Towards Evaluation of Social Robotic Telepresence based on Measures of Social and Spatial Presence

Annica Kristoffersson, Silvia Coradeschi and Amy Loutfi
AASS
Örebro University
Örebro, Sweden
first.last@oru.se

Kerstin Severinson Eklundh
School of Computer Science and Communication
Royal Institute of Technology
Stockholm, Sweden
kse@csc.kth.se

ABSTRACT

In this paper we present data collected at a training session for health care personnel and alarm operators in steering a mobile social robotic telepresence robot for the first time. The purpose of the system is to be used as a communicative tool particularly when interacting with an elderly audience. The results are based on questionnaires which includes questions about experienced social and spatial presence from the Temple Presence Inventory as well as the Networked Minds Social Presence Inventory. Also investigated in this study is how intuitive the system was to use as well as how attentive the users were to what was going on in the environment. Over thirty healthcare personnel and alarm operators participate in the study and the overall results presented in the paper suggest that the two questionnaires are indeed suitable for use also in the social robotic telepresence domain for providing indications on both social and spatial presence.

Keywords

robotic telepresence, human-robot interaction, social presence

1. INTRODUCTION

Social robotic telepresence systems can be an effective tool to promote and encourage social interaction particularly between an elderly audience and their respective caregivers. In this context, a robot is present in the home of the elderly and a caregiver connects to the system to remotely visit the elderly. There are several factors which suggest that such a system could provide an added benefit for the elderly as well as for service providers. Firstly, an elderly person interacts with the robot in a natural and intuitive manner, and little additional learning is required. Secondly, the caregiver connecting to the robot from a remote location gains a greater level of control as they are allowed to move in the environment. This adds an enhanced level of safety and enables the caregiver to better assess potentially dangerous situations. Thirdly, the technology is suitable for a diverse group of elderly people ranging from the very mobile, who want to maintain contact with their local caregiver, to those who are less mobile, who want to gain a greater sense of safety. Despite this potential, the deployment of social robotic telepresence systems in elderly homes must be backed by user evaluations from first time as well as long-term experiences in order to best understand problems and user needs.

An interesting aspect about performing user evaluations with social robotic telepresence systems particularly with an elderly focus is the need to consider feedback from a variety of different users. User definition encompasses not only the elderly person but also the different types of caregivers. Primary caregivers consider professionals such as doctors, nurses, occupational therapists, audiologists and alarm operators. While secondary caregivers encompass informal caregivers such as acquaintances and families. Within the user evaluation, the nature of social contact that occur via the robot as well as the usual work behavior and processes for each user should also be considered. For example, a physician would normally contact an elderly person to mainly discuss health status while a family member may be more focused on psychological status and general well being. At the same time, a young individual who has had experience in using video conferencing systems such as Skype may find it easier and more natural to communicate via the robot than a health care professional who is unfamiliar with technological tools and does not use them in his/her normal work processes.

In this paper we focus on evaluating social robotic telepresence when used as a tool to communicate to the elderly. The user evaluations focus on the perspective of the primary caregivers. In correspondence with Mutlu and Forlizzi [3], we motivate the selection of this group by the desire to include an organizational perspective in our overall evaluations. Specifically we have chosen to focus on health care professionals assisting elderly who still live in their own homes as well as alarm operators to which elderly can connect by pushing an alarm button when in need of assistance. Presented in this paper are the results from a set of experiments whereby health care professionals and alarm operators used the system for the first time to make a remote visit to a smart-home environment with strong similarities to an elderly home. Different groups received different introduction to the system, one group received information about the

system via a instructional manual prior to using the robot, while another group received no information about the system. In the study specific focus is placed on how *present* the primary caregivers felt in the remote environment during the training session. To this effect, the results are based on questionnaires which includes questions about experienced social and spatial presence from the Temple Presence Inventory [1] as well as the Networked Minds Social Presence Inventory [2]. Also investigated in this study is how intuitive the system was to use as well as how attentive the users were to what was going on in the environment. Questions about how health care personnel perceived the different actions available for the robot during their training session were asked and field notes on how attentive they were to a number of artificially created events were recorded.

The paper is organized as follows. The experiment is described in Section 2 and it includes information about the Giraff robot, the experimental environment, test scenario, ways of observing as well as information about the chosen questionnaires. The results obtained can be found in Section 3. We discuss the results in Section 4 and present a future outlook on the works.

2. EXPERIMENTAL SETUP

In this section, a brief overview of the social robotic telepresence device called the Giraff is given, followed by a description of the experiments in which we focus on social and spatial presence. The methodology consisted of inviting the primary caregivers to a training session where they were allowed to use the system to make a remote visit to an elderly home. This visit was presented as an opportunity to train on steering and using the Giraff. The training session also served the purpose to collect data via questionnaires and observations.

This study is part of a larger project initiative called ExCITE[1] to evaluate social robotic telepresence from a longitudinal perspective at a number of test sites in Europe. The specific groups present in this study will also take part at the Swedish test sites in the ExCITE-project.

2.1 The Giraff

The Giraff provides a means for achieving remote communication between two parties. A mobile robotic base equipped with a web camera, a microphone and a screen is placed at the local user site. The local user interacts in a natural way through the robotic device with a remote peer who connects through a client interface. The client interface is on the remote user site and allows this user to teleoperate the Giraff and navigate around while speaking through a microphone and a web camera. Graphical depictions of the robotic device are given in Fig. 1 (a) and (b). As can be seen in the figure, the camera and screen are mounted on a tilt unit allowing the remote user to control the field of view. A snapshot of the client interface is shown in Fig. 1 (c). The interface is designed to be used in an intuitive manner and to not require an extensive period of learning and currently runs on all later versions of Windows. A standard computer, its pointing device, a headset and a web camera is sufficient. Navigation is done by pointing with the mouse cursor on the real-time video image received from the Giraff while pushing the left button of the mouse. The speed of the robot is automatically adjusted and dependent on how far from the Giraff's current position the user is pointing. The Giraff can move both forwards and backwards.

2.2 Subjects

The users invited to the training session were 11 people working as health care professionals in homes of elderly and 21 alarm operators responding to alarms from elderly. The health care professionals are all assisting the elderly couple at the first Swedish test site in ExCITE.

The health care professionals (group A) came to the training session in groups of two or three people and had not seen the existing instruction movie or written manual before the training session while the alarm operators (group B) came individually and had all seen the instruction movie. The average age of the users was $\mu_A = 43$, $\sigma_A = 9.04$ and $\mu_B = 42.19$, $\sigma_B = 10.34$ respectively. There were in total three men and 29 women therefore no comparison between genders is done in this experiment.

None of the users had previously used Skype or similar systems for videoconferencing where $\mu_A = 1.18$, $\sigma_A = 0.40$ and $\mu_B = 1.90$, $\sigma_B = 1.67$ on a likert scale 1-7 where 1 = not at all and 7 = to a very high degree. This implies there was a dual novelty of using videoconferencing technology along with the added mobility of the robot.

2.3 The Experiment

The training session took place in a smart-home environment at AASS[2]. The primary caregivers were placed in a room where a laptop connected to a headset and a mouse was installed. From this location, the users connected to the Giraff which was located in the smart-home environment. In the smart home environment, an actor was present sitting in a wheelchair. The actor interacted with the caregiver via the Giraff. The use of an actor was essential to script the visit and ensure that the interaction was as similar as possible between visits from the different caregivers. The script used is further detailed in Section 2.4. A graphical overview of the remotely visited room can be found in Fig. 2. In the figure, the cylinder represents the Giraff and the office chair represents the wheelchair. As can be seen in Fig. 2, the smart-home has both bedroom, living room and kitchen. It was possible to record what was happening by using the existing ten cameras that are positioned in different locations and directions in the ceiling of the home. A snapshot from one of the recordings can be seen in Fig. 3.

The same test scenario was used for all subjects but as the health care professionals arrived in groups, the people not currently driving the robot were allowed to follow the scenario beside the Giraff user. The session began with informing the participants about the computer and its connected devices and then instructing them to make a visit to a remote home with the Giraff. They were also asked to be attentive during the training session and to inform the elderly if they

[1]For more information about the project, visit http://www.excite-project.org.

[2]AASS stands for Center for Applied Autonomous Sensor Systems (AASS) at Örebro University.

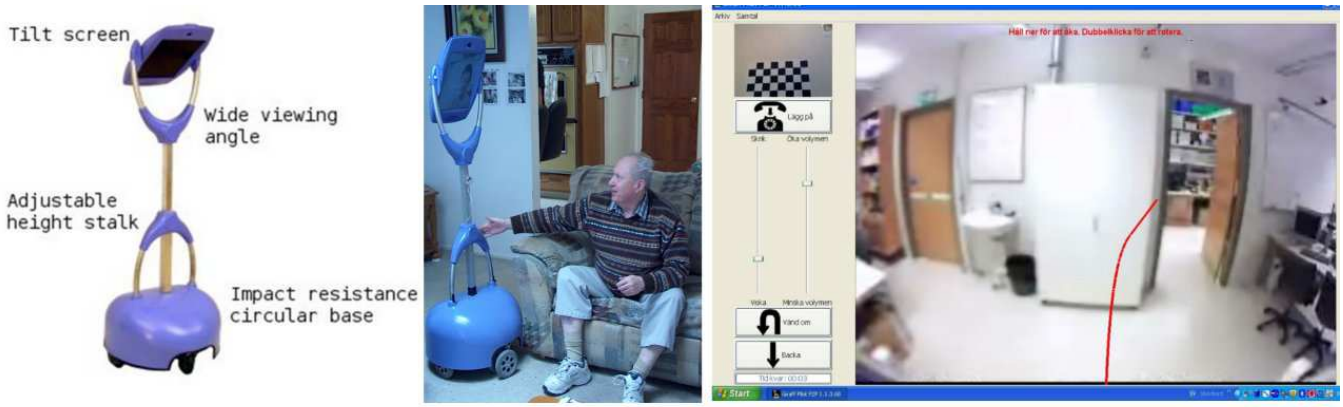


Figure 1: (a) The Giraff robot (b) A conversation between an embodied remote operator and a local elderly (c) The Giraff client interface

Table 1: Summary of response among the presence dimensions in the TPI for health care professionals A and alarm operators B

Dimension	μ_A	σ_A	α_A	μ_B	σ_B	α_B
Spatial Presence	3.07	1.62	0.75	3.05	1.80	0.76
Social Passive Interpersonal	5.36	1.45	0.84	4.98	1.46	0.84
Engagement (Mental Immersion)	5.16	1.47	0.82	4.98	1.60	0.85
Object Realism	3.84	2.30	0.65	3.61	2.12	0.43
Person Realism	4.09	2.27	0.56	3.99	2.24	0.28
Social Richness	5.00	1.31	0.91	4.63	1.48	0.88

noticed any particularities in the remote environment such as an open fridge door and the sound of an alarm. Further, they were informed that they would be asked to fill in a questionnaire after completing the training session.

2.4 Script

The procedure outlined below was used for each visit. Here we use the term *users* to denote the primary caregiver, the term *elderly* to denote the actor in the smart home and the term *operator* to denote the researcher who sat together with the user and provided technical support in case of difficulties (e.g. cannot find the docking station, knowing which buttons to press on the client side software).

1. The users are instructed by the operator to start the Giraff application, log on to the Giraff server. The 173 cm tall Giraff was positioned in its docking station that was placed against the wall beside the refrigerator.

2. Once connected remotely, the users were facing the wall as they would in a scenario visiting a real home. They were told by the operator to undock the Giraff from the docking station by pushing the buttons **Backward** and **Turn** in the Giraff Application and to locate the elderly person in the remote environment. The elderly would be found in the bed. No guidance in where to find the elderly was given except if the user drove outside the apartment with the Giraff.

3. Upon finding the elderly, the elderly person would move over to the wheelchair and ask the user to follow. The elderly went to the kitchen where the door to the fridge had been left open.

4. While in the kitchen, the elderly started a discussion about a medical issue.

5. The elderly then asked for help to find the remote control for the television which in most cases had been forgotten on the floor in between the sofa table and the television.[3]

6. After finding the remote control, an alarm rang in the bedroom which was artificially triggered. Depending on the users response, the elderly person would find an appropriate means to conclude the conversation and ask the user to return to the docking station.

7. The user returns to the docking station and disconnects from the Giraff with help from the operator if necessary.

2.5 Data Collection

The questionnaire used at the training sessions consisted of four parts: (A) Modified questions originating from the Temple Presence Inventory (TPI) [2] regarding the dimensions “Spatial Presence”, “Social Presence”, “Social Passive

[3]Minor adjustments of the remote control position were made for the 2nd and 3rd users during the training sessions with health care professionals.

Table 2: Summary of response among the presence dimensions in the Networked Minds

Dimension	μ_A	σ_A	α_A	μ_B	σ_B	α_B
Emotional Contagion	2.59	1.87	0.95	2.96	1.80	0.91
Comprehension	5.23	1.57	0.74	4.84	1.58	0.77
Behavioral Interdependence	4.6	1.77	0.80	4.70	1.81	0.91

Figure 2: A model of the visited smart home

Interpersonal", "Engagement (Mental Immersion)", "Perceptual Realism" and "Social Richness", (B) Questions regarding the perceived intuitiveness during the training session, (C) Modified questions originating from the Networked Minds Social Presence Inventory (Networked Minds) [1] regarding the dimensions "Co-presence" and "Psycho-behavioral Interaction" and (D) Questions about demographics and previous experience of videoconferencing. We further videotaped the training session at the remote location as well as made field notes based on user observations regarding their attentiveness to the open fridge door and the sound of the alarm. The modifications of the TPI mainly consisted of scaling and reformulating questions e.g. if a question involved both person and object it was given as two separate questions. For Networked Minds, the scales were not presented clearly in its documentation and one scale was chosen for all questions.

All questions in part A and C were to be answered on a likert-scale 1-7 where 1 = not at all and 7 = to a very high degree except for in the dimension Social richness where the users were asked to choose answers in opposite couples where 7 referred to feelings such as sensitive and social and 1 referred to feelings such as insensitive and non social.

In part B, the users were asked to answer how easy or difficult it was to perform different tasks with the Giraff which they had been asked to perform during their training session such as "How was it to start the Giraff application?", "How was it to do a u-turn?". It also included questions about if had been easy to find errors in the kitchen as well as to hear an alarm from the bedroom. The questions in part B were also to be answered on a likert-scale 1-7 where 1 = very difficult and 7 = very easy.

3. RESULTS

3.1 Summary of presence questionnaires

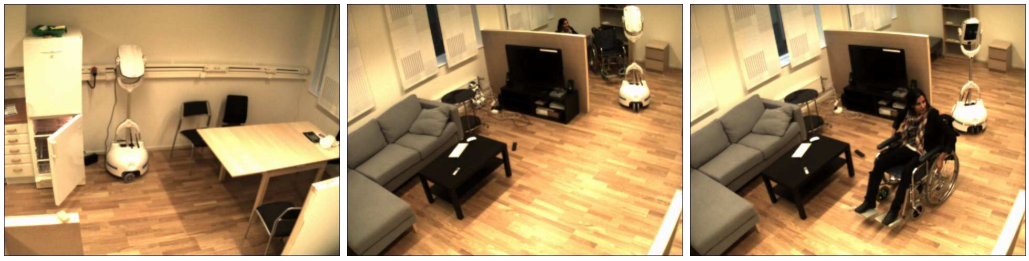
Two different questionnaires were used to collect the experienced presence from the users. These were the TPI and Networked Minds. The response on the TPI dimensions are presented in Table 1. Along with mean and standard deviations, the internal consistency Cronbach's α is presented. Generally, a Cronbach α between 0.6 and 0.7 can be considered acceptable, an α between 0.8 and 0.95 as good and an α higher than 0.95 as high. As can be seen in Table 1, the reliability on the dimensions Object and Person Realism, originating from the TPI dimension Perceptual realism is low. One reason for this is that it includes a question about smell which is not possible to sense via the Giraff. However, the experience shows that objects and person were realistic to some extent. Objects looked like($\mu_A = 5.5$,$\sigma_A = 1.51$,$\mu_B = 4.81$,$\sigma_B = 1.69$) and sounded like($\mu_A = 5.1$,$\sigma_A = 1.20$,$\mu_B = 4.76$,$\sigma_B = 1.51$) they would have in reality. The same was the case for the elderly who was perceived by the user to look like($\mu_A = 5.64$,$\sigma_A = 1.29$,$\mu_B = 5.43$,$\sigma_B = 1.33$) and sound like ($\mu_A = 5.18$,$\sigma_A = 1.33$,$\mu_B = 5.19$,$\sigma_B = 1.44$) it would have in reality. A one-way ANOVA test showed that there is no significant difference between the groups concerning realism.

Choice of scaling on some questions in the dimensions Co-presence and Psycho-behavioral Interaction - Attention originating from Networked Minds unfortunately caused a low reliability when performing the same calculations as for the TPI. Thus the response for these dimensions is not possible to analyze. A higher internal consistency scoring from acceptable to high was calculated for the Psycho-behavioral dimensions - Emotional Contagion, Comprehension and Behavioral interdependence. The results for these are presented in Table 2.

Using the Networked Minds scale also brings the opportunity to look for a symmetry among the response namely whether the user perceives that the elderly shares the same state of Social Presence. In our experiment the users responded in such a way that they perceived the elderly to share the same state of Social Presence. This was an unexpected result as for many of the users this training occasion was their first use of a video conferencing system, a robot and also the encounter with the specific elderly (actor).

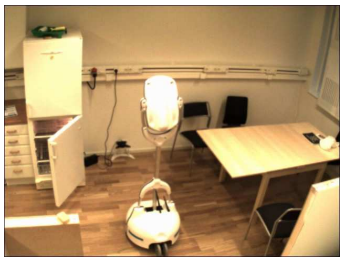
3.2 Summary of perceived intuitiveness

A summary of the perceived intuitiveness of the Giraff system among the health care professionals A and alarm operators B is given in Table 3. The attentive questions: finding errors in kitchen and hearing alarm from bedroom can be compared with the observed notifications to the elderly during the training sessions. Only 38% in total notified the elderly about the open fridge door, the number was similar for both groups. 56% of the health care personnel and 71% of the alarm operators notified the elderly about the alarm. As previously stated, only the alarm operators had seen the instruction movie before doing the training, an interesting outcome was that none of the health care personnel knew



(a) Undocking (b) Meeting the elderly (c) Follow to kitchen

(d) In the kitchen (e) Locate Remote Control (f) Farewell



(g) Docking

Figure 3: Snapshots from a Trial Run

Table 3: A summary of the response from health care professionals A and the alarm operators B for perceived intuitiveness.

How it was to	μ_A	σ_A	μ_B	σ_B
start Giraff Application	6.45	0.82	6.52	0.81
connect to the Giraff	6.36	0.81	6.38	0.86
leave docking station	5.73	1.19	6.38	1.07
make a u turn	5.64	1.12	5.71	1.27
find the person you met	5.45	1.04	5.90	0.89
stop	5.91	0.94	6.38	0.86
go backwards	5.36	1.21	5.43	1.36
follow the person you met	5.27	1.35	5.62	1.07
find errors in kitchen	4.64	1.50	5.35	1.41
help the person in the living room	5.64	1.03	5.95	0.92
hear alarm from bedroom	5.56	2.13	5.35	1.69
go back and dock the Giraff	4.64	1.57	5.86	1.11
finish the call	6.27	0.90	6.29	1.01
hear what the person said	5.20	1.75	5.52	1.29
see the person I met	6.10	0.74	5.24	1.55
keep appropriate distance to the wheelchair	4.27	1.10	4.71	1.49

where the docking station was located when asked to go back to the docking station while 43% of the alarm operators were observed knowing its location.

3.3 Summary of one-way ANOVAS

The field notes made regarding if the users heard the alarm as well as noticed the open fridge door consisted of two 1/0-variables. The variables were then used in one-way ANOVA tests which examined the correlations between the response in Questionnaire part B - Perceived intuitiveness and how attentive they actually were during the training session. The response from the healthcare personnel shows that there is no significant difference in the response with the dependent variable "observed seeing the open fridge door". The response from the alarm operators shows a significant difference on the question find errors in kitchen. There are significant differences in the response from the healthcare personnel with the dependent variable "observed hearing the alarm" regarding how intuitive it was to start the Giraff application, connect to the Giraff, undock, help out in the living room, hear the alarm from the bedroom as well as to hang up the Giraff call. The only significant difference for the alarm operators is how it was perceived to undock. A one-way ANOVA test further examined the differences in response from the groups of users. We had suspected there would be a difference in the response as users in group A had not seen the instruction movie while users in group B had seen the instruction movie. However there was only a significant difference in the response to how it was perceived to dock the Giraff. This corresponds to the observation that no user in group A knew the position of the docking station while 47 % of the users in group B knew the position when asked to go dock.

4. CONCLUSIONS

One note about the methodology used in this paper is that the use of subjective paper-and-pencil questionnaires which follow an experiment can potentially be subject to biased answers. However, the possibility of prompting the users with questions during their interaction with the elderly was rejected as it was thought this could have disturbed their interaction and influence the results. Overall, the users seem to have a realistic sense of their spatial presence meaning they understood they could not reach or touch objects and persons. The realism was also considered hight as objects and persons looked and smelled like they would in reality. Also the social, passive interpersonal dimension which was about what the users could observe considering facial expressions, body language, changes in tone of voice and way of dressing scored well. The two TPI-dimensions social richness and engagement that are measuring more social factors points to that the social presence during the training session was high. Instead looking at the Networked Minds scores the score for emotional contagion (or how the user considered it had influence on the mood of the elderly and vice versa) was low. However, one should have in mind that the situation was unnatural as it was the first time of use of the Giraff as well as the first time the user met the elderly. On the other hand, the scores for comprehension as well as dependent behavior (behavioral interdependence) were higher. Overall, it does seem as if the two questionnaires are suitable for use also in the social robotic telepresence domain.

5. ACKNOWLEDGMENTS

The authors would like to acknowledge the PEIS (physically embedded intelligent systems) project at Örebro University [4] for use of the smart-home. This work was funded by the AAL Call 2 project ExCITE - Enabling Social Interaction Through Embodiment.

6. REFERENCES

[1] F. Biocca and C. Harms. Networked minds social presence inventory (scales only version 1.2). Technical report, M.I.N.D. Labs, Michican State University, Michigan, USA, 2002.

[2] M. Lombard and T. Ditton. A literature-based presence measurement instrument the temple presence inventory (tpi) (beta). Technical report, M.I.N.D. Labs, Temple University, Pennsylvania, USA, 2004.

[3] B. Mutlu and J. Forlizzi. Robots in organizations: the

role of workflow, social, and environmental factors in human-robot interaction. In *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*, HRI '08, pages 287–294, New York, NY, USA, 2008. ACM.

[4] A. Saffiotti and M. Broxvall. PEIS ecologies: Ambient intelligence meets autonomous robotics. In *Proc of the Int Conf on Smart Objects and Ambient Intelligence (sOc-EUSAI)*, pages 275–280, 2005.