On the Effect of Operator Modality on Social and Spatial Presence during Teleoperation of a Human-Like Robot

Christian Becker-Asano[1] and **Severin Gustorff**[1] and **Kai O. Arras**[2] and **Bernhard Nebel**[1]

Abstract. With an increasing availability of affordable and effective robotic telepresence systems, key questions in the design of such systems arise, in particular when they aim at untrained users. Previous research regarding telepresence systems has focused on the (mobile) robotic platforms themselves or the differences between virtual as compared to physical representations thereof. The design space of the operator interface, i.e., the operator modality, with its potential impact on presence, however, has not been explored systematically.

This paper reports results of an empirical study investigating how two different operator modalities impact the perceived spatial and social presence of operators in dyadic remote multi-modal interaction. The robot Daryl, used as telepresence medium in our study, features three degrees of freedom in its head unit as well as a stereo camera system. This enabled the transmission of a stereo, first-person perspective, which was used by the operator in combination with a head-mounted display whose movements were tracked to drive the robot's head. Compared to a previously realized console-based operator interface, our results show significantly higher spatial as well as social presence for the head-mounted display modality while no significant difference in task performance was found. We conclude that for robotic telepresence platforms with mobile head units and stereo camera systems it seems advisable to use a head-mounted display as part of the teleoperation interface in order to provide operators with a particularly immersive remote presence experience.

1 Introduction

Research in teleoperation of vehicles, aircrafts, and robots has received great attention over the last decades and the technologies used to teleoperate such systems have been studied intensively. While teleoperation typically involves physical interaction of a remotely controlled machine with its environment, telepresence refers to a wider range of interaction types including human-human communication. Recently, a new generation of affordable mobile telepresence robots are being announced by an increasing number of companies. With the promises of "replicating a person in a distant location" they are currently entering the market of domestic and professional domains including remote working, assisted living or telemedicine[3]. At the same time, the performance of head-mounted displays (HMDs) also increased together with a significant drop in purchase price – recent examples include Oculus Rift [16] or Sony HMZ-T2 [25] –, motivating a reevaluation of its potential for such telepresence applications.

Apart from these commercial efforts, the biggest challenge to developing telepresence lies in achieving a sense of "being there" at a remote environment [17]. Previous studies have investigated this question with respect to mobile robotic platforms [14, 8], anthropomorphic robot design [7], and the difference between virtual and physical representation of the robot [11]. Unlike those works, we focus on the operator's properties and the human interface technology that enables the operator to interact through such robots. Thus, we investigate the impact that the operator modality has on the perceived social and spatial presence in remote human-human communication.

Research in virtual reality agents typically concentrates on the operation of virtual avatars in order to refine the concept of 'presence' or 'telepresence' [9, 27, 20, 24, 6, 18]. Here, presence is understood as a central indicator of how well a person's mind and attention is attracted by a remote or virtual environment with its people and events. It is also assumed that higher levels of presence might lead to better performance in carrying out different kinds of tasks. Lately, researchers tried to maximize this subjective impression in operators of physical avatars, robots respectively [14, 11, 7]. However, few efforts have been made to explore how a human-like robot's physical abilities together with different types of operator modalities affect social and spatial presence.

In our study reported here, a mildly humanized robot is teleoperated by subjects over head motion (three degrees of freedom) and bidirectional, stereo audio and video to empirically investigate the impact of different operator modalities on presence while carrying out a combined task with spatial and social interaction features. This work follows our proposed research questions related to cross-combinations of modalities and human-like robots, outlined in [3, 1]. Note that by using the terms 'mildly humanized' or 'human-like', we refer to the robot's motion capabilities to provide a human-like configuration of degrees of freedoms, particularly in the head unit, and not to the robot's overall appearance.

2 Related work

In a number of Virtual Reality setups it was found that task performance is lower when an HMD is used as presentation device as compared to standard desktop interfaces [20, 18]. However, the tasks involved in these studies did not include any social component and differences in an operator's subjective level of presence were not investigated. Therefore, it seems interesting to ask, if similar effects on task performance can be observed for (1) tele-robotic applications that (2) include a social component.

[1] Artificial Intelligence Lab, University of Freiburg, 79110 Freiburg, Germany, email: [basano,gustorfs,nebel]@informatik.uni-freiburg.de

[2] Social Robotics Lab, University of Freiburg, 79110 Freiburg, Germany, email: arras@informatik.uni-freiburg.de

[3] Most prominent product announcements as also cited in [13] include: Ava by iRobot, Beam by Suitable Technologies (former Texai by Willow Garage), VGO by Vgo Communications, Giraff by HeadThere, RP-VITA by InTouch Health, TiLR by RoboDynamics, Double by Double Robotics

To measure the inherently subjective phenomenon "presence" a multitude of different methods, including at least 28 questionnaires, already had been devised as early as in 2004 [29]. Over the last years, research on telerobotic applications adopted the general distinction of two types of presence, namely *spatial* and *social presence* [11, 5].

Spatial presence refers to the illusion of physical presence at a remote or virtual environment and has been described [5] as bearing similarities to the term "telepresence" [17]. In telepresence systems, immersive technology is used to focus our attention towards the distant physical or virtual environment by bidirectional transmission of sensory stimuli. However, an operator's feeling of *spatial presence* is also supported by high levels of interactivity [27] letting it depend on the task as well [11]. The factors latency, number of possible actions, and accuracy of mappings are described as measurable features of interactive systems [27]. The *Temple Presence Inventory* (TPI) [15] became a quasi-standard questionnaire in presence research to assess subjective levels of spatial presence.

Social presence, in contrast, relates to the sense of not only "being there" but moreover "being together with others." An operator socially present (or *co-present*) in a distant location reacts to social cues and "generates models of the intentionality of others" [5]. This viewpoint on presence incorporating other people is different to *spatial presence* since *social presence* is also possible when not feeling physically present. In telepresence research, the *Networked Minds Social Presence Questionnaire* (NMSPQ) [4] became a standard tool for assessing subjective levels of social presence.

With respect to presence in telerobotics applications, a study with a task featuring social and mainly physical interaction (monitoring elderly people) utilized the 'Giraff' mobile telepresence robot together with a console interface [14]. Parts of the TPI and NMSPQ questionnaires were used to measure both spatial and social presence felt by the operator. As the Giraff robotic platform is mobile, the study focused on six types of formations realized by the robot operator together with the interlocutors. A number of correlations between the operator's spatial/social presence and these formations are reported, which are interesting even though they were detected post-hoc.

With a focus on a teleoperator's subjective experience and ease of use of the mobile robot "Robonaut" when being supplied with a first person view, the effects of employing a head-mounted display (HMD) in combination with other sophisticated components as telepresence hardware are described in [8]. The video streams of the two cameras mounted in Robonaut's head were wirelessly transmitted to the HMD and the operator's head movements tracked to drive the robot's rotational head movements accordingly. The subjective experiences, described only on an informal level, indicate that an HMD-based interface is easy to manage and intuitively operated.

A study conducted by Hoffmann and Krämer [11] using the NMSPQ revealed no difference in social presence felt by an interlocutor, when either confronted with a virtual or a physical representation of a small toy robot. Their results suggest, however, that an interaction partner's degree of acceptance and affective state is influenced by the interaction task with a more task-oriented scenario resulting in more positive feelings and better acceptance of the robot than in case of a persuasive-conversational one.

In a recent two-by-two empirical study a total of 32 participants used the "double telepresence robot" to collaboratively construct either small or big versions of a geometrical object. The impact of operator (whole-body) mobility on task performance and presence was investigated following a two-by-two experimental design. Although an operator's feeling of presence is higher in the two mobile conditions (as compared to the two stationary conditions), task completion times for the two versions of the high-mobility tasks as compared to the low-mobility tasks were on average significantly lower [21].

Leaving aside the telerobotic platform's mobility but taking its degree of anthropomorphism to the extremes, android robots such as "Geminoid HI-1" [19] have been developed to investigate, how social presence is influenced by very anthropomorphic design. Comparing Geminoid HI-1 against a video conferencing system and a speakerphone conversation [22] suggests that an operator can convey the highest level of presence when using a console interface to teleoperate Geminoid HI-1.

In summary, previous research on spatial and social aspects of robotic telepresence mainly focused on the robotic platforms themselves and not on the interface an operator is provided with. The technological design space on the operator's side was hardly investigated as another important factor enabling social interaction through these systems. So, we set out to perform a methodologically sound investigation of potential changes in spatial and social presence resulting from more or less immersive human-system interfaces as provided to the operator.

Figure 1. The mildly humanized robot Daryl

3 Telepresence with robot Daryl

For the investigation of the cross-combination of different operator modalities with different telerobotic embodiments [1], the mildly humanized mobile robot Daryl, built at the Social Robotics Laboratory, University of Freiburg, Germany, serves as the operator's avatar, see Fig. 1. Daryl is a custom-made platform with an abstracted yet antropomorphized design for the purpose of research in robotics and particularly in human-robot interaction. It features several degrees of freedom in wheels, torso and head, three of which are used by our telepresence setup. Two cameras in the robot's eyes and two binaural microphones in the robot's ear-like modalities provide stereoscopic vision and hearing. Spoken audio of the operator is outputted to onboard loudspeaker. The setup enables an operator to perceive the remote place through the 'eyes' and 'ears' of Daryl and move its head in three dimensions according to his or her own head motion.

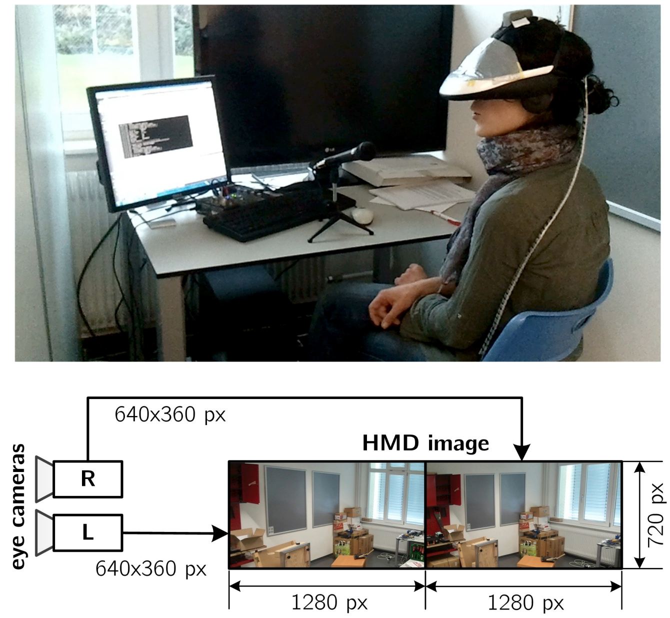

Figure 2. A participant using the HMD (top) and a sketch of the video transformation applied between Daryl's eye cameras and the HMD (bottom). The cameras' native resolution (640x480 px) is first cropped to 640x360 px and then scaled to the HMD's internal resolution of 1280x720 px [25]

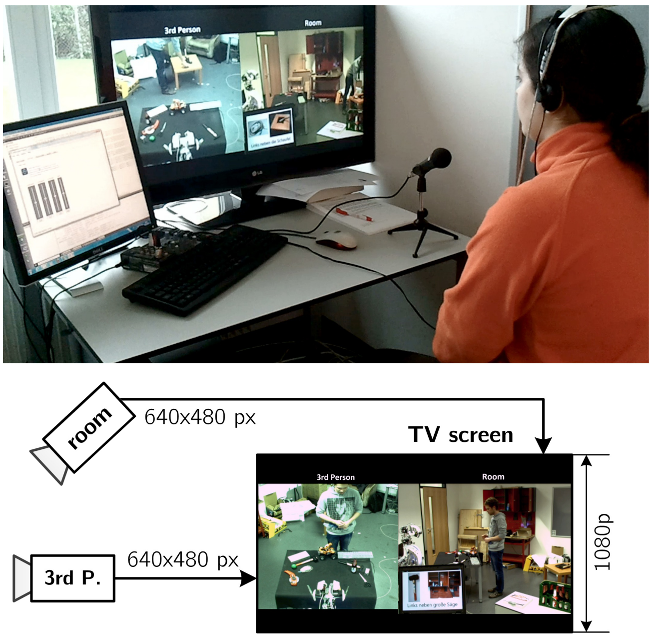

Figure 3. A participant in front of the console interface (top) and a sketch of the video transformation applied to the two static cameras presented on a TV screen (bottom)

3.1 Hardware setup

In this section we describe the two modalities which were realized to analyze for differences in an operator's felt level of presence in particular for remote human-human interaction.

Both modalities use an inertial measurement unit (IMU) [28] to capture the operator's head movements. The IMU is a small USB device mounted on top of the HMD or on top of the headphones in the case of the console modality. It provides angular rotation measurements in three dimensions and control the yaw, tilt, and roll axes in Daryl's head. Measurements are retrieved from the sensor at 50 Hz and are transmitted to the robot for execution.

HMD-modality In this setup, the operator is equipped with a head-mounted display (HMD, Sony HMZ-T1 [25]). The HMD has two integrated displays each with a resolution of 1280x720 pixels, one for each eye and a field of view of 45°. This roughly translates into looking at a 32 inch TV display sitting one meter away from it. As such, the presentation size is much the same as that of the console modality described below. The device also provides high-quality headphones. The HMD shows the live video captured by Daryl's eye cameras merging the frames of both sensors to a stereoscopic field of view, see Fig. 2. The HMD's integrated eye lenses were adjusted to match the inter-pupillary distance of the participant's eyes. A bidirectional audio link is established by a microphone on the operator's desk whose signal is sent to the speaker on board the robot and by the ear microphones in Daryl's head whose signals are transmitted to the operator's headphones.

Console-modality In this modalitiy, our operators use a TV screen and wear IMU-equipped headphones, see Fig. 3 (top). In the remote place, two static cameras capture the scene: one camera has a third-person view onto the scene and is placed above behind the robot, the other camera oversees the setup in the remote room and gives a side view of the robot, see Fig. 3 (bottom) and also Fig. 7. As presentation device, a 32 inch TV display is used. The choice of this setup is motivated for the sake of comparability of results with related work in which the same operator modality was used to remote control android robots [2]. The bidirectional auditory cue is implemented in the same way as for the HMD-modality. Note that the use of static cameras in this condition means that movements of the robot's head do not change the operator's visual feedback. Head motion only serves to support interaction with remote communication partners, e.g., by executing head gestures or looking at objects.

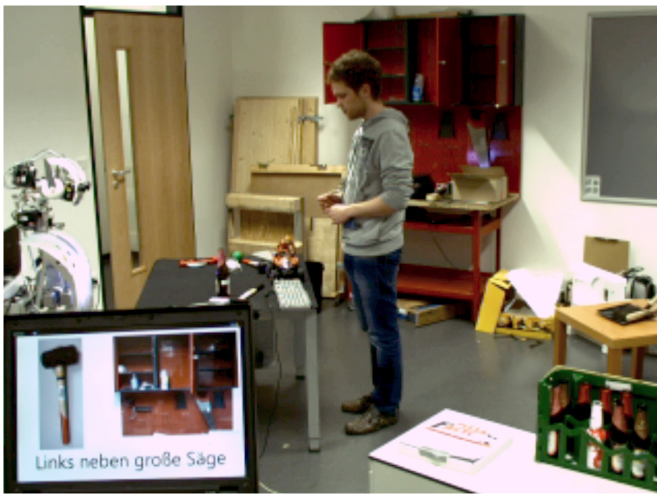

Figure 4. Overview of the experiment room (as captured by the "room" camera for the Console-modality) with the robot on the left side and the interaction partner in front of it. The operator instructions are displayed on a notebook, which is placed as to cover the lower left corner in this view

Figure 4 provides an example of the view provided by the "room" camera used for the Console-modality. An overview of the hardware setup with its components and their interplay is shown in Fig. 5.

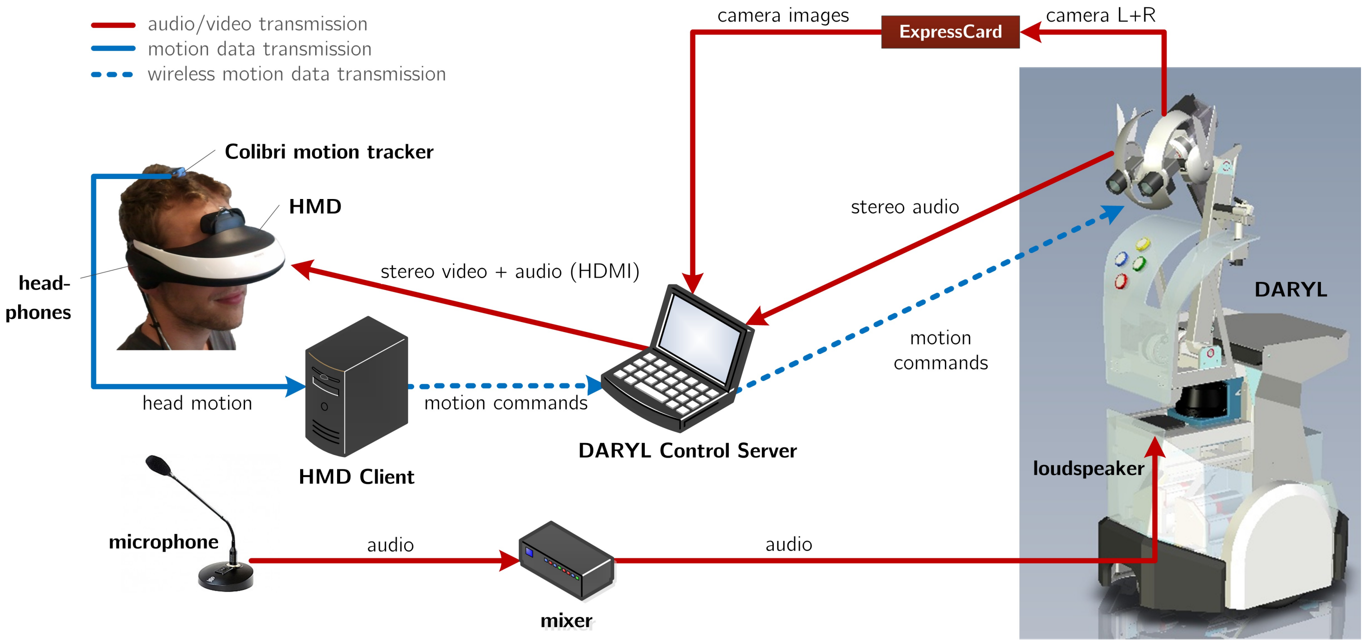

Figure 5. Hardware setup and its components (shown for the case of the HMD modality). The operator's head motion is tracked by a "Colibri" motion tracker [28] and translated into motion commands, which are wirelessly transferred to Daryl to be executed. In parallel the operator's voice is amplified and replayed by Daryl's loudspeaker. Stereo audio and video captured from Daryl's hardware is directly transmitted via HDMI to the operator's HMD to avoid latency issues

3.2 Software setup

Figure 6 shows a flow diagram of the setup's software system. The video streams acquired by Daryl's eye cameras or alternatively by the two static cameras in the console modality are captured and combined into one stream by the broadcasting software *XSplit* [26] at the robot's location. At the operator's location this combined stream is displayed either on the HMD or the TV screen. The angular readings of the IMU attached to the operator's head are translated into motion commands, which are then transmitted over network to the robot's embedded controller. The embedded controller runs a real-time operating system for axis control. As schematized in Fig. 6, several communication protocols are involved including the Geminoid Protocol [19] which is employed since it allows reuse of the same control setup to operate android robots of the Geminoid series planned in future work.

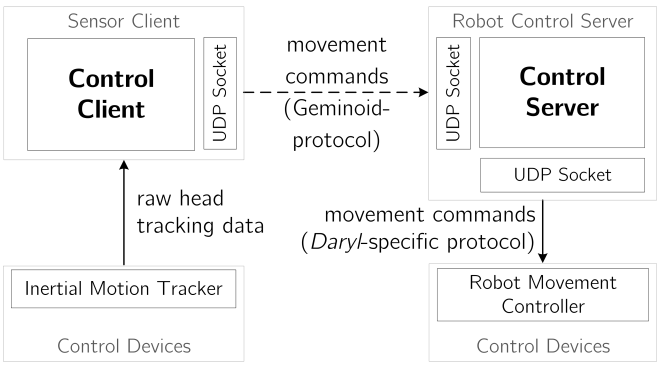

Figure 6. Schematic overview of the client-server software setup

4 Empirical study

In order to find possible differences in the feeling of presence depending on operator modalities, we conducted an empirical study.

4.1 Hypotheses

An operator's ability to move Daryl's head in an intuitive manner is expected to produce a stronger identification with the robotic avatar and to support an operator's sense of spatial presence. In particular, the first-person perspective provided by the stereoscopic eye cameras of Daryl in the HMD modality along with the attenuation of ambient distractions of the operator is expected to result in a stronger sense of spatial presence. These considerations lead to the formulation of our first hypothesis:

Hypothesis 1 (H1): Spatial *presence is higher in telerobotic systems with human-like head motion when using the HMD modality as compared to the console modality.*

The HMD modality enables an operator to "directly look into the eyes" of an interaction partner potentially increasing mutual awareness and co-presence within the team. Operators may feel more engaged in the task, because their attention is more attracted by the remote environment, events, and people there. This might contribute to mutual allocation and understanding, and *social presence*. Hence, the second hypothesis:

Hypothesis 2 (H2): Social *presence is higher in telerobotic systems with human-like head motion when using the HMD modality as compared to the console modality.*

In addition to these fundamental design question for telepresence systems, and in particular in view of the increasing number of commercial systems for this purpose, a key questions is, if the operator

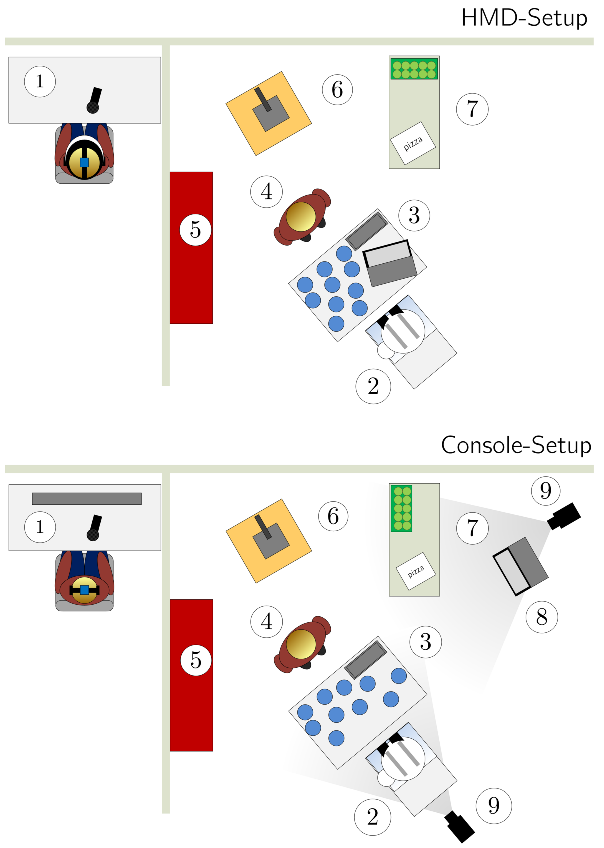

Figure 7. Task setups for the two operator modalities. In both setups, the operator is located in a separate room (left) isolating him or her from the room, where the experiment takes place (right). (1) Operator desk with the operator wearing either the HMD (top) or headphones (bottom), (2) Daryl, (3) the table on which ten items are initially located (indicated by blue circles) and a keyboard used to trigger the presentation of the next item with its target location, (4) the interaction partner, (5) a tool cabinet, (6) a small table with a dustpan, (7) drawer cabinet with a pizza box and a beverage crate on top, (8) in the Console-setup, the notebook presenting information about items and locations is placed next to one of the cameras (9), whereas in the HMD-setup it is placed on the table (3)

modality has an effect on the operator's task performance in social telepresence settings. Based on previous results [20, 18], we postulate the third hypothesis as follows:

Hypothesis 3 (H3): *An operator's task performance is lower in the HMD modality as compared to the console modality.*

4.2 Data Acquisition

In order to assess the operators' subjective levels of presence, the *Temple Presence Inventory* (TPI) [15] was combined with the *Networked Minds Social Presence Questionnaire* (NMSPQ) [4]. Both questionnaires were developed in the context of virtual reality research and have proven to yield reliable results. The TPI contains many items regarding spatial presence and realism of the telepresence experience as well as some social presence related questions including the bipolar measures proposed by [23]. To capture features related to social presence, such as mutual awareness, mutual understanding, and behavioral interdependence, the NMSPQ is used, because it provides more precise measurements regarding social interaction.

Both questionnaires use 7-point Likert scales, ranging from "Not at all" to "Very much", or "Never" to "Always". The two questionnaires were translated into German in collaboration with an English language professional preserving the meaning of all items.

Eleven additional items not contained in the original questionnaires were added. They capture the participant's media experience, quality of audiovisual information, perception of delay, and comfort while interacting through the robot. At the end of the combined questionnaire, participants were asked to provide general information such as age, gender, education, and whether they experienced motion sickness during the interaction.

4.3 Design and procedure

The study was conducted at the Social Robotics Laboratory, University of Freiburg, where a larger lab room was prepared for the sessions as the remote environment. A second room nearby served as the operator environment. Figure 7 shows the setups for the two operator modalities.

The levels of spatial and social presence felt by an operator might also vary with the task he or she has to perform through the remote embodiment. As proposed previously [1], for a specific, remote task (e.g., playing chess) "task efficiency" might be considered more relevant than achieving high levels of social presence. Task efficiency, in turn, is expected to be supported by an operator modality that achieves a higher degree of spatial presence. Therefore, we chose a cooperative task, for which the spatial layout of objects surrounding the robot is most relevant. Our task involves two people, namely the operator and an interaction partner, who collaboratively tidy up the remote environment in a master-slave manner. The operator instructs the partner how to place ten objects at three target locations. Because no arms are attached to Daryl, the operator cannot physically interact with the remote environment. This limits the operator to look at items and locations in the room and to supervise the partner's actions. Thus, the interaction partner acts as a kind of "butler", who might need to ask clarifying questions so that also some aspects of social interaction are involved.

The objects of interest are placed on a table in front of Daryl and range from everyday objects, such as an empty beer bottle and two types of hammers, to rather unusual objects, e.g., an aluminum profile and a professional measuring tool. The participants were first asked to complete this task in a face-to-face manner, before completing it again through Daryl. Although each object was assigned to one specific target location, to avoid learning effects, their placement at each location was randomized by describing it in relation to static reference objects. For example, a hammer had to be placed in the tool cabinet (see Fig. 7, (5)) to the left or right of a saw. This information was given to the operator by means of a notebook placed in the room of the robot and the butler had to press a key on a keyboard (see Fig. 7, (3)) to trigger the presentation of the next item with its target location, after the previous one was placed correctly.

4.4 Participants

For comparability with planned future studies involving the female android "Geminoid F" [1], only female participants were recruited for this study. Twenty-nine female participants were randomly assigned to one of the two conditions (mostly liberal arts students). Fourteen were assigned to the HMD-modality (age: $\mu = 23$ years,

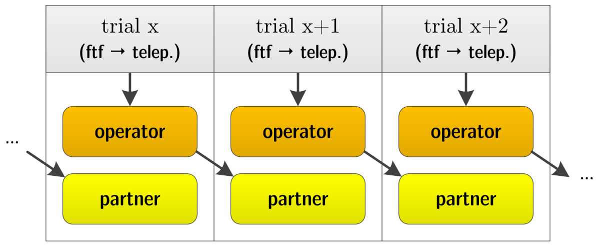

Figure 8. For each trial the previous operator served as the new interaction partner for the face-to-face (ftf) followed by the telepresence (telep.) part

$\sigma = 3.25$ years) and 15 to the Console-modality (age: $\mu = 22.5$ years, $\sigma = 2.45$ years). In each trial, the previous participant served as the new interaction partner; see Fig. 8. When there was no previous participant available, a person from the lab acted in this role.

After arrival at the lab, participants were asked to read an information sheet and to sign a consent form. They got further oral instructions about the technology and the robot before the face-to-face part of the trial started. Subsequently, the telepresence part employing one of the operator modalities began. Immediately after the second part the participants were requested to fill out the combined questionnaire.

4.5 Results

A reliability analysis using Cronbach's Alpha on the questionnaire data to verify internal consistency resulted in alpha values of above .70 for all 13 dimensions, which is considered a reliable and acceptable measure. Only the TPI dimension *Engagement* yielded a slightly lower alpha value of .69, cf. Fig. 9.

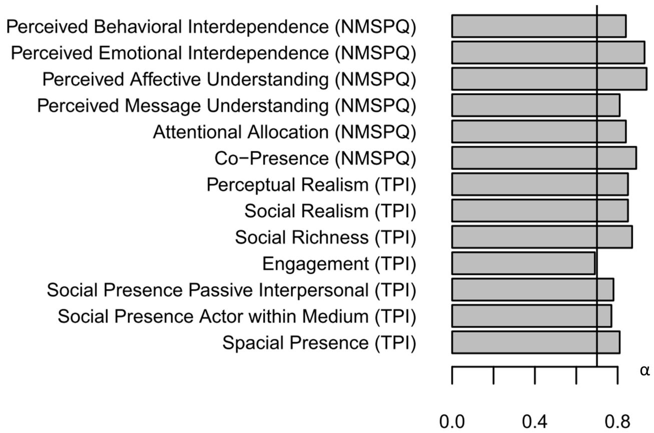

Figure 9. Cronbach's Alpha values for all presence dimensions of the combined questionnaire

To analyze whether participants of the HMD-modality experience higher levels of presence as compared to those of the Console-modality (see **H1** in Section 4.1), a between-group, one-tailed t-test is performed. As indicated in Table 1, the three dimensions *Spatial Presence*, *Social Presence Actor within Medium*, and *Engagement* show significant differences for the two operator modalities. The HMD-modality always scores significantly higher than the Console-modality, see Figure 10.

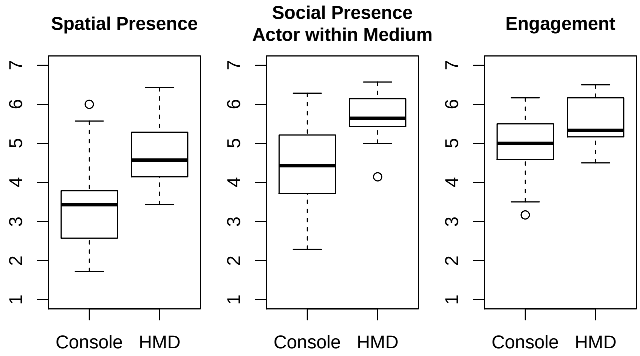

Figure 10. Box plots presenting those three dimensions of the questionnaire that resulted in significant differences between modalities

The average difference of 1.32 points for *Spatial Presence* confirms our first hypothesis that an operator's spatial presence is higher for the HMD-modality. In addition, operators of the HMD group experienced the interaction as significantly more engaging.

Regarding differences in social presence (see **H2** in Section 4.1), only the dimension *Social Presence Actor within Medium* results in a significant inter-modality difference in the expected direction. For the related dimensions *Co-Presence* and *Social Richness* significant trends can be observed indicating for both dimensions a higher average score for the HMD-modality.

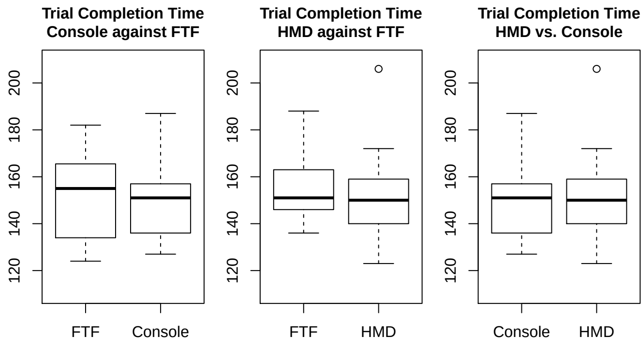

Figure 11. Box plots of the performance results comparing completion times wrt. operator modality against face-to-face (FTF) and between modalities

Regarding differences in performance between modalities (see **H3** in Section 4.1), we compare the total trial completion times [30, 20]. Within-subject, paired one-tailed t-tests are conducted with operator modality as the independent variable comparing each participant's face-to-face performance against her respective mediated performance. Comparing face-to-face against Console-modality completion time yields $p = 0.528$ ($\mu_{ftf} = 2{:}33$ min, $\sigma_{ftf} = 19.46$ s; $\mu_{console} = 2{:}30$ min, $\sigma_{console} = 18.40$ s; see Fig. 11, left) and comparing face-to-face against HMD-modality completion time yields $p = 0.522$ ($\mu_{ftf} = 2{:}35$ min, $\sigma_{ftf} = 16.20$ s; $\mu_{hmd} = 2{:}33$ min, $\sigma_{hmd} = 21.39$ s; see Fig. 11, middle). Also, a between-groups, one-tailed t-test assuming better performance results for the console modality yields no significant difference ($p = 0.644$; $\mu_{console} = 2{:}30$ min, $\sigma_{console} = 18.40$ s; $\mu_{hmd} = 2{:}33$ min, $\sigma_{hmd} = 21.39$ s; cf. Fig. 11, right). Thus, our analysis fails to indicate any significant differences in task performance and *Hypothesis 3* is not confirmed.

Table 1. Results of Welch one-tailed t-tests between the two operator modalities for all presence dimensions. (*) indicates a significant difference on the 5% level, (**) a significant difference on the 1% level, ($^{+}$) a significant trend (p ¡ 10%)

Dimension	t	df	p	Console μ	Console σ	HMD μ	HMD σ
Spatial Presence	-3.39	25.08	0.001 (**)	3.38	1.22	4.70	0.85
Social Presence Actor within Medium	-3.71	22.52	0.001 (**)	4.43	1.11	5.66	0.63
Social Presence Passive Interpersonal	-0.18	26.20	0.428	5.08	1.28	5.16	1.00
Engagement	-2.09	24.85	0.024 (*)	4.93	0.86	5.50	0.59
Social Richness	-1.62	25.78	0.059 ($^{+}$)	4.37	1.25	5.03	0.93
Social Realism	-0.87	22.85	0.197	5.07	1.56	5.48	0.91
Perceptual Realism	0.51	19.71	0.693	3.73	1.47	3.51	0.66
Co-Presence	-1.35	24.98	0.094 ($^{+}$)	5.59	1.01	6.02	0.70
Attentional Allocation	0.77	26.59	0.776	5.67	1.00	5.38	1.06
Perceived Message Understanding	-1.45	22.36	0.080 ($^{+}$)	5.66	1.07	6.12	0.60
Perceived Affective Understanding	-0.94	26.47	0.177	3.03	1.51	3.51	1.22
Perceived Emotional Interdependence	0.41	26.41	0.657	3.53	1.71	3.26	1.85
Perceived Behavioral Interdependence	-0.27	26.98	0.396	5.79	0.95	5.88	0.91

5 Discussion and conclusions

We set out to investigate the effects of two different operator modalities on levels of spatial and social presence as experienced by operators of a robotic avatar that features a head unit with a stereo camera system and three degrees of freedom. Higher levels of both spatial (H1) and social presence (H2) were expected for the HMD-modality, because in this setup the robot's head movements are intuitively controlled by an operator, who is supplied with a stereo, first-person view. In addition, previous studies indicated better task performance for the console modality in terms of total completion times and we expected to find similar effects for our collaborative tidy up task (H3).

The statistical analyses of the questionnaire data suggest that H1 can be confirmed. For H2 regarding social presence, however, only four of the subset of ten dimensions clearly related to this phenomenon show at least a significant trend in the expected direction with only one dimension being clearly significant on a 1% level. Nevertheless, these results seem to suggest that, at least for robotic avatars with a mobile head unit and a stereo camera system, an HMD-based teleoperation interface is advisable when aiming at high levels of telepresence.

Interestingly, although both questionnaires (TPI and NMSPQ) are tailored towards virtual reality setups, high reliability scores were achieved suggesting that they can also be applied to robotic telepresence systems. Furthermore, our results for the Console-modality are very similar to those reported in [14]. The Giraff tele-presence platform, however, does not allow for the use of an HMD-based interface due to the lack of a stereo camera.

The results of our study seem to be an important complement to those reported in [21]. In their study, the task itself was changed to afford high- versus low-mobility in the tele-robotic interface, but the operator modality remained unchanged. Our HMD-modality condition bears similarity to their high-mobility version of the collaborative task with the decisive difference that our participants, first, perceived the remote environment more directly in stereo audio and video and, second, could change their viewing direction intuitively. It would now be interesting to also investigate task performance effects in a small version of our scenario, or after switching to the same scenario as described in [21]. In fact, our task affords rather high levels of spatial awareness, but the social component is quite limited. In contrast to previous results [20, 18], the HMD-modality did not impede an operator's performance in this task. Complex tasks with high degrees of uncertainty as, for example, the construction task of [21] might benefit more from high presence interfaces than less complex ones [10]. Indeed, a recent empirical study comparing two interaction tasks [11] came to similar results. Perhaps, the better display quality and lower latency provided by modern HMDs significantly reduces their negative effects such as nausea and physical demand, which were previously reported as important factors [20, 18]. In fact, none of our HMD-participants reported of any such negative effects.

In a different scenario featuring higher degrees of social interaction, e.g., the one described in [2, 22], the advantage of an HMD-modality over a Console-modality has to be reevaluated. In addition to this task dependence, it remains to be explored, how a robot's degree of anthropomorphism affects an operator's felt presence, if at all. Our results are also limited to technical solutions that feature low latency audio/video links. It remains a challenging question to include latency as a factor in robotic telepresence systems. We are currently planning to conduct studies tackling these questions in cooperation with Japanese researchers [1].

In general, we are confident that this line of research helps to delineate paths towards the design of telecommunication devices that might eventually even exceed face-to-face interaction in terms of social as well as spatial presence as predicted by Hollan and Stornetta in 1992 [12].

6 Acknowledgments

We would like to thank Parastou Kohvaei and Miriam Helbig for their help with conducting the study and translating the questionnaires.

REFERENCES

[1] C. Becker-Asano, S. Gustorff, Kai O. Arras, K. Ogawa, S. Nishio, H. Ishiguro, and B. Nebel, 'Robot embodiment, operator modality, and social interaction in tele-existence: A project outline', in *Proc. of Intl. Conf. on Human-Robot Interaction (HRI2013)*, pp. 79–80, (2013).

[2] C. Becker-Asano, S. Nishio, K. Ogawa, and H. Ishiguro, 'Exploring the Uncanny Valley with Geminoid HI-1 in a real-world application', in *IADIS Intl. Conf. on Interfaces and Human Computer Interaction*, pp. 121–128, Freiburg, (2010).

[3] Christian Becker-Asano, K.O. Arras, Bernhard Nebel, and Hiroshi Ishiguro, 'The effect of anthropomorphism on social tele-embodiment', in *IROS 2012 Workshop on Human-Agent Interaction*, Vilamoura, Portugal, (2012).

[4] Frank Biocca and Chad Harms, 'Internal consistency and reliability of the networked minds social presence measure', *Seventh Annual International Workshop on Presence*, (2004).

[5] Frank Biocca, Chad Harms, and Judee K. Burgoon, 'Toward a more robust theory and measure of social presence: review and suggested criteria', *Presence: Teleoper. Virtual Environ.*, **12**(5), 456–480, (October 2003).

[6] Frank Biocca, Chad Harms, and Jenn Gregg, 'The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity', in *4th annual Intl. Workshop on Presence*, Philadelphia, PA, (2001).

[7] Charith Lasantha Fernando, Masahiro Furukawa, Tadatoshi Kurogi, Kyo Hirota, Sho Kamuro, Katsunari Sato, Kouta Minamizawa, and Susumu Tachi, 'TELESAR V: TELExistence surrogate anthropomorphic robot', in *SIGGRAPH Emerging Technologies*, p. 23. ACM, (2012).

[8] S. M. Goza, R. O. Ambrose, M. A. Diftler, and I. M. Spain, 'Telepresence control of the NASA/DARPA robonaut on a mobility platform', in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '04, pp. 623–629, New York, NY, USA, (2004). ACM.

[9] Carrie Heeter, 'Being there: The subjective experience of presence', *Presence: Teleoperators and virtual environments*, (1992).

[10] R Held and Nathaniel I. Durlach, 'Telepresence', *Presence: Teleoperators and Virtual Environments*, **1**(1), 109–112, (1992).

[11] Laura Hoffmann and Nicole C. Krämer, 'Investigating the effects of physical and virtual embodiment in task-oriented and conversational contexts', *International Journal of Human-Computer Studies*, **71**, 763 – 774, (2013).

[12] Jim Hollan and Scott Stornetta, 'Beyond being there', in *Procs. of the SIGCHI conf. on human factors in computing systems*, pp. 119–125, New York, (1992). ACM Press.

[13] Annica Kristoffersson, Silvia Coradeschi, and Amy Loutfi, 'A review of mobile robotic telepresence', *Advances in Human-Computer Interaction*, **2013**, (2013).

[14] Annica Kristoffersson, Kerstin Severinson Eklundh, and Amy Loutfi, 'Measuring the quality of interaction in mobile robotic telepresence: A pilot's perspective', *International Journal of Social Robotics*, **5**(1), 89–101, (2013).

[15] Matthew Lombard and TBDL Weinstein, 'Measuring Presence: The Temple Presence Inventory', *Proc. of Intl. Wksp. on Presence*, (2009).

[16] Palmer Luckey. Oculus Rift by Oculus VR, Inc. http://www.oculusvr.com/, April 2013.

[17] Marvin Minsky, 'Telepresence', *OMNI magazine*, (1980).

[18] M. Ephimia Morphew, Jay R. Shively, and Daniel Casey, 'Helmet-mounted displays for unmanned aerial vehicle control', in *Defense and Security*, pp. 93–103. International Society for Optics and Photonics, (2004).

[19] Shuichi Nishio, Hiroshi Ishiguro, and Norihiro Hagita, 'Geminoid: Teleoperated android of an existing person', in *Humanoid Robots: New Developments*, number June, 343–352, I-Tech Education and Publishing, Vienna, Austria, (2007).

[20] Randy Pausch, M. Anne Shackelford, and Dennis Proffitt, 'A user study comparing head-mounted and stationary displays', in *Virtual Reality, 1993. Proceedings., IEEE 1993 Symposium on Research Frontiers in*, pp. 41–45. IEEE, (1993).

[21] Irene Rae, Bilge Mutlu, and Leila Takayama, 'Bodies in motion: Mobility, presence, and task awareness in telepresence', in *Procs. of CHI2014*, (2014).

[22] Daisuke Sakamoto, Takayuki Kanda, Tetsuo Ono, Hiroshi Ishiguro, and Norihiro Hagita, 'Android as a telecommunication medium with a human-like presence', in *Proc. of the ACM/IEEE Intl. Conf. on Human-robot Interaction*, HRI '07, pp. 193–200, New York, NY, USA, (2007). ACM.

[23] J. Short, E. Williams, and B. Christie, *The social psychology of telecommunications*, volume 1, John Wiley & Sons, London, New York, Sydney, Toronto, 1 edn., 1976.

[24] Mel Slater, Martin Usoh, and Anthony Steed, 'Depth of presence in virtual environments', *Presence: Teleoperators and Virtual Environments*, 1–33, (1994).

[25] Sony. Sony HMZ-T2 product website. http://www.sony.co.uk/product/personal-3d-viewer/hmz-t2, September 2013.

[26] SPLITMEDIA. XSplit Broadcaster (http://www.xsplit.com/ [last accessed: 03/18/2013]), 2013.

[27] Jonathan Steuer, 'Defining Virtual Reality: Dimensions Determining Telepresence', *Journal of Communication*, **42**(4), 73–93, (December 1992).

[28] Trivisio Prototyping GmbH. Colibri - inertial motion tracker. http://www.trivisio.com/products/motiontracking/colibri, July 2012.

[29] Joy van Baren and Wijnand IJsselsteijn, 'Measuring presence: A guide to current measurement approaches', Technical report, Deliverable 5, IST-FET OMNIPRES Project, (2004).

[30] BG Witmer and MF Singer, 'Measuring presence in virtual environments', Technical Report October, U.S. Army Research Institute, (1994).