MeBot: A robotic platform for socially embodied telepresence

Massachusetts Institute of Technology

MeBot
A Robotic Platform for Socially Embodied Telepresence

Sigurdur Orn Adalgeirsson
MIT Media Laboratory
20 Ames Street E15-468
Cambridge, MA 02139
siggi@media.mit.edu

Cynthia Breazeal
MIT Media Laboratory
20 Ames Street E15-468
Cambridge, MA 02139
cynthiab@media.mit.edu

Figure 1.   A picture of the MeBot V4.

Abstract—**Telepresence refers to a set of technologies that allow users to feel present at a distant location; telerobotics is a subfield of telepresence. This paper presents the design and evaluation of a telepresence robot which allows for social expression. Our hypothesis is that a telerobot that communicates more than simply audio or video but also expressive gestures, body pose and proxemics, will allow for a more engaging and enjoyable interaction. An iterative design process of the MeBot platform is described in detail, as well as the design of supporting systems and various control interfaces. We conducted a human subject study where the effects of expressivity were measured. Our results show that a socially expressive robot was found to be more engaging and likable than a static one. It was also found that expressiveness contributes to more psychological involvement and better cooperation.**

Keywords-**Human robot interaction; telepresence; robot-mediated communication; embodied videoconferencing;**

I. INTRODUCTION

The fundamental aim of telepresence research is to allow people to be in two places at the same time. There are many reasons why we might want to occupy two spaces at the same time, examples include wanting to provide safer working environments, perform surveillance, attend meetings or simply spend time with our loved ones. Different situations pose different requirements of the communication medium, and therefore many different telepresence systems with different capabilities have been developed.

Face-to-face interaction is still the golden-standard in communication, against which all platforms are compared. This is partly due to the rich set of social behaviors and cues that we as humans know and share. The reason why face-to-face interaction is preferred might be that the non-verbal cues that are exchanged, can contribute to feelings of engagement, liking, trust, persuasion etc.

Embodiment and *immersion* are concepts that are frequently used in the telepresence literature. *Embodiment* refers to the level presence that people interacting with the robot experience, *immersion* refers to the level of engagement or involvement the operator experiences. Many systems focus on providing deep levels of immersion and much research has gone into haptic feedback systems towards

that goal [1]. Embodiment has been the focus of many systems for different purposes, some applications require a high level of dexterity at the remote site and therefore systems are developed that provide high resolution in motion [2]. Telerobots meant for communication need to embody the operator in a way that provides them with adequate representation in the remote space so that they can take a fully involved part in the interaction and be perceived by their collaborators as being equally present. It is the belief of the authors that a socially expressive embodiment is needed.

This paper presents the design of a telepresence robot that allows the operator to express some of the non-verbal behavior that people use in face-to-face interactions such as hand and head gestures, postural mirroring, interpersonal distance, and eye contact. This is a novel design that integrates video and audio of the remote operator's face with mechanical embodiment of physical gestures of the arms and head, and desk top mobility. The platform is also easily portable which increases its range of applications by allowing for "roaming" interactions.

Novel interfaces for intuitive control of the robot are introduced as well as means to address the issue of eye-

contact in video-conferencing. These interfaces are designed to mitigate cognitive load in controlling many degrees of freedom (DOFs).

We present a study that evaluates the expressive embodiment of the robot and found that these expressive degrees of freedom contribute to greater engagement, cooperation and enjoyment of the people interacting with the robot-mediated operator.

II. BACKGROUND

A. Telerobots for Communication

Early telepresence robots, designed to explore the social aspects of remote presence, were developed in the mid-1990s. Some of the first experiments were performed by Eric Paulos and John Canny at UC Berkeley [3]. Their initial telerobots were blimps fitted with webcams, microphones and speakers but later they developed the *Personal Roving Presences* or PRoPs which allowed people to roam about an office and provided some embodiment. Much research effort has gone into telerobotics for healthcare. Examples include home care assistance [4], interpersonal communication for elderly care [5] and a robotic teddy bear for early education, family communication and therapeutic purposes [6].

Field trials have been performed to evaluate the public's acceptance of tele-operated service robots in public places such as malls or subway stations [7]. For these particular experiments, researchers developed autonomously generated non-verbal behavior to accompany speech from the operator.

An effort was made to develop a conversational robot (autonomous system, not for communication) that made use of non-verbal channels like facial expression, pointing and posture [8]. The researchers showed that people who had conversations with the robot when it was expressive reported a higher level of conversational turn-taking, more natural behavior and more natural utterances than people who conversed with a static robot.

A few papers have explored remote teleconferencing using physical embodiment. Sakamoto et al. compared normal videoconferencing to embodiment via humanoid robot of identical appearance. Results showed that people experienced much stronger presence with the android but that they also felt that it was very uncanny [9]. Some work was done to investigate the difference between 3D movement and embodiment that reached out of the computer screen compared with conventional 3D rendered on a 2D substrate [10].

B. Immersive Virtual Environments

IVEs have been used extensively by researchers in the fields of *Communication* and *Social psychology* to measure different levels of presence and what affects it. Bailenson et al. investigated how virtual reality avatar realism both in behavior and appearance, affected levels of presence-metrics in subjects [11]. Researchers have shown an increase in the measure of social presence between 2D and 3D videoconferencing environments [12]. Increase in social presence and interpersonal trust has been shown to result from the use of virtual avatars for net-based collaborations [13].

III. SYSTEM DESIGN

Our longer term goals are to investigate which aspects of physical embodiment are most important for conveying social expressiveness as well as identifying applications for which socially expressive telerobots are especially useful. Our goals for the work presented in this paper were twofold: First, to develop a desktop telerobot to investigate the impact of physically embodied social gestures and movements on remote collaboration. Secondly, to develop interfaces for the remote operator that allow them to convey a rich repertoire of non-verbal cues while mitigating cognitive load of controlling many degrees of freedom remotely.

A. Robot Design

The design of the robot went through several iterations before we concluded that we had a sufficiently capable platform for evaluating our core research questions. The prototypes can be seen in figure 2.

1) Early Prototypes: The first prototype was a simple 3-DOF robot that mobilized a *Nokia N810* device. This robot provided mobility and a *head-tilt* DOF for looking up and down (the design is further explained in [14]). By building and testing this prototype we received a lot of feedback that moved us in the direction of the next version.

The second prototype included significant redesign, including a custom DC motor control scheme called *MCB Mini*, a different mobile device (*OQO model 02*) and a suite of sensors in the base: range sensors for obstacle avoidance, and edge detection sensors for not falling off the edge of the table. We added a *head-pan* DOF to increase the expressive capabilities of the robot. We also added two arms consisting on one DOF each to enable simple gesturing capabilities.

The third prototype improved upon V2 by elevating the head and separating the *head-pan* DOF from the shoulders as well as adding an *arm-rotate* DOF to each arm. This prototype was completed in design but never implemented because of anticipated physical stability issues.

2) Final Version: The fourth and final prototype of the robot was the result of a complete re-design of the system. This design was a mobile and portable base that had seven range finding sensors for obstacle avoidance and edge-detection for not driving off the table. It had two 3 DOF arms which allowed for arbitrary pointing direction and a fairly rich set of expressive hand gestures. The face of the robot was mounted on a 3 DOF neck (*head-pan*, *head-tilt* and *neck-forward*). The neck allowed a range of expressions (shyness/timidness/reservedness vs. curiosity/excitement/engagement) as well as iconic gestures such as nodding and shaking in addition to arbitrary looking directions.
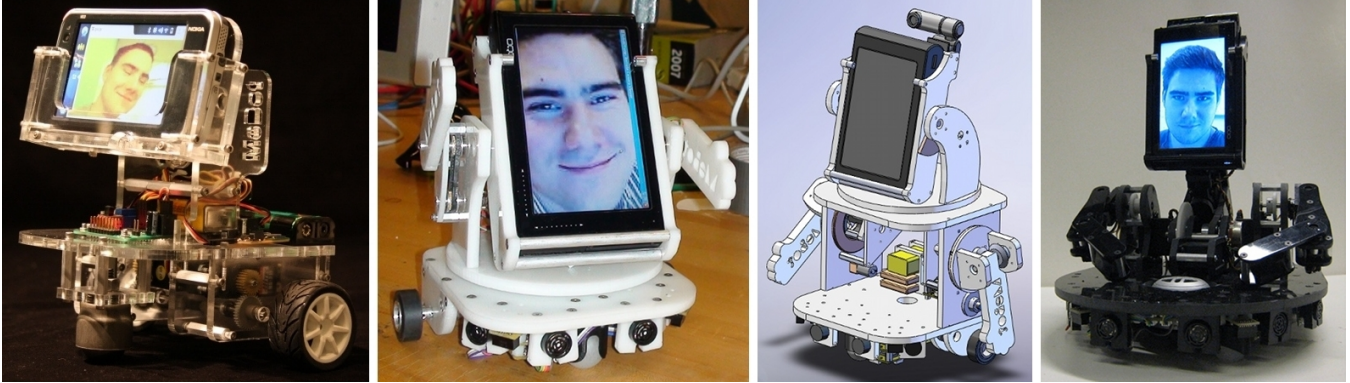
Figure 2. Pictures of the succession of robot prototypes. From left to right: Mebot V1 through V4.
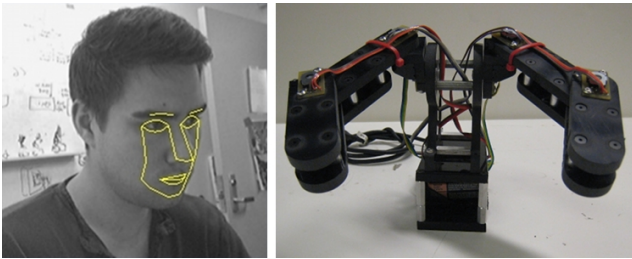

Figure 3. On the left: A view of the FaceAPI interface, the operator can control the neck and head of the robot simply by moving their own head. On the right: On the right: The passive model controller, the operator can control the robot's arms and shoulders by manipulating this passive model.

B. Interfaces and Control

1) Head and Neck Control: When people communicate face to face, a lot of information is encoded in the movements and posture of the head. Obviously facial expressions play an important role in setting the context for the conversation, dictating non-standard meanings of words and phrases, establishing turn-taking, and revealing internal state like confusion, agreement etc. Head and neck movement can in a similar manner influence or assert all of those qualities (head nodding and shaking being the most obvious examples).

In our initial design we considered giving the user explicit controls for the head movement. We soon abandoned that idea as we anticipated that if the user had to think about making an expressive gesture and then performing it then it would already be too late for it to be synchronized with the dialog. Also the operator might not even be fully aware of all expressive movements of their head as some of them come so naturally to them.

We chose to capture these subtle movements automatically and have the robot's neck and head move according to its operator's head. We found a software library that performs head-pose and orientation estimation given a video feed, called *FaceAPI* by *Seeing Machines*. This library was used

to sense the head movements of the operator and then we performed the necessary mappings to have the robot's head and neck move in the same way.

The effect of this mode of control was fairly natural looking head movement of the robot as well as a natural interface for the operator to control a part of the robot without much cognitive load or even the use of their hands.

2) Arm and Shoulder Control: Providing intuitive and easy-to-use control of the arms and shoulders of the robot can be difficult and we considered multiple different methods for achieving this. We thought about whether we should provide direct control to the arms in a similar way to how the head was controlled. Towards this end we tested optical and inertial measurement systems. The optical tracking was achieved by having the operator wear gloves with passive reflective markers and sense them using a *Vicon* optical tracking system. For inertial measurement we used gloves with 9-DOF *Inertial Measurements Units* (IMUs). Both of these systems required the operator to wear sensors and both of them were fairly expensive. They also introduce the added difficulty of managing which movements should be mapped onto the robot and which ones should not; that is, how should the operators "decouple" themselves from the system to perform ordinary tasks such as typing on a keyboard etc.

We decided that a less intrusive and possibly easier method of control would be to build a passive model controller for the arms of the robot. The controller was a passive model of the robot that had joints in the all the same places and when they were moved, the robot would move in a corresponding way. When the operator would let go of the joints, say to perform a task using their hands that doesn't have anything to do with the interaction, the model simply stays in place and the movement of the robot is not affected. This design was inspired by a similar method of control for a robot called The Huggable [6]. The passive model controller can be seen in figure 3.

3) Navigation: For navigating the robot around, the operator used a device called *Space Navigator* by *3DConnexion*.

Figure 4. On the left: The fully articulated 3D model of the robot as it is displayed to the operators. On the right: The custom built camera-embedded display, as the operator observes the remote scene, their video is being captured in a way that establish eye-contact between the participants.

This device is a 3D mouse and it allowed the operator to rotate and translate a target location relative to the robot's current location. This was visualized on an overhead display that also showed the sonar range data as well as other logistic information such as the battery voltages.

4) Eye-Contact: Many current videoconferencing systems and telepresence robots mismanage eye-contact in the interaction. Most people that have used videoconferencing tools agree that the effect of having a discussion with somebody who doesn't seem to look you in the eye can be very disruptive for the quality of the interaction. It was our belief that by simply fixing the eye-contact problem with our system, we could improve the quality of telepresence interactions drastically.

To address this problem we designed and built a display with a camera embedded in the center of it. The remote video window was projected onto the center of this display by a video-projector placed behind or above the operator. When the operator was controlling the robot and talking to their partner while watching them in the video window, the camera was looking right back at them under an angle that is very close to $0°$. This produced the effect that the operator was looking straight forward from the perspective of a person watching the robot and when the head of the robot is faced towards a local participant, eye-contact could be established. A picture of the camera-embedded display can be seen in figure 4.

C. Software

Each prototype of the robot needed different software to interface with the different style of motor control schemas and mobile devices. This section will describe parts of the software for the fourth prototype of the robot.

1) Real-Time Media Streaming: Java was used for the bulk of the code that was written for this project and it leveraged the *Personal Robots Group's* codebase called *c6*. The real-time streaming of audio and video was performed using the *Java Media Framework (JMF)*. *JMF* provides a nice structure for media transfer that has a pipeline analogy.

Programmers can edit the media data by creating custom filters and insert them at appropriate locations within the pipeline.

2) Face-Cropping: In an effort to have the person who is interacting with the robot more clearly perceive the robot operator as embodied by the robot we decided to only stream the region of the video that contains the face of the operator. We built a custom *JMF* video-filter that uses *OpenCV* and some persistence filtering to find and track the face as well as extracting that portion of the video and streaming it. This way the operator could comfortably focus their attention on the interaction and control of the robot. The *Face-Cropping* module would make sure that even if the operator was moving around slightly, their face would be centered and rendered in a full-screen view on the robot.

3) Visual Control Feedback: Since in many practical applications the robot could be out of the operator's view, we are faced with the problem of the operator not fully understanding the effect of their controls. To "close the feedback loop" so to speak, we designed a fully articulated 3D model of the robot and displayed it to the operator. Using this model, the operator could directly observe the effects of their control. The 3D model can be viewed in figure 4.

IV. Experiment

Our hypothesis is that making telepresence systems socially expressive by affording them the ability to convey their operators' non-verbal behavior such as gestures, body language and proxemics, can make remote interactions more present, more engaging and more enjoyable. We also believe that a system that allows for social expression will foster collaboration and cooperation.

To test this claim we designed an experiment where we could evaluate the experience of a local collaborator when interacting with a robot-mediated operator. We specifically wanted to learn about how the social expressiveness affected the collaborators experience so we decided to run a between-subjects study with two conditions:

1) **Expressive body condition:** The arms of the robot move according to the control input from the operator via the passive model controller and the head moves according to the head movements of the operator.

2) **Static body condition:** The robot is in a neutral and still pose during the whole interaction.

A. Hypotheses

We set forth several hypotheses for the outcome of our experiment. The actual outcome and validity of these claims is investigated in the *Discussion* part of this section.

H1 - Co-presence: People would experience stronger co-presence when they interacted with an expressive telerobot

H2 - Psychological involvement: People would experience more psychological involvement with their partner when they interacted with an expressive telerobot.
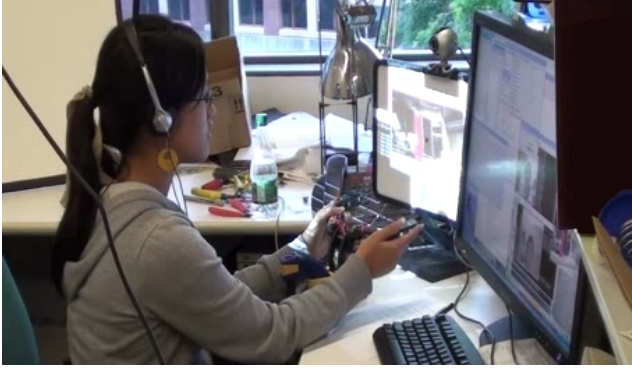
18

Figure 5. A picture of the operator and the setup in the operator station. The operator has the control interface in front of her, as well as the space navigator for mobility and the sympathetic for expression. She also has the remote scene projected on the custom camera-embedded screen.

H3 - Trust: People would trust their partner more when they interacted with an expressive telerobot.

H4 - Engagement: People would feel more engaged with their partner when they interacted with an expressive telerobot.

H5 - Cooperation: People would cooperate better with their partner when they interacted with an expressive telerobot.

H6 - Enjoyment: People would enjoy their interaction more when they interacted with an expressive telerobot.

B. Operator

The study was a doubly blind study, that is to say that the operator was not aware of the hypotheses nor the conditions in any of the experiments. For consistency, the robot was operated by the same person in both conditions and for all participants. The operator controlled the robot in all of the interactions from a private office in our lab. The office setup can be seen in figure 5.

The office door was closed so that participants wouldn't see the operator in person until after the experiment. The operator used three interfaces to control the robot, a passive model controller for the shoulders and arms of the robot, head-pose and orientation for controlling the neck and head and a graphical interface to control those DOFs in case either of the other two would stop functioning.

Much emphasis was placed on the operator performing consistently between interactions and especially between conditions. Effort was made to make the operator unaware of which condition was being performed. The camera on the robot was moved from the head, which is where it would normally be, down to the base so that the operator wouldn't notice the camera view change as she moved the head of the robot. Noise was added to the operators audio signal so that she would not hear the motor noise that resulted from her moving the arms or head of the robot. The operator

was videotaped so that her behavior could be monitored for consistency between conditions.

C. Study Task

We chose to use a variation on *The Desert Survival Problem* as a task for the operator and participant to work on during their interaction. We chose this task because it is a fairly established method in the field and it has the following characteristics which are beneficial to our evaluation:

- It sparks an active conversation between operator and collaborator.
- It involves real world objects which call for some spatial reasoning and sharing of workspace.
- It has some quantifiable task measures for success.
- It takes under about thirty minutes to finish.

The Desert Survival Problem was originally developed by Lafferty and Eady [15] but has since been used by several social scientists and roboticists [16][13][17][12][18]. We used a slightly modified version of the task, as designed by Takayama et al. [16].

D. Study Setup

The participants were seated in a space that had been partitioned from the rest of the lab by black curtains. They would be seated at the center of the long side of an approximately six foot long table. Across from the participant, on the other side of the table was a smaller table on which the robot was placed. Along the long table were twelve items arranged neatly in six pairs of two items. The pairs contained the following items: A tarp and a canvas, a knife and a gun, a chocolate bar and a water bottle, a compass and a mirror, a flashlight and matches, a map and a book on edible animals.

The participants were asked to think about which item out of every pair they would prefer to have with them in the desert for survival. They would discuss their choices with the operator and she would provide arguments either for or against bringing those items.

The operator performed as closely as she possibly could in accordance with a script for every interaction. The only reason for her to digress from the script would be to respond to unanticipated questions so that the participant experienced natural dialog. The script contained arguments for and against every item that was available for the participants to select. The operator would always agree with the participants' choices on the same two pairs, but disagree on the rest. This way, she would have disagreed with every participant on the same four pairs and we could then investigate how many participants would decide to change their initial selections on these four items after having heard the scripted arguments for the other items.

E. Dependent Measures

We decided that the following measures would be relevant to our experiment:

Dep. var.	S $\overline{M}$	S σ	E $\overline{M}$	E σ	$p <$
Co-presence	5.373	0.936	5.742	0.749	0.164
Psych. inv.	5.241	0.548	5.714	0.711	0.022
Beh. eng.	4.509	0.544	4.968	0.694	0.024
Trust	5.632	0.950	5.612	1.437	0.959
Gen. eng.	4.700	1.089	5.496	0.944	0.015
Cooperation	4.982	0.959	5.870	0.978	0.005
Enjoyment	5.474	1.172	6.261	0.619	0.008
Items changed	1.842	1.015	1.696	0.974	0.637

1) Social Presence: Social presence has frequently been used as a quality measure of different communication media [12][13] and it is particularly relevant to our system for what we want to facilitate is exactly social expression and behavior. We used a widely accepted method for measuring social presence, introduced by Biocca et al. [17]. Biocca et al. define the fundamental dimensions of social presence as co-presence, psychological involvement and behavioral engagement.

2) Trust: An important application for telepresence are business meetings. This is the most common application of telepresence systems and has been the major focus for commercialization of products in this domain. With modern globalization of industries and governments, business is conducted between distant regions very frequently and in such transactions, trust is vital. This is why we thought that the measure of trust would be a useful metric to assess the success of our system. We use the word trust, but more specifically we mean trust as it applies to reliability. We used Wheeless's and Grotz's *Individualized Trust Scale* as reported in [19] to measure this metric.

3) Cooperation: Collaborative meetings are another application of telepresence that hasn't received quite the same amount of attention as business meetings but is equally or even more relevant to this technology. By collaborative we mean a more active meeting, such as that of designers, artists or developers. These types of meetings usually involve the active development of either a tangible product or idea that requires the complete engagement and participation of every collaborator. Cooperation is a measure that could be vital to the assessment of our system's compliance to the requirements of these types of situations. We used an assembly of questions from Takayama et al. [16] to measure this metric.

4) Engagement: Throughout the design and idealization of this project we have had family communication in mind as an important application that needs significant attention. We want our system to facilitate family communication that allows a deeper engagement and a more enjoyable experience than that which can be obtained via simple videoconferencing. Engagement is a relevant measure in this sense. We used questions from [20] to measure this metric.

5) Number of Changed Items: This measure comes from the specific task that the operator and local collaborator work on together. This measure reveals how many items the collaborator changed after hearing the operator's arguments for the alternative one.

F. Results

1) Participants: A total of 48 people participated in the study, six participants' data was excluded from the analysis leaving a total of $n = 42$ participants (24 female, 18 male). Exclusion criteria was set by two rules: If participant experienced severe technical difficulty with the robot or if it was discovered that the participant already knew the remote operator personally prior to the study, they would be excluded.

Participants were asked to rank their knowledge about robotics on a seven point scale ($\overline{M} = 3.80$, $\sigma = 1.52$). They were also asked to rank their knowledge about computers on a seven point scale ($\overline{M} = 5.18$, $\sigma = 1.36$). The participants were asked for their age in years ($\overline{M} = 23.21$, $\sigma = 8.92$).

2) Study Results: The results of the ANOVA analysis as well as means and standard deviations for all dependent variables can be seen in figures 6 and 7 as well as in table I.
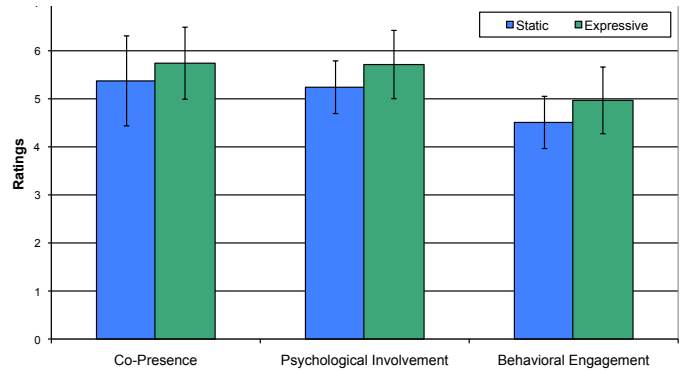

Figure 6. A graph comparing the average ratings for measures of co-presence, psychological involvement and behavioral engagement. Error bars indicate $\pm\sigma$.

G. Discussion

1) Hypotheses Summary: The first section in the Experiment chapter sets forth six hypotheses for what we expect to read into the outcome of the evaluation. Here are the results:

H1 - Co-presence: Hypothesis not upheld.
H2 - Psychological involvement: Hypothesis upheld.
H3 - Trust: Hypothesis not upheld.
H4 - Engagement: Hypothesis upheld.
H5 - Cooperation: Hypothesis upheld.
H6 - Enjoyment: Hypothesis upheld.

As we can see, the measures of co-presence and trust did not yield the expected results, that is neither of those metrics
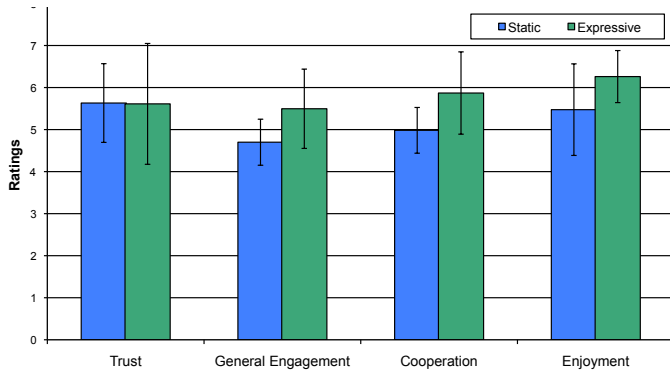
Figure 7. A graph comparing the average ratings for measures of trust, general engagement and cooperation. Error bars indicate $\pm\sigma$.

measured a statistically significant difference between the two conditions of the study.

2) Hypotheses Details:

H1 - Co-presence: Not a significant difference. The authors wondered if the fact that both conditions consisted of an embodied partner via a mobile robot which looks capable of expressive communication had a strong enough effect on the measure of presence that the difference between an expressive and static robot was too "fine grain" to detect a difference. This issue will be addressed in a future study with the same interaction, without the robot, only a graphical device and speaker.

The authors also noticed that this portion of the questionnaire, as adapted from [17], seemed to be better suited for virtual environments than for robot-mediated communication. These questions had an ambiguous meaning in this context and caused confusion with the participants.

H2 - Psychological involvement: Expressive case rated higher. This factor of the study speaks strongly for socially expressive telerobots as participants who experienced the expressive case of the study reported higher values of understanding, clarity and emotional influence in the interaction.

H3 - Trust: Not a significant difference. This metric was measured using fifteen sets of 7 point Likert-scales with antonyms (ex. unreliable - reliable) at each extreme. The participants would rate how they experienced the operator through this system on those scales. The result did not show a statistically significant difference.

The authors noticed that most participants simply selected the most positive options in this section. Participants were notified that the results were anonymous but there is still a possibility that they got too involved with the operator during the interaction to be willing to give her a "bad" rating in the questionnaire. Possibly more care should have been taken to explain that this was not a personal measure of the operator but more so an evaluation the system. This was a complicated boundary to manage.

H4 - Engagement: Expressive case rated higher. Mutual assistance was the subgroup of behavioral engagement that showed the strongest difference between the conditions.

The result suggests that a socially expressive robot elicits more helpful behavior from its partner as well as being perceived more helpful by the partner. This is an interesting finding as one could just as well have thought that a robot that seems less capable might elicit more helpful behavior than the one that looks like it is more capable of helping itself.

H5 - Cooperation: Expressive case rated higher. Strong statistical difference was measured in favor of the expressive case. Cooperation in this sense means the willingness to cooperate with the robot-mediated partner as well as the perceived willingness of the partner to cooperate.

Behavior that is descriptive of an authoritative figure like a policeman or superior at work or a parent is usually a firm, assertive and static posture while that of a peer, be it a friend, a coworker or a sibling is usually more animated and playful. This might affect the perceived hierarchy between the robot-mediated partner and the participant and make the static robot look more authoritative while the expressive robot could be perceived more as a collaborator or peer.

This could in turn have an effect of the perceived quality of cooperation in the interaction.

H6 - Enjoyment: Expressive case rated higher. This metric was only measured by one question as opposed to averaged groups of questions and its result should therefore be interpreted more as an indicator than a concrete measurement.

The indication was that people who experience the expressive robot report higher levels of enjoyment from the interaction with a confident difference. This result above all others pleased the robot designer.

V. CONCLUSIONS

We presented a description of an iterative design process of a portable and mobile, socially expressive telerobot. The robot was built to ask specific questions about the effect of physically embodying the operators in a stronger way than video allows and enabling them to express some of the non-verbal ques that are commonly used in face-to-face interactions.

We conducted an experiment that evaluated how people perceived a robot-mediated operator differently when they used a static telerobot versus a physically embodied and expressive telerobot. Results showed that people felt more psychologically involved and more engaged in the interaction with their remote partners when they were embodied in a socially expressive way. People also reported much higher levels of cooperation both on their own part and their partners as well as a higher score for enjoyment in the interaction.

The present work strongly indicates that telepresence technologies could benefit from enabling their users to express their non-verbal behavior in addition to simply passing their audio and video data. This is particularly true in applications where deep engagement in the interaction is important as well as cooperation. The obvious examples are business and organizational meetings but even more so collaborative meetings and other events that demand more active participation from their members.

More work must be done to better understand the effects of social expression in telerobots. A similar study to the one presented, but with no robot, only video-conferencing, would help us understand how much the mere embodiment of a robot affects the interaction. Further research could be done to understand better which aspects of expression contribute the most to the observed improvements in quality of interaction.

REFERENCES

[1] A. Fisch, C. Mavroidis, J. Melli-Huber, and Y. Bar-Cohen, "Haptic devices for virtual reality, telepresence, and human-assistive robotics," *Biologically Inspired Intelligent Robots*, p. 73, 2003.

[2] L. Li, B. Cox, M. Diftler, S. Shelton, and B. Rogers, "Development of a telepresence controlled ambidextrous robot for space applications," in *IEEE International Conference on Robotics and Automation*, 1996, pp. 58–63.

[3] E. Paulos and J. Canny, "Social tele-embodiment: Understanding presence," *Autonomous Robots*, vol. 11, no. 1, pp. 87–95, 2001.

[4] F. Michaud, P. Boissy, H. Corriveau, A. Grant, M. Lauria, D. Labonte, R. Cloutier, M. Roux, M. Royer, and D. Iannuzzi, "Telepresence robot for home care assistance," *Procceedings of AAAI*, 2006.

[5] T. Tsai, Y. Hsu, A. Ma, T. King, and C. Wu, "Developing a telepresence robot for interpersonal communication with the elderly in a home environment," *Telemedicine and e-Health*, vol. 13, no. 4, pp. 407–424, 2007.

[6] W. Stiehl, J. Lieberman, C. Breazeal, L. Basel, R. Cooper, H. Knight, L. Lalla, A. Maymin, and S. Purchase, "The huggable: a therapeutic robotic companion for relational, affective touch," in *3rd IEEE Consumer Communications and Networking Conference, 2006. CCNC 2006*, vol. 2, 2006.

[7] S. Koizumi, T. Kanda, M. Shiomi, H. Ishiguro, and N. Hagita, "Preliminary field trial for teleoperated communication robots," in *Robot and Human Interactive Communication, 2006. ROMAN 2006. The 15th IEEE International Symposium on*, 2006, pp. 145–150.

[8] T. Tojo, Y. Matsusaka, T. Ishii, and T. Kobayashi, "A conversational robot utilizing facial and body expressions," in *2000 IEEE International Conference on Systems, Man, and Cybernetics*, vol. 2, 2000.

[9] D. Sakamoto, T. Kanda, T. Ono, H. Ishiguro, and N. Hagita, "Android as a telecommunication medium with a human-like presence," in *Proceedings of the ACM/IEEE international conference on Human-robot interaction*. ACM New York, NY, USA, 2007, pp. 193–200.

[10] N. Negroponte, "Talking heads," *Discursions (Video Disc), Side One track 14*, 1983.

[11] J. Bailenson, K. Swinth, C. Hoyt, S. Persky, A. Dimov, and J. Blascovich, "The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments," *Presence: Teleoperators & Virtual Environments*, vol. 14, no. 4, pp. 379–393, 2005.

[12] J. Hauber, H. Regenbrecht, A. Hills, A. Cockburn, and M. Billinghurst, "Social presence in two- and three-dimensional videoconferencing," in *PRESENCE 2005: The 8th Annual International Workshop on Presence*, 2005.

[13] G. Bente, S. Rüggenberg, and N. Krämer, "Social presence and interpersonal trust in avatar-based, collaborative net-communications," in *7th Annual International Workshop on Presence*, 2004.

[14] Y. Gu, "Semi-autonomous mobile phone communication avatar for enhanced interaction," June 2008, undergraduate Thesis for the Department of Mechanical Engineering, MIT.

[15] E. P. Lafferty, J.C. and J. Elmers, *The Desert Survival Problem*, ser. Experimental Learning Methods, Plymouth, Michigan, 1974.

[16] L. Takayama, V. Groom, and C. Nass, "I'm sorry, Dave: i'm afraid i won't do that: social aspects of human-agent conflict," in *Proceedings of the 27th international conference on Human factors in computing systems*. ACM New York, NY, USA, 2009, pp. 2099–2108.

[17] F. Biocca, C. Harms, and J. Gregg, "The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity," in *Presence 2001, 4th Annual International Workshop on Presence*, 2001, pp. 21–23.

[18] C. Kidd and C. Breazeal, "Effect of a robot on user perceptions," in *IEEE/RSJ International Conference on Intelligent Robots and Systems, 2004. (IROS 2004). Proceedings*, vol. 4, 2004.

[19] R. Rubin, P. Palmgreen, H. Sypher, and M. Beatty, *Communication research measures: A sourcebook*. Lawrence Erlbaum, 2004.

[20] M. Lombard, T. Ditton, D. Crane, B. Davis, G. Gil-Egui, K. Horvath, J. Rossman, and S. Park, "Measuring presence," in *Third International Workshop on Presence, Delft, The Netherlands*, vol. 240. Citeseer, 2000.