modes. Hence they do not match the dexterity and accuracy of a human hand. As a result, apparently simple manipulation tasks can be challenging to operate.

The operator is provided with sensory feedback via a Microsoft Kinect 2 on the robot's chest (see Fig. 2a) and two Intel RealSense F200 3D cameras (see Fig. 2b) attached to the robot's wrists. Each contributes to video streams and 3D visualizations shown on the operator's console.

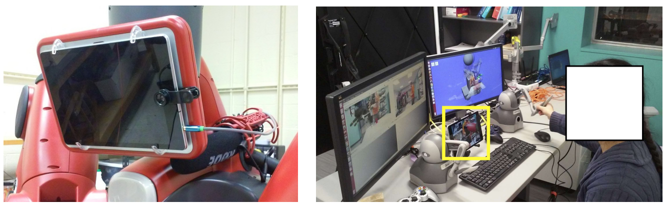
B. Operator Console

The operator has the option to control the robot via several input devices, but in this study, we only use a pair of 6-DOF haptic devices (Geomagic Touch). Each device controls the motions of one robot hand, including (1) the position & orientation of the robot hands, and (2) the opening and closing of the compliant gripper. The three fingers compliant



(a) (b)

Fig. 2: (a) The experimental platform is a mobile manipulator that consists of an off-the-shelf humanoid robot and an omnidirectional mobile base. (b) Compliant grippers are attached as the robot's hands for manipulation.



(a) User's side (b) Operator's side

Fig. 3: The bidirectional telepresence system. Two tablets for audio and video communication are set (a) on the robot's face and (b) at operator's workstation (marked with yellow rectangle), respectively.

hand can be formed into power grasping or precise grasping. The mobile base is not controlled.

Bidirectional telepresence is achieved using two Google Nexus 9 tablets, one attached to the robot's face screen, and the other with the operator at the operator console. The robot-mounted tablet has a wide angle lens and is connected to a directional microphone to reduce noise interference from the robot's motors and fans.

The operator console supports the control of the robot and communication modes (Fig. 3). The control GUI displays controller states, camera streams, 3D maps, and robot status in real-time. The robot's chest camera provides a front view of the subject, while a stand-alone camera on a tripod provides a side view. Video streams and the 3D displays can be expanded and shrunk as desired. Although the system supports adjustable autonomy, the operator was deliberately restricted to using direct teleoperation for this experiment.

C. Experimental Setup

Novice subjects are recruited in pairs. For each pair, we randomly assigned one subject to be the **operator** (O) that controls the robot remotely from a workstation, and the other to be the **partner** (P) who is present in the robot's environment. Fig. 3 shows the setup from O's side. O sits in a separated space at the console, and chooses either the left or the right haptic device according to personal preference. O is blocked from direct line of sight of P and the robot with a screen. Audio communication, if provided, is conducted over audio headsets. Fig. 4 shows the setup at P's side. Before a task starts, P stands facing the robot across a table. On the table, a green line and a red line spaced 25 cm apart denote the robot's side and P's side, respectively. P is prevented from reaching his/her hand over the red line, while the robot's hand should always start its motion behind the green line.

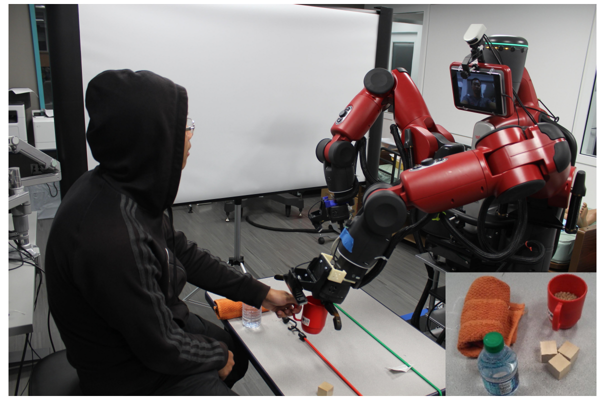


Fig. 4: Experiment setup — Partner's side

A successful handover trial consists of the following steps (Fig. 5): 1) P picks up an object from the table, 2) O controls the robot from a starting area to receive the object from P, 3) O retracts the hand back to the starting area, 4) O gives the object back to P, and 5) O retreats to the starting area and P places the object on the table. The subjects are asked to try to finish each trial as fast as possible. A trial is be considered to be failed if the object is dropped either by the robot or P.

By enforcing a separation between the robot starting area and P's area, we require O to command substantial arm movements to the robot in both receiving and giving. It also limits the amount of assistance that P provides to the robot during receiving, and prevents O from simply dropping the object into P's hand during giving.

D. Experimental Conditions

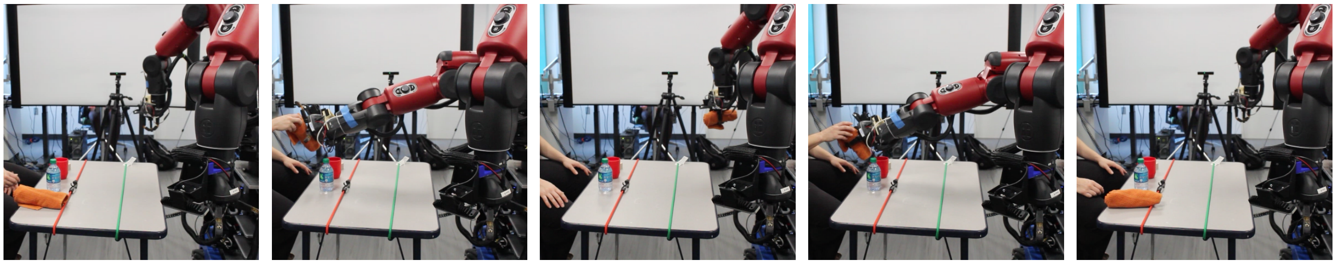
In each session, the subjects were given the following levels of bidirectional telepresence, tested in random order:

- **No Communication(NC)**- Neither audio nor video communication is available through the telepresence system.
- **Audio (A)**-Only audio communication is available.
- **Video (V)**-Only video communication is available.
- **Audio and Video (AV)**-Both audio and video communication is available.

In all conditions, the operator can see the partner through the video feeds on the operator console, but in NC and A the partner is asked not to attempt to use gestures to communicate with the operator. Altogether, 3 trials $\times$ 4 objects $\times$ 4 conditions $\times$ 10 experiments = 480 trials were conducted in our experiment.

E. Task Description

Each task involves repeated handover of a single object. Four objects were used: towel, water bottle, plastic cup and small wooden block (see Fig. 4, inset). These were chosen to vary in weight, necessary orientation precision and position



(a) Idle (b) Receive (c) Retract (d) Give back (e) Idle

Fig. 5: Phases in a successful handover trial.

precision, which affects control difficulty and the amount of communication necessary to produce a successful handover. For example, the bottle is heavy while the towel is light. The cup filled with lentils requires orientation precision to avoid spilling, while the small wooden block needs precise positioning. O also must switch between the two grasp modes (power-mode and precision mode) for different objects.

The objects are placed on the table on P's side when the experiment starts. In each session, tasks were presented in the fixed order of 1) towel, 2) bottle, 3) cup, and 4) block.

F. Intake Survey

After obtaining informed consent, experimenters administered a survey to each subject to collect demographic information (age, gender, handedness) and answers to the following questions:

1) How many years of nursing or health care professional experience do you have?
2) How many hours a week, on average, do you spend on the computer?
3) I am experienced with (a) remote-controlled toys, (b) first-person perspective video games (e.g. Half-Life, Call of Duty), (c) real time strategy games (e.g. StarCraft, Warcraft)

Questions 3.a–c are rated on a 7-point Likert scale ranging from "Strongly disagree" (1) to "Strongly agree" (7).

G. Training Phase

In the training phase, experimenters first explained the experimental procedures to the subjects as follows:

- Introducing the robot and the objects.
- Explaining the process of a handover trial and conditions for success.
- Rules were described for each experimental condition. Specifically, in NC and V, both subjects were asked not to communicate verbally, and in NC and A, P was asked not to communicate via gestures.

Subject pairs then underwent a 15-min training session followed by a competency test. During the training phase, one experimenter taught O how to control the robot, while another instructed P about the rules for completing a handover trial. The competency test asked the pair to perform tasks of handing over a towel and a wood block within 1.5 min and 2 min, respectively. The subject pair could have two attempts for each object. Those that failed the test could practice 10 more minutes before taking the same test again. Failing the second test would exclude the pair from participating in the rest of the study.

H. Post-Session Surveys

Subjects take a break after each session and evaluate their experience under the telepresence condition in last session. We use the NASA-Task Load Index (NASA-TLX) to measure the workload for each task (object) with the following questions:

- **Mental Demand**: How mentally demanding was the task?
- **Physical Demand**: How physically demanding was the task?
- **Temporal Demand**: To what extent did you feel hurried during the task?
- **Performance**: How successful do you consider yourself in accomplishing the task?
- **Effort**: How much effort did you personally have to put forth to accomplish your level of performance?
- **Frustration**: To what extent did you feel insecure, discouraged, irritated, stressed, and annoyed during the task?

Ratings were reported on a five-point Likert scale, anchored by "Not at all"(1) to "Extremely"(5).

A standard pair-wise questionnaire (6 subscales forming 15 pairs) was used to obtain a weight for each TLX subscale at the end of the experiment, following [23]. For each pair, subjects choose the subscale that contributes more to overall workload. Each subscale is assigned a weight equal to the number of times it is chosen, and the overall workload score is the sum of weighted subscale scores divided by 15.

After each session, subjects were asked the following user experience questions about the communication channel:

- This level of communication helped you perform the task more precisely
- This level of communication helped you perform the task more quickly

After the session under condition AV, the subjects would respond to the following additional question:

- How much attention did you pay to the audio channel?
- How much attention did you pay to the video channel?

The answers to these questions were on a five-point Likert scale, anchored by "Not at all" (1) to "A great deal" (5).

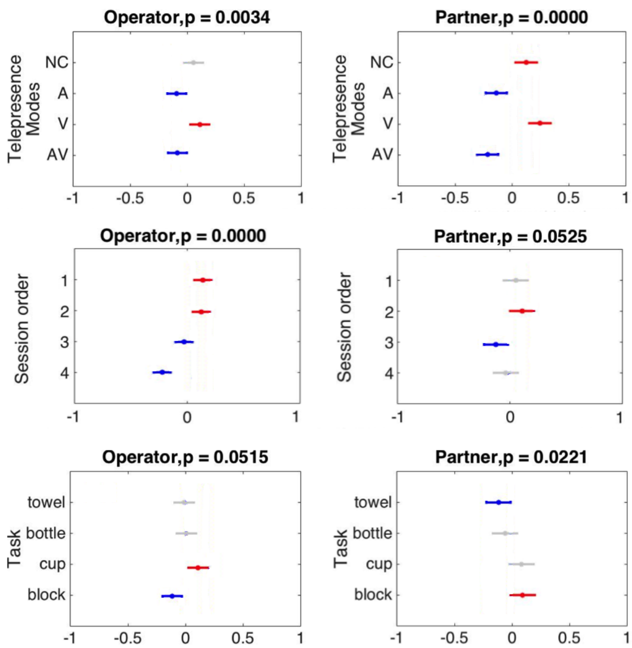


Fig. 6: Normalized NASA TLX workload ANOVA results

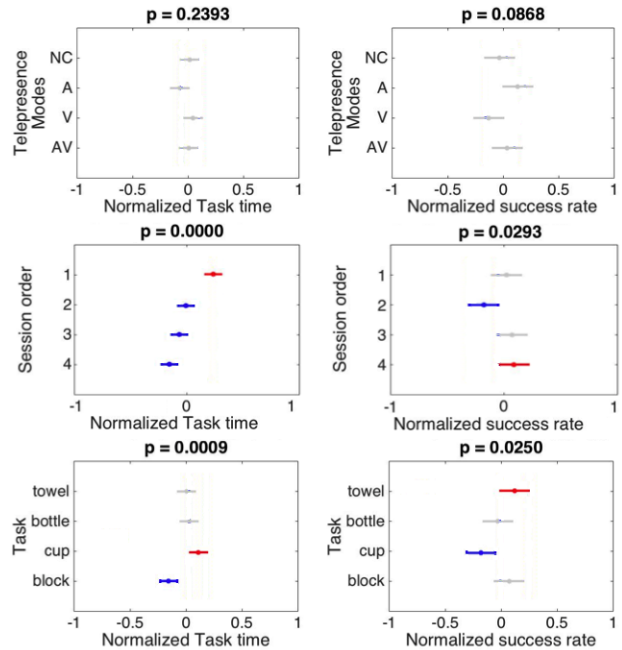


Fig. 7: Objective performance measure ANOVA results

The partner was asked two more questions about their intimacy rating for the robot and the operator:

- You felt closely engaged in the interaction with the robot under this experimental condition
- You felt more intimate with the operator under this experimental condition

Ratings were given on a seven-point Likert scale anchored by “Strongly disagree” (1) to “Strongly agree” (7).

I. Objective Measures

We videotaped the operator and the partner during the experimental phase. To measure objective fluency, we extracted the time spent for each trial and counted success/failure trails from the video. We also recorded the full body motion and audio of the participants, and counted the number and types of communications performed (verbalization, hand motion, body gesture, and facial expression).

V. Analysis

A. Data Pre-processing

We found that the measures from same subject are relatively consistent, but vary largely between subjects. For example, some subjects may rate 0 for the easiest task while others may still rate it some small number. Since we are trying to compare the change in the measures under different conditions, normalization is used to reduce some of the baseline variability. Specifically, we divided a subject’s score by the mean of all that subject’s scores for the same questions and subtracted 1.

B. Order, Task, and Experimental Condition

Many factors affect objective and subjective performance. For example, certain objects are harder than others, and later sessions are easier than earlier ones as the participants begins to learn the procedure. We use one-way ANOVA with multi-comparison to analyze how the following factors normalized measures:

- **Session Order.** The order in which sessions took place.
- **Task Type.** The object handed-over.
- **Experimental Condition.** The telepresence communication channel used.

C. Audio/Video Coding

Coding of audio and video channels was conducted under the following guidelines. One audio communication was counted if the participant uttered a sentence with clear meaning, such as:

- Question & Answer: “Did you get it?”, “Yes/No”
- Instruction: “Move it closer,” “Please open the hand.”
- Comment: “Good job,” “This is fun.”

One video communication was counted if the participant used a clearly observable posture/gesture, such as:

- Hand gesture: Open/close hand. Thumb up/down.
- Head gesture: Nod/shake head.

VI. Results

A. Subject Population

We recruited twenty healthy subjects through a partnership with the nursing school at our university, using emails and word-of-mouth. The subject pool consists of 18 females and 2 males, age range 22–61, mean age 28.85 with σ = 9.00. 18 are right-handed, 1 left-handed, and 1 is ambidextrous. All subjects are students of the nursing school, 15 with less than 1 year of experience, 3 with 1–3 years, and 2 with 3+ years. Number of computer hours and video game experience showed a roughly even distribution. We used an online poll for scheduling and pairing our subjects. All subject pairs passed the competency task and were retained for the study.

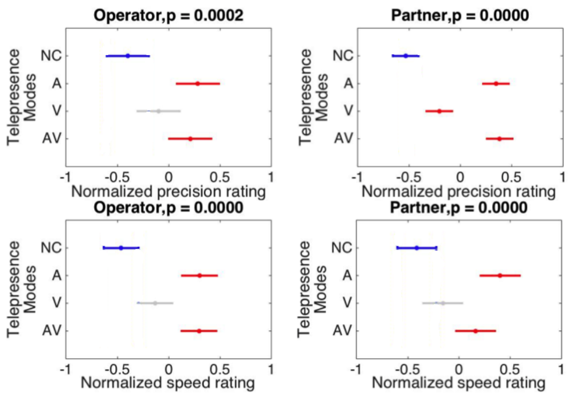


Fig. 8: User experience ANOVA results

B. Task Workload

Fig. 6 shows results of ANOVA with multi-comparison for normalized task workload measures. In this and the remaining ANOVA figures, groups shown in blue have significantly different means with groups shown in red at the α=0.05 level. Horizontal bars indicate comparison intervals for statistically significant differences in population means.

Session order and telepresence mode had significant effects on both operator and partner's workload, while task type only had significant effects on the partner. The lower workload in later sessions indicates that learning occurs over time. The experimental condition had a surprising effect. For both operator and partner, **the presence of an audio channel reduced workload significantly**. However, the video-only condition slightly (but insignificantly) increased the workload from NC. A potential explanation is that communication is more difficult using only expressions and gestures.

C. Objective Fluency Measures

ANOVA results for normalized task performance are shown in Fig. 7. Session order and task type had significant effects. The direction and magnitude of the task type effects are consistent with the task difficulty, and the order-induced change are due to familiarity. However, **experimental condition did not significantly affect any objective performance measure**.

D. User Experience

ANOVA results for normalized *perceived* task performance are shown in Fig. 8. Both operator and partner felt they were faster and more accurate with telepresence provided, with significant effects when audio was present. (This is surprising in light of the insignificant changes in objective performance) Effects of telepresence condition on the partner's overall intimacy rating are shown in Fig. 9. Telepresence significantly improved intimacy with the robot as well as the operator. In summary, **bidirectional telepresence generally improves user experience, and although audio has a larger effect, the video channel does contribute to greater intimacy**.

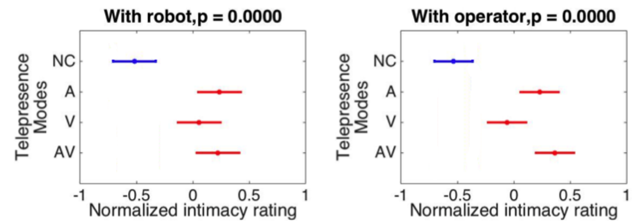


Fig. 9: Partner intimacy rating ANOVA results

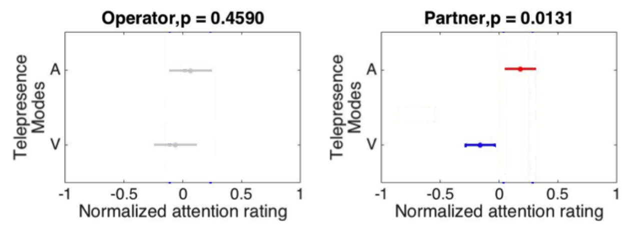


Fig. 10: Attention paid to each channel when both Audio and Video were provided

E. Audio vs Video Communication

The ANOVA results shown in Fig. 10 indicate that partners pay significantly more attention to audio than video channels. No statistically significant differences were observed in the operator. (We note that this may be due ambiguous wording, since we did not specify whether "Video" referred to the body camera streams or the telepresence screen).

Descriptive statistics about the communication type and frequency are shown in Fig. 11 and Fig. 12. In general, we found huge variability in communication strategies from pair to pair. The average number of communication instances were similar under different conditions, but participants preferred to use audio over video in the AV condition.

Fig. 12 shows the distribution of communications made by subjects in each role. Partners made most of the communications, particularly in the video category. This is likely because the operator focused more on controlling the robot.

VII. Discussion

Overall, these results supports some of our hypotheses:

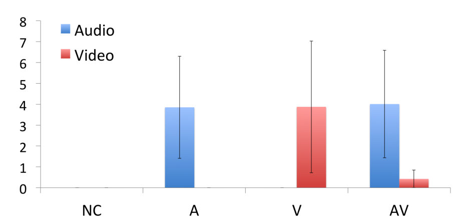


Fig. 11: Average number of communications per trial (+/- 1 S.D.)

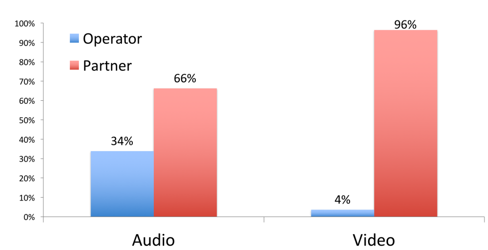


Fig. 12: Percentage of communications made by operator/partner

- H1: Unsupported. Objective performance was not significantly affected by the communication channel.
- H2: Mixed. Subjects reported a lower workload when an audio channel was provided. However the workload was (insignificantly) higher when only video was provided.
- H 3: Supported. All the measures for user experience were improved with telepresence.
- H4: Supported. Both audio and video channel improved performance and experience, yet the audio only condition always outperform video only condition by a large margin. Moreover, subjects preferred to use audio channel over video channel when both were provided.

The increased reliance on audio over visual communication is not surprising for this task, since it involves visual attention to spatial coordination. However we were surprised at the dramatic improvements that telepresence had on subjective measures of user experience and perceived fluency, in the face of negligible objective effects on fluency.

VIII. Conclusion

This paper studied how bidirectional telepresence channels affect the performance and subjective experience of robot-mediated object handover. A user study was conducted with a population of nursing students on a tele-nursing robot. Results suggest that additional communication channels improve intimacy in the face of a robot embodiment, which is particularly important in the case of robots interacting with sensitive populations, e.g., patients in a hospital. Workload was also reduced when audio communication was included.

Handover fluency was not significantly affected by any telepresence channel, which suggests that communication was not the performance-limiting factor in our task. Curiously, more telepresence channels increased *perceived* fluency. In future work, we are interested in improving the fluency of handover and other tele-manipulations by changing the control interface, either by making the operator's console more immersive, or by using semi-autonomous behaviors in a supervisory control mode. For example, tactile sensation is an important aspect of human manipulation that is not available on our system, and could hypothetically supplement audio/visual communication for improved performance.

A limitation of this work is that it considers a relatively impersonal and relaxed context of handover. In future work, we are interested in studying BTPTA in more communication-intensive or time-sensitive scenarios, like patient treatment, in which the telepresence channel may have a more substantial effect on objective performance.

References

[1] F. Tanaka, T. Takahashi, S. Matsuzoe, N. Tazawa, and M. Morita, "Telepresence robot helps children in communicating with teachers who speak a different language," in *Proc. ACM/IEEE Int. Conf Human-robot Interaction*, 2014, pp. 399–406.

[2] Indiegogo, "Origibot: Remote telepresence robot w gripper," accessed May 2016. [Online]. Available: https://www.indiegogo.com/projects/origibot-remote-telepresence-robot-w-gripper/

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

[4] Apple, "Apple introduces ichat av and isight. press release, june 2003." accessed Jan. 2016. [Online]. Available: http://www.apple.com/pr/library/2003/jun/23ichat.html

[5] G. Venolia, J. Tang, R. Cervantes, S. Bly, G. Robertson, B. Lee, and K. Inkpen, "Embodied social proxy: mediating interpersonal connection in hub-and-satellite teams," in *Proc. SIGCHI Conf. Human Factors in Computing Systems*. ACM, 2010, pp. 1049–1058.

[6] S. Lachapelle., "Talk face-to-face right from within gmail. official google blog, nov. 2008." accessed Jan. 2016. [Online]. Available: http://googleblog.blogspot.com/2008/11/talk-face-to-face-right-from-within.html

[7] Skype, "Skype introduces video calling for macintosh users. press release, sept. 2006." [Online]. Available: http://www.skype.com/en/about/

[8] A. Kristoffersson, S. Coradeschi, and A. Loutfi, "A review of mobile robotic telepresence," *Advances in Human-Computer Interaction*, vol. 2013, p. 3, 2013.

[9] G. Baltus, D. Fox, F. Gemperle, J. Goetz, T. Hirsch, D. Magaritis, M. Montemerlo, J. Pineau, N. Roy, J. Schulte, *et al.*, "Towards personal service robots for the elderly," in *Proc. Workshop on Interactive Robotics and Entertainment (WIRE-2000)*, 2000.

[10] D. Sampsel, P. Vermeersch, and C. R. Doarn, "Utility and effectiveness of a remote telepresence robotic system in nursing education in a simulated care environment," *Telemedicine and e-Health*, vol. 20, no. 11, pp. 1015–1020, 2014.

[11] P. M. Vespa, C. Miller, X. Hu, V. Nenov, F. Buxey, and N. A. Martin, "Intensive care unit robotic telepresence facilitates rapid physician response to unstable patients and decreased cost in neurointensive care," *Surgical Neurology*, vol. 67, no. 4, pp. 331–337, 2007.

[12] H. Nakanishi, Y. Murakami, D. Nogami, and H. Ishiguro, "Minimum movement matters: Impact of robot-mounted cameras on social telepresence," in *Proc. ACM Conf. Computer Supported Cooperative Work*, ser. CSCW '08. New York, NY, USA: ACM, 2008, pp. 303–312. [Online]. Available: http://doi.acm.org/10.1145/1460563.1460614

[13] M. Buss, A. Peer, T. Schauß, N. Stefanov, U. Unterhinninghofen, S. Behrendt, J. Leupold, M. Durkovic, and M. Sarkis, "Development of a multi-modal multi-user telepresence and teleaction system," *Int. J. Robotics Research*, vol. 29, no. 10, pp. 1298–1316, 2010.

[14] A. Kron, G. Schmidt, B. Petzold, M. Zäh, P. Hinterseer, E. Steinbach, *et al.*, "Disposal of explosive ordnances by use of a bimanual haptic telepresence system," in *IEEE Int. Conf. Robotics and Automation*, vol. 2. IEEE, 2004, pp. 1968–1973.

[15] W. R. Chitwood Jr, L. W. Nifong, W. H. Chapman, J. E. Felger, B. M. Bailey, T. Ballint, K. G. Mendleson, V. B. Kim, J. A. Young, and R. A. Albrecht, "Robotic surgical training in an academic institution," *Annals of Surgery*, vol. 234, no. 4, pp. 475–486, 2001.

[16] E. C. Grigore, K. Eder, A. G. Pipe, C. Melhuish, and U. Leonards, "Joint action understanding improves robot-to-human object handover," in *IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*. IEEE, 2013, pp. 4622–4629.

[17] C.-M. Huang, M. Cakmak, and B. Mutlu, "Adaptive coordination strategies for human-robot handovers," in *Proc. Robotics: Science and Systems*, 2015.

[18] F. Dehais, E. A. Sisbot, R. Alami, and M. Causse, "Physiological and subjective evaluation of a human–robot object hand-over task," *Applied Ergonomics*, vol. 42, no. 6, pp. 785–791, 2011.

[19] M. Huber, H. Radrich, M. Rickert, A. Knoll, T. Brandt, and S. Glasauer, "Human-robot interaction in handing-over tasks." in *RO-MAN 2008*, Munich, Germany, Aug. 2008, pp. 107–112.

[20] J. Mainprice, M. Gharbi, T. Simeon, and R. Alami, "Sharing effort in planning human-robot handover tasks." in *Int. Symp. Robot and Human Interactive Communication (RO-MAN)*, Kissimmee, FL, Sept. 2012, pp. 764–770.

[21] A. Koene and M. P. Remazeilles, "Relative importance of spatial and temporal precision for user satisfaction in human-robot object handover interactions," in *Third Int. Symp. New Frontiers in Human-Robot Interaction*, 2014.

[22] Z. Li, P. Moran, Q. Dong, R. J. Shaw, and K. Hauser, "Development of a tele-nursing mobile manipulator for remote care-giving in quarantine areas," in *IEEE Int. Conf. Robotics and Automation*, 2017.

[23] A. Cao, K. K. Chintamani, A. K. Pandya, and R. D. Ellis, "Nasa tlx: Software for assessing subjective mental workload," *Behavior research methods*, vol. 41, no. 1, pp. 113–117, 2009.