ADDRESSING THE LONG-TERM EVALUATION OF A TELEPRESENCE ROBOT FOR THE ELDERLY

Amedeo Cesta[1], Gabriella Cortellessa[1], Andrea Orlandini[1,2] and Lorenza Tiberio[1]
[1]*CNR – Consiglio Nazionale delle Ricerche, ISTC, Rome, Italy*
[2]*CNR – Consiglio Nazionale delle Ricerche, ITIA, Milan, Italy*

Keywords: Human-Robot interaction, Social robotics, Telepresence robot, Long-term evaluation.

Abstract: This paper presents aspects of an ongoing work for a *long-term* evaluation of a telepresence robot named GIRAFF, as a tool for facilitating interaction and support delivery to older people living at home. Most robotic systems are usually used for short periods of time and in laboratory settings, while this paper describes the challenges, both technological and related to the user evaluation that human-robot interaction should addressed in view of a real use of the technology for a long time span outside the laboratory. The work describes our experience in developing testing sites and in designing an evaluation plan to assess the potential of the GIRAFF platform for telepresence. We highlight open points related to the transition from a limited use in time (*short term*) to a significant period of time (*long term*). From a human-robot interaction perspective, we first introduce some results from the short term evaluation, obtained by interviewing 26 nurses as possible *clients* (people connecting to the robot) and 10 older adults as possible *end users* (people receiving visits through the robot). The paper describes then a complete evaluation plan designed for the long term assessment. From a technological point of view a set of mandatory "intelligent features" are taken into account that could enable a better real world deployment by inheriting capabilities form state-of-the-art autonomous intelligent robots.

1 INTRODUCTION

The area of social robotics is receiving increasing attention and the task of "robot as companions" is becoming a relevant reference problem also driven by funding agencies (see for example the EU within FP7 - Cognitive Systems). This attention has focalized not only research groups but also small and medium size companies that created new platforms (e.g., the NAO by Albebaran Robotics, the PR2 by Willow-Garage), the appearance of new areas with conferences and journals (e.g., the label Human-Robot Interaction spelled out in different ways), and several projects to build different type of solution of robots that both interact with humans and are connected to heterogeneous technology to build innovative solutions (e.g., (Pineau et al., 2003; Cesta et al., 2011; Saffiotti, 2009)).

In this paper we want to emphasize one aspect that deserves special attention: the study of *reaction of people who share the environments in which the robot operates* over a *long periods of time*.

The reason to underscore this aspect is because, in robotics, there is a deep-rooted tradition in developing technology usually shown in sporadic events and for short periods, i.e., for demos or live show cases, which are intended to demonstrate the "enhanced" characteristics of a prototype, making them attractive while "hiding" or at least "containing" the technical problems connected with any long term use within a comprehensive application. It is worth noting that in the case of social companions (e.g., robot assisting old people at home) one relevant issue for the robotic technology is their continuous operation, their robustness and the continuous interaction with humans over time. Which are the implications of such a continuity of use? And, in particular, how can we develop a methodology for assessing human reactions with respect to this technology? These two questions are becoming very relevant. The challenges for the Intelligent Technology and the Human Robot Interaction researchers are numerous and mainly related to two aspects: (a) in terms of *technology*, the need exists to create robust, efficient and secure solutions; (b) in terms of *users perspective*, robots must adhere to user requirements and be acceptable in the long term.

Cesta A., Cortellessa G., Orlandini A. and Tiberio L..
ADDRESSING THE LONG-TERM EVALUATION OF A TELEPRESENCE ROBOT FOR THE ELDERLY.
DOI: 10.5220/0003884306520663
In *Proceedings of the 4th International Conference on Agents and Artificial Intelligence* (SSIR-2012), pages 652-663
ISBN: 978-989-8425-95-9

More specifically, the transition from a use in the laboratory to an actual deployment into real contexts, highlights the need for a transition from short term experiments to a long term experience hence requiring a shift of attention that is highlighted in this paper. In particular we underscore how *long-term use and evaluation* are in our opinion key points to be addressed to ensure that intelligent robotic technology can actually make a leap forward and be used in real environments.

The motivation of this paper stems in our work in the ExCITE project [1]. The project focuses around an industrial mobile telepresence platform called GIRAFF and aims at fielding nine robots in three different European countries for an evaluation in real context of use. The evaluation aims at taking social and psychological factors into account and at analyzing also the emergence of "undesired behaviors" like technological weaknesses in continuous operation, human rejection, etc. Starting from the project experience, this paper analyzes and reasons about the work to both realize experiments with real users outside the laboratories and to develop a methodology for addressing long term evaluation with people. The paper also reasons on some technological challenges for telepresence robots when fielded in real environments.

The paper is organized as follows: Sections 2 and 3 enrich the context for this work by introducing respectively a telepresence robot and an overview and the ExCITE aims. Section 4 describes the design and current result of the evaluation of the GIRAFF at work. Based on these results, Section 5 discusses some technical challenges that need to be addressed for long term use. Some conclusions end the paper.

2 WHY MOBILE ROBOT TELEPRESENCE

Mobile robotic telepresence is a sub area in robotics that is progressively producing available systems and tools for real use (i.e., Anybots QB (Anybots, 2011), VGo Communications VGo (VGo, 2011), Willow Garages Texai (WillowGarage, 2011)). The different platforms combine a robotic mobile base with a remote video conference facility for the communication between distributed worker-teams, relatives or health professionals and elderly people at home or at healthcare facilities – see examples in (Lee and Takayama, 2011; Tsui et al., 2011b).

One of the most recurrently questions when a telepresence robot is proposed to a potential user is related to the difference with traditional video conferencing systems. Telepresence is a new and different manner of telecommunication because it establishes a real sense of shared space between geographically remote people duplicating the three-dimensional experience through "real human face to face interaction". Schloerb (Schloerb, 1995) describes telepresence like as the possibility for an individual to be "objectively present in a real environment that is physically separated from the person in space". In addition, one of the main feature of telepresence systems is the sense of robots presence that emerges when humans interact with and via a telepresence robotic system. Spatial presence refers to the "the sense of being physically located somewhere" as well as social presence refers to "being with others" in a mediated environment (Heeter, 1992; Ijsselsteijn et al., 2000).

One other aspect that is worth being mentioned concerns the use of mobile telepresence for old people support. One of the current limitation in proposing an autonomous robotic platform as companions at home (the basic idea proposed in (Pineau et al., 2003) for example) is not only the current costs and brittleness but also the lack of human empathy and the de-personalization of the current social robots. Indeed, some research initiatives are investigating the human-like aspects of robotics (e.g., (Kanda et al., 2005) and others) but this is seen as a medium to long term achievement. On the contrary the use of robot as a media for facilitating remote human presence in a physical space can be seen as a current intermediate solution able to maintain the "human touch" without the need of "being there".

In the telepresence research area, there have been few examples of robotic technology as tool for home care assistance for the elderly and for their interpersonal communication with caregivers, relatives and healthcare professionals. InTouch Health Company, in cooperation with Johns Hopkins University, has developed the "Physician-Robot" allowing physicians to visit more regularly their hospitalized patients. Results from an evaluation of Johns Hopkins University showed that 80% of the patients felt that Physician-Robot increases the interaction between physicians and patients (Thacker, 2005). InTouch Health has also developed the RP-7 platform for physician-patient consults allowing to remotely monitor patients offering ongoing support in terms of primary care and rehabilitation through the robot (InTouch, 2011). Some other examples are also worth being mentioned: (a) TeCaRob system provides continuous remote physical assistance performing tasks such as to transfer and move end-user, to carry out activities in end-user environment, to interact closely with end-

[1] http://www.excite-project.eu

user body and to communicate and monitor end-user (Helal and Abdulrazak, 2006); (b) the Telepresence Robot for Interpersonal Communication (TRIC) is a telepresence robot for interpersonal communication that allows older adults to remain in contact with family members and caregiver from their home environments via traditional methods (Tsai et al., 2007). Indeed, more recent products are offering a larger opportunity for addressing the problem of presence at elderly home in a more systematic way.

3 THE EXCITE PROJECT

The EXCITE project is winner of one of the calls of the EU Ambient Assisted Living Joint Programme[2], the aim of which is to *enhance the quality of life of older people and strengthen the industrial base in Europe through the use of Information and Communication Technologies (ICT)*. The specific topic of the call within which the project has been selected was focused on the idea of using innovative technology for the "Advancement of Social Interaction of Elderly People". In line with this objective, the key idea of EXCITE is to promote the use of a telepresence robot called GIRAFF to foster interaction and social participation of older people as well as to provide an easy means to possible caregivers to visit and interact with their assisted persons in their living environment (Cesta et al., 2010).

As said before, telepresence robots have been increasingly proposed to be used in workplace and Mobile Remote Presence (MRP) systems have been studied as a means to enable remote collaboration among co-workers (Lee and Takayama, 2011; Tsui et al., 2011a). Furthermore, MRP systems are also being used to provide support to elderly people. In this respect, some initial research exists which aims to understand the acceptance of older adults, their concerns and attitude toward the adoption of MRP systems (Beer and Takayama, 2011; Kristoffersson et al., 2011; Tiberio et al., 2011). The challenge of EXCITE is to carry out a set of fielded experiments with the elderly in their living environment, exploring the ability of the robot to reduce the sense of social isolation by bridging distances, facilitating interaction and communication. It is worth highlighting how this paper identifies a path to go beyond the current results available in literature. Our current work is focused on the establishment of a methodology for assessing the long term experience of MRPs on behalf of older users, thus fostering a longitudinal and ethnographic approach based on a rigorous evaluation plan. Before giving more details on this objective and the work we are performing we first introduce the MRP platform.

3.1 The GIRAFF Telepresence Platform

GIRAFF is a mobile remote presence robotic platform. It is manufactured by GIRAFF Technologies AB [3] a Swedish company founded by a team of Silicon Valley entrepreneurs. Specifically, it is a remotely controlled mobile, human-height physical avatar integrated with a videoconferencing system (including a camera, display, speaker and microphone). It is powered by motors that can propel and turn the device in any direction. An LCD panel is incorporated into the head unit. The GIRAFF's 14kg weight and integrated carrying handle allow it to easily be placed in a car or carried up stairs. A remote user can charge the GIRAFF by driving it onto the docking station. The docking station (see Figure 1, lower part) charges the batteries in under two hours. A full charge is sufficient to allow the GIRAFF to work for over two hours.

The robotic platform is accessed and controlled via a standard computer/ laptop using a software that can be downloaded on any PC from the GIRAFF web site. From a remote location the *client* (member of family or healthcare professionals) with limited prior computer training teleoperates the robotic platform while older people (*end users*) living in their own home (where the robot is placed) can receive their visit through GIRAFF.

Figure 1: Upper part: a screenshot form the GIRAFF Client. Lower part: details of the GIRAFF docking station for recharging.

[2] http://www.aal-europe.eu/

[3] http://www.giraff.org

The software operates essentially like a videoconferencing system. It is necessary to obtain authorization from individual care provider, and then connect to the person you want to visit through the GIRAFF. The client interface is designed to be as easy to use as possible (see Figure 1, upper part). A standard computer and its pointing device (such as a mouse) is sufficient. By clicking the mouse on the real time video image, the robot will follow the selected direction (red line decide the direction you want to go, green line when you actually move the robot). If you want to turn in place, a dragging motion with the mouse on the real time video image can be made. The GIRAFF automatically moves until the camera is centered at the end point of the drag. When GIRAFF is not in use its display stays in a “sleeping” position and it faces the wall to ensure individual privacy.

3.2 Key Ideas of the ExCITE Project

The GIRAFF robotic platform is a rather robust telepresence prototype, that can be used for a long-term experience in real world context. Starting from this platform, the ExCITE project aims at assessing its validity in the field of elderly support. The main innovative concepts the project is based on are the following:

– *User Centered Product Refinement*. This approach is based on the idea of obtaining users feedback during the time they use the robot and cyclically refine the prototype in order to address specific needs;

– *User tests outside labs*, rather then testing the system in laboratory setting, the robotic platform is placed in a real context of use. This approach is in line with several research that highlights how systems that work well in the lab are often less successful in real world environments (Sabanovic et al., 2006). The evaluation of robots made in a laboratory environment does not favor the emergence of robotic aid suitability to support elders who are able to stay in their own homes. For this reason an essential step is to assess the technology in the specific contexts in which the technology is supposed to be used (Hutchins, 1995);

– *Use on a time period long enough*, to allow habituation and possible rejection to appear. Indeed, interviews and survey conducted after a short period of time can be limited and can prevent other effects to emerge. On the contrary, a key aspect of relationship is that it is a persistent construct, spanning multiple interactions (Bickmore and Picard, 2005). In this light, in order to assess the human-robot interaction it is important to investigate how people interact with robots over long periods of time.

– *Analysis of cultural and societal differences*, an interesting part of our project stems in the idea of comparing the long term deployment of the telepresence platform in different countries so as to allow an analysis of cultural and societal differences over European countries.

Figure 2 gives a s brief sketch of the whole project: different GIRAFF prototypes are being deployed for long periods of time (at least three months, and possibly 1 year) in three different countries (Italy, Spain and Sweden) in real context of use. Feedback obtained from the users (both older users having the robot at home and the clients, that is people connecting and visiting the older) is used to technically improve the robot. Currently already three different versions of the prototypes have been realized which include several improvements.

During this first year of the project we have done specific work to set up the different test sites. This allowed us to obtain useful feedback and suggestions for both technological development and human-robot interaction features. In what follows, we describe our progressive work toward a long-term human-robot interaction assessment and then we discuss some lessons learned from a more focused technological point of view.

4 THE EVALUATION EFFORT

During the project activities we conceived a twofold path for evaluating the Human-Robot Interaction gathering both feedback from short interactions between potential users and the GIRAFF robot and also focusing on a long term assessment plan.

More specifically we identified two tracks for our effort:

– *Short Term Evaluation*, that consists of a collection of immediate feedback of users (both robot users and remote client users) on the GIRAFF platform, connected to different aspects of the interaction mainly related to the users opinion judgments and expectation on the GIRAFF platform and the interaction with it. The first part of the evaluation entails small periods of interaction between GIRAFF and the users to gather immediate feedback on some aspects identified as relevant.

– *Long Term Evaluation*, which relates to the study of the long-term impact of GIRAFF on elderly users using the system both to communicate with

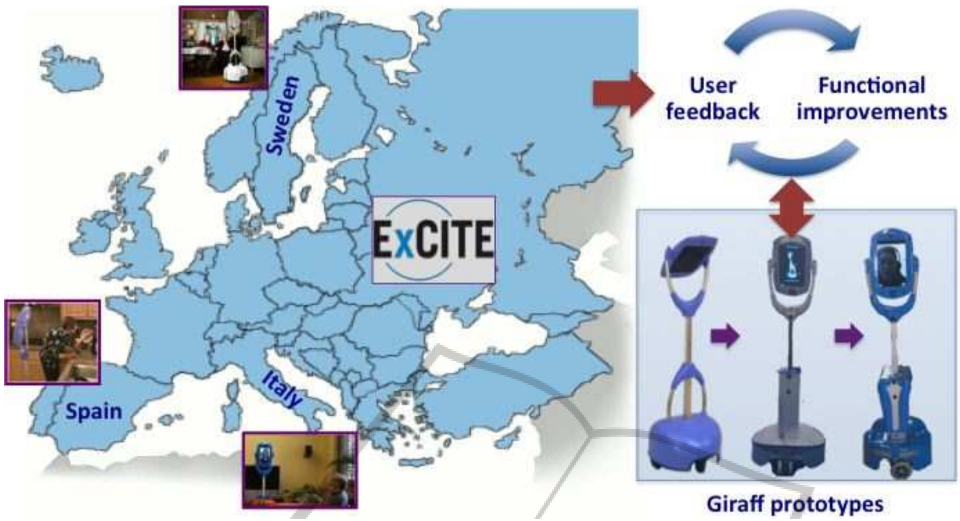


Figure 2: The ExCITE approach.

their relatives and friends and to receive visits from health care providers and in general care givers.

The short term evaluation effort, though not sufficient alone, still provides immediate feedback that can be used to quickly improve the technological development, to eventually add functionalities to the system or to simply confirm the validity of some technological choice. In addition it can give valuable guidance to the long-term assessment. For this reason we adopted a combined approach and we are currently involving participants representative of both types of users: both for the *client side* and the *end user side*.

Different initiative are ongoing to cover this twofold approach. We present here some preliminary results for the short term evaluation and our complete design for a methodology to assess the long term-impact.

4.1 Short Term Evaluation

As stated in (Beer and Takayama, 2011), before intelligent technologies would be accepted, it is important to understand their perception of the benefits, concern and adoption criteria. Those authors recruited 12 old participants giving them first-hand experience with both meeting a visitors and driving the system.

In our study, we aim at reproducing as much as possible an "ecological" setting for the experiment. To this purpose we distinguished the role of the users and recruited different participants according to their expected role. Specifically for the client side, we recruited users representative of the potential visitors of the elderly users among caregivers, nurses, health workers, etc. For the end user side we interviewed older adults living alone, or possibly receiving some kind of health care assistance.

4.1.1 Method

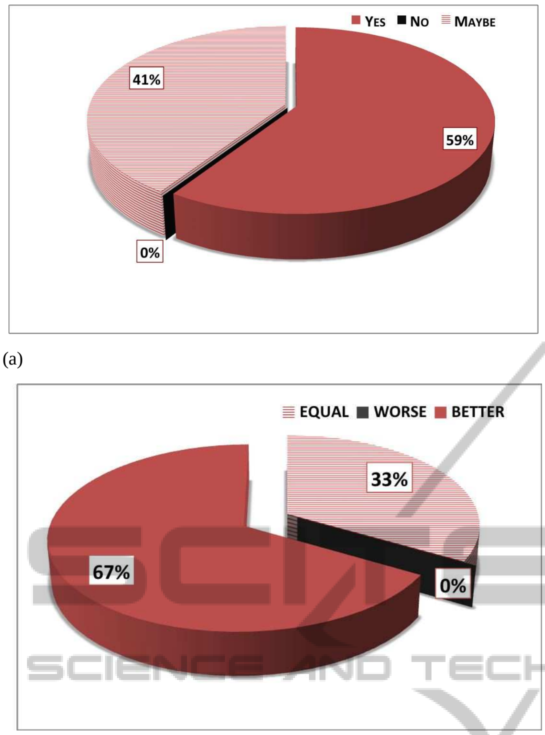
This evaluation was aimed at assessing users reaction toward the possible adoption of the GIRAFF system as a means to visit or provide some kind of service to the elderly users. Aspects investigated were *willingness to adopt the robotic solution*, *advantages* and *disadvantages* as well as possible *suggestions for improvements*.

Nurses as Clients

Participants and Procedure. Twenty six nurses form different specialist areas were recruited for this study. The sample interviewed so far is composed 11 women and 16 men with a mean age of 42 years.

The meeting entailed a tutorial presentation of 20 minutes to describe features and functionalities of the GIRAFF robot. After this tutorial, a practical session allowed nurses to operate the system and experience the different functionalities. Following the tutorial a focus group was conducted and a final questionnaire was administrated to assess possible possible applications of the telepresence robot, the perceived advantages and disadvantages of the system, the patient profile best suited to benefit from the use of an aid-based on telepresence.

Results. A first analysis of the results showed a positive reaction of the participants to the GIRAFF system (see Figure 3). In particular 59% of participants



(a)

(b)

Figure 3: General assessment of the GIRAFF system: (a) willingness to adopt it; (b) qualitative comparison with traditional teleconference system like skype.

would be willing to use GIRAFF as an aid support in his/her profession. No one opposes the use of robots.

In addition most of them judges the GIRAFF system as a better tool with respect to traditional teleconference system like Skype. Among the advantages of the system they mention the mobility of the platform that allows to have more control over the life of the end user and a better monitoring of the domestic environment.

The participants are in favor of the use of robots to train the family caregiver to small nursing tasks and constant contact with assisted people. The possibility of *continuous monitoring* (see Figure 4) of the patient at home is considered to be most useful in applications (48% of participants were in favor of this kind of application), with *support* application following at 37%, while the *companionship* functionality is less envisaged, maybe also due to the specific type of professional activities of participants [4].

The size of the robot (in particular the size of its

[4]It is worth saying that this type of "real users" preferences are exactly the type of outcomes that are looked for in order to prioritize issues for a comprehensive roadmap of robotic research.

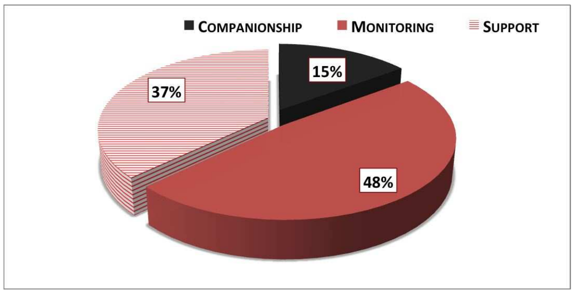


Figure 4: Favorite domains of application for the GIRAFF.

base) is one of the most significant disadvantages reported by the sample. Furthermore, a perceived disadvantage, shared by most participants is related to the *privacy* aspect, especially connected to the possible family members who share living space with the elderly patient. A difference emerges in the sample in relation to gender: women are more willing than men to use GIRAFF ($F(1,26) = 4.38, p < .05$). In addition they declared that GIRAFF would reinforce their sense of safety in their work ($F(1,26) = 4.938, p < .05$).

The focus group conducted at the end of this analysis, highlighted some aspects considered as particularly relevant for using GIRAFF in the healthcare domain for long term period. These aspects refer to the need to integrate additional functionalities to:

– support the night vision of the telepresence robot
– enable the system to divert calls when the client is not connected to the robot via a PC
– integrate the GIRAFF functionalities with other domotic devices that may be available at home
– integrate the GIRAFF functionalities with physiological sensors to monitor vital parameters of the assisted person

All these suggestions for technical improvements are currently inspiring the future modifications of the GIRAFF system in line with the user centered approach pursued in EXCITE [5].

Older Adults as End-users

Participants and Procedure. To investigate aspects connected to the end-user interaction with the telepresence system we contacted 10 older adults. Four of them were potential end users who have been asked to participate in the long-term evaluation described

[5]It is worth adding that the integration of our MRP in a sensorized home environment for continuous monitoring and health care of the old person is the goal of the recently approved *Giraff+* project that also pursues the goal of a long term evaluation.

later in this paper [6]. The remaining participants are involved in a parallel study, also connected to the project that aims to validate the GIRAFF system as a tool for providing remote rehabilitation (Tiberio et al., 2011).

The procedure followed in this qualitative research entailed an explanation of the the main idea underlying the telepresence system, showing some descriptive materials, a video of the GIRAFF system and, where possible, a practical demonstration of the system itself. The selection of the material and the modality to present the system were decided according to the time availability, and the specific situation presented in each evaluation session. Overall, we here opted for a qualitative analysis given the relatively small number of the sample.

Results. A qualitative analysis of the interview have been conducted and the most relevant feedback are here reported in terms of positive and negative aspects of the GIRAFF system.

Positive Aspects. Among the positive aspects most of the subjets reported the following:

- participants judged the visit through GIRAFF as engaging and "real"
- the robot was pleasant to see
- the ability of the robot to move in the environment was positively assessed
- users felt physically involved during the interaction
- participants think that the robot would help someone living alone at home to feel safer
- participants judged positively both the audio and the video functionalities
- participants think that interaction through the robot was spontaneous

Negative Aspects. Among the most negative aspects we mention:

- the GIRAFF system is too big and consequently may be not well integrated in a domestic environment due to its size
- the battery power may be too short
- there may be some problems due to the privacy issue
- there were some concern related to the safe movement of the robot and to its ability of obstacle avoidance
- some "intelligent features", like the autonomous recharging of the battery, are missing
- difficulty in connecting the robot to the docking station

Also this effort showed an overall positive reaction to the system, even though some improvements are desired in view of a real usage of the system. It is worth underscoring that the key point here is the fact that these qualitative data has been gathered not by interviewing generic people but "real potential users" like for example a group of caregivers and old people that can receive visit through the robot.

4.2 Long Term Evaluation Design

One of the original features of the EXCITE project consists of realizing long-term experiments involving older people using GIRAFF from their environment both to communicate with others and to receive assistance services.

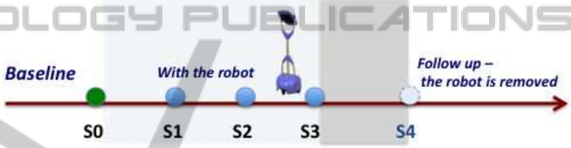


Figure 5: The Long Term Evaluation timeline.

4.2.1 Method

Figure 5 gives a general idea of the method. The evaluation entails a period of N months (with $3 \leq N \leq 12$) during which the end user will have GIRAFF at home and the clients can visit him/her through GIRAFF. Assessment happens at milestones S_i. Specifically, after an initial assessment (S0 in figure) at the beginning of the experimentation (*baseline*), the variables of interest are measured at regular intervals (S1-3) to observe changes over time. At the last month GIRAFF will be removed from the end user apartment and the same variables will be again assessed after 2 months from this removal (S4). Overall the general idea is to use a repeated measures method to see changes over time during the long term usage of the robot.

Participants and Procedure. Table 1 lists the possible types of both clients, who can be a caregiver (both formal and informal) or a relative/friend, and the end user who is an elderly living at home. The three different cases have been identified to cover different situations in which GIRAFF can be deployed. The type of material used in the long term evaluation for both the client and the end user depends upon the type of interaction for which GIRAFF is used. For

[6] Two out of these four older adults accepted to have the robot at home for the long term assessment.

Table 1: Different cases for the long term evaluation.

Type of users	
CLIENT	
Health care institution	In this case the client is a professional care giver (*formal caregiver*) who uses GIRAFF as a means to communicate at home with the elder (end user)
Family member (caregiver)	In this case a family member who takes care of the elder (*informal caregiver*) uses also GIRAFF to visit the elder
Other relatives or friends	In this case the client is a family member and/or a friend who connects through GIRAFF to the elder. The family member is not a caregiver in this last case
END USER	
Older Adult	In all cases the end user is an elderly living at home and receiving visits either form a caregiver (formal/informal) to receive assistance, or by relatives and friends just for social interaction

this reason, for each of the three mentioned situations we had developed (and or selected) a set of questionnaires (almost all validated in the three languages of the involved countries) aimed at monitoring specific variables and to be administrated at specific time both to end users and to clients.

Material. For each of the described case we prepared the material to assess the variables under study at the specified intervals.

Table 2 lists in detail the different variables and the related instruments that will be used to measure the variables over time.

Client Side. Specifically on the client side, during the initial step (S0), we will use: (a) an informed *consent form* describing the aim and procedure of the study; (b) the *socio demographic data* form to gather some relevant information on the user; (c) we developed on purpose a questionnaire aimed at assessing the client expectation on the GIRAFF's ability to ease the support (*Support Expectation*). It is worth highlighting that we developed two slightly different types of questionnaires for the formal and informal caregivers, while for the third type of client we designed a questionnaire (*Influence on Relationship Expectation*) on the expectation on GIRAFF as a means to ease and support the remote communication and consequently the social relationship.

During the following step (S1), for all three types of clients we will use: (a) questionnaires based on the SUMI inventory (Sumi, 2011) to assess the *usability* of the client software: (b) will ask participant to keep a *diary* to register the "salient" events of the visit through GIRAFF in terms of encountered problems, good features and so on.

During the subsequent step (S2), in addition to the diary that client will have to keep along the whole experience with the robot, we make a first assessment of ability of GIRAFF to ease the support (or the communication) between the client and the end user through the *Support Assessment* and *Impact on Relationship Assessment* questionnaires. In addition, during this phase we will also use the Temple Presence Inventory (Lombard et al., 2009) that is a tool to measure dimensions of (tele)presence.

At step S3 we use the Positive Affect Negative Affect Scale, PANAS, (Terracciano et al., 2003), and a final structured interview to assess the overall experience in terms of the most relevant variables considered in the study.

After two months from the removal of GIRAFF, S4 will allow assessing the impact of the GIRAFF absence through the *Support Assessment* questionnaire.

End User Side. For the end user receiving the robot we followed a similar approach, but we focused on some additional variables that is worth dwelling on. Specifically, we measure (a) the *perceived loneliness* through the UCLA Loneliness Scale (Russell et al., 1980), which was developed to assess subjective feelings of loneliness or social isolation;(b) the perceived health status through the SF12 (Ware et al., 1996); (c) the Multidimensional Scale of Perceived Social Support (Zimet et al., 1988); (d) Geriatric Depression Scale (Yesavage et al., 1983): a modified version of the Health Service Satisfaction Inventory. Finally the Almere (Heerink et al., 2010) model will allow assessing dimensions of technology acceptance.

In the table, measures highlighted in bold will ensure the repeated measures thus allowing to observe the influence of changes over time. It is worth underscoring how this evaluation methodology will allow monitoring the human-robot interaction over time, thus contributing to understand the long term impact of a fully deployed robotic solution.

With the help of the other partner we currently have all the materials in the three languages and we are starting the first phase of the long-term evaluation.

The actual implementation of this plan in three different European countries will also support a cross-cultural analysis, continuing some work started on this specific topics (Cortellessa et al., 2008).

Table 2: Long term evaluation: variables measured along the phases (S0–S4) and related material.

PHASES	S0	S1	S2	S3	S4
CLIENT SIDE					
Healthcare	Consent Form, Socio-Demographics Data Form, **Support Expectation**, **Diary**	Usability, **Diary**	**Support assessment**, Temple Presence Inventory, **Diary**,	PANAS, PIADS, Final Interview, **Diary**	**Support Assessment**
Family member caregiver	Consent Form, Socio-Demographics Data Form, **Support Expectation**(informal carer), **Diary**	Usability, **Diary**	**Support assessment** (informal carer), Temple Presence Inventory, **Diary**	PANAS, PIADS, Final Interview, **Diary**,	**Support Assessment** (informal carer)
Friends and/or relatives	Consent Form, Socio-Demographics Data Form, **Influence on Relationship Expectation**, Diary	Usability, **Diary**	**Influence on Relationship assessment**(informal carer), Temple Presence Inventory, **Diary**	PANAS, PIADS, Final Interview, **Diary**	**Influence on Relationship Assessment**
END USER SIDE					
Older adult	Consent Form, Socio-Demographics Data Form, **Loneliness (UCLA)**, **Quality of Life (SF12)**, **Multidimensional Scale of Perceived Social Support**, **Geriatric Depression Scale**, **Almere model**, **Health Service Satisfaction Inventory** (if applies)	**Loneliness (UCLA)**, **Multidimensional Scale of Perceived Social Support**, **Geriatric Depression Scale**, Attitude_Acceptance, **Health Service Satisfaction Inventory** (if applies)	Temple Presence Inventory, **Almere model**	**Loneliness (UCLA)**, **Quality of Life (SF12)**, **Multidimensional Scale of Perceived Social Support**, **Geriatric Depression Scale**, **Almere model**, PANAS, PIADS, Final Interview	**Loneliness (UCLA)**, **Quality of Life (SF12)**, **Multidimensional Scale of Perceived Social Support**, **Geriatric Depression Scale**, **Health Service Satisfaction Inventory** (if applies)

5 TECHNICAL CHALLENGES FOR LONG TERM USE

Most of the paper is dedicated to introduce methodology for evaluation. There is an additional aspect worth commenting on which is connected to the goal of performing long term evaluation in real robotic contexts of use: the technological set up for those real contexts. There is quite an amount of work to be done in creating robust contexts of use for the MRP and some general comments can be derived from such an experience even at an intermediate stage of the project.

Broadly speaking, the old population demonstrate limited acceptance of new technology, mainly due to the fear of complexity in its use. This is particularly true for robotics, as they are concerned about the system being difficult to learn (e.g., (Demiris et al., 2004)). Nevertheless, older adults have demonstrated a willingness to use such technology, when the goal is assistance for their independent living (Sharit et al., 2004; Cesta et al., 2007). In the case of MRP is also important that client users, that teleoperate the robot, are facilitated in such operation. It is important to allow them to focus attention on the assistive interaction with the old person and, thus, experiencing an increasing (and easy) projection/immersion in the environment.

It is worth observing how a Mobile Remote Presence robot is a relatively simple system. It is completely operated by the remote user and thus it is usually not endowed with autonomous behaviors. However, in our work for fielding the GIRAFF in operational contexts we gathered incremental evidence that situations exist in which some technical advancements usually connected with autonomous behavior can add robustness to the whole system in an application area where frail users are involved. For example, the robot should be able to act autonomously when the operator cannot control the telepresence robot properly, or when data transmission is lost. Additionaly, from the client users' view, a limited autonomous behavior can increase their projection capability and achieve a safe and reliable operation of the telepresence robot in a (potentially) dynamic environment (Tsai et al., 2007).

5.1 Gaining Robustness by Introducing Autonomous Capabilities

In this subsection we present some contextualized use cases for autonomous capabilities that have emerged from our experience.

Robot Encumbrance. One situation, quite common in Mediterranean countries, is due to the small size of apartments where old people live. The GIRAFF robot should be installed and operated also in domestic places of such limited size. Even skilled client users may have difficulties in controlling the robot within such small environments.

In this regard, the basic requirements needed to endow a robotic platform with autonomous behaviors are the capabilities of (a) mapping the environment, (b) localizing itself within a (partially) known environment and (c) avoiding the (possibly moving) obstacles. Usually, SLAM techniques (Smith et al., 1990) are exploited to obtain localization and mapping capabilities. Furthermore, many different solutions can be used to implement obstacle avoidance functionalities (Khatib, 1986). Additionally, the system can be equipped with some autonomous navigation ability to safely guide the telepresence robot to some specific home location requested by the operator. For instance, the operator may request the MRP to reach the kitchen in the apartment so that she can visually check the status of the stoves.

Docking. A crucial location, for the robot at home, results to be the *docking station* used to recharge batteries and park if idle. In fact, since the GIRAFF must not be left out of the docking station without control, this is the most important location and the robot should always be able to reach it. The robot should also be able to detect the status of its battery and, whether below a given threshold, automatically reach the docking station. Indeed, the return to docking is considered the GIRAFF aspect that requires synthesis of new solutions. Again here some autonomous capabilities would be of help.

Connectivity. Another important issue of a long term domestic test site is that very ofter it is not possible to rely on a continuous active WiFi internet connection (in particular, this issue has been detected as a quite critical one in both Italy and Spain). Some time, sudden communication breakdowns may leave the robot still with no active control. This is highly risky since the robot may remain stuck in an unsafe position also preventing the movements of the old person at home. In this case, whenever the data transmission is lost, the robot should automatically reach and, then, plug at the docking station. Hence, minimizing the risk to leave the GIRAFF with no charge in the middle of an apartment.

People Searching. During an emergency call [7] a client user should take control of the robot and find the elderly as soon as possible in order to check her health conditions. In such cases, the GIRAFF platform could be endowed with the capability of autonomously looking for the elder in the apartment instead of requiring the operator to (potentially) visit the whole house at random. Then, once the operator takes the control of the robot, it will find automatically the proper position to start a new dialogue, minimizing the movement of the robot in the environment and increasing the effectiveness of the interaction.

People Following. During a dialogue, the operator would be allowed to focus her attention to the old person movements/gestures by adjusting its position, instead of continuously controlling the robot. An interesting feature would be to enable GIRAFF to automatically identify the direction of the old person who is speaking and/or autonomously adapt to the position of the person (for example to maintain the person centered on the robot's camera).

Safety. A final point worth being mentioned is the paramount importance of the safety of platform basic movements to avoid any scaring movement in proximity of the old person. In this respect recent technology improvements for fault-free low level behavior like those described in (Bensalem et al., 2010) can potentially result very useful if integrated.

It is worth observing how all these identified cases can be already addressed with a smart integration of functionalities that are within the state-of-the-art of current autonomous robotics.

[7] This is a specific capability that allows to force the usual operational functionalities. It enables an authorized client to bypass the old person authorization and connect with the home environment.

6 CONCLUSIONS

This paper describes the ongoing work performed within the EXCITE project that is trying to assess a robotic telepresence system within the elderly domain. Specifically, two important aspects are presented, that can be considered as mandatory steps for both a general roadmap in robotics and our specific work.

On one hand, we have highlighted the importance of performing *ecological experiments*, i.e., which reproduce as much as possible the actual conditions of use of robotic technology, in terms for instance of real people who use it and real context of use. Although still simple in the results, analysis of the short-term evaluation provides a number of indications “from the field” that are representative of the actual users’ expectations, both in relation to the human-robot interaction and to the most urgent technological improvements essential for an effective deployment. Nurses for example, expressed a number of requests that would be important to fruitfully use the GIRAFF system as a means to support their work. At the same time, the technological tests done in real homes, highlighted technological barriers that must be necessarily overcome.

The article’s second contribution concerns our effort toward a long-term assessment. Other works in the area have highlighted this need but in this article we have proposed a rather elaborated and detailed methodology for the long-term evaluation that, to the best of our knowledge, is quite original.

In addressing these two points, there were also a number of technological challenges and requirements for “intelligent features” that the technology should incorporate and that could contribute to solve some of the open challenges in moving from a short-term demonstration to a real and continuos use.

At present, we are starting the first stages of the long-term evaluation and we do expect useful and valuable information from this effort.

ACKNOWLEDGEMENTS

They are partially supported by the EU under the Ambient Assisted Living Joint Program – EXCITE Project (AAL-2009-2-125) and under the ICT Call 7 – *Giraff+* Project (GA 288173). Authors are indebted to all the project partners for the stimulating work environment. Interactions with the colleagues from Örebro University have been particularly fruitful to synthesize the evaluation plan. Authors would also like to thank Vittoria Giuliani for comments to preliminary versions of their work.

REFERENCES

Anybots (2011). Anybots – Your Personal Avatar. http://www.anybots.com/. Last checked: March, 2011.

Beer, J. B. and Takayama, L. (2011). Mobile remote presence for older adults: Acceptance, benefits, and concerns. In *Proceedings of Human Robot Interaction: HRI 2011*, pages 19–26, Lausanne, CH.

Bensalem, S., de Silva, L., Gallien, M., Ingrand, F., and Yan, R. (2010). “Rock Solid” Software: A Verifiable and Correct-by-Construction Controller for Rover and Spacecraft Functional Levels. In *i-SAIRAS-10. Proc. of the 10th Int. Symp. on Artificial Intelligence, Robotics and Automation in Space*.

Bickmore, T. W. and Picard, R. W. (2005). Establishing and Maintaining Long-Term Human-Computer Relationships. *ACM Transactions on Computer Human Interaction*, 12:293–327.

Cesta, A., Coradeschi, S., Cortellessa, G., Gonzalez, J., Tiberio, L., and Von Rump, S. (2010). Enabling Social Interaction Through Embodiment in ExCITE. In *ForItAAL. Second Italian Forum on Ambient Assisted Living, Trento, Italy, October 5-7*.

Cesta, A., Cortellessa, G., Giuliani, M., Pecora, F., Scopelliti, M., and Tiberio, L. (2007). Psychological implications of domestic assistive technology for the elderly. *PsychNology Journal*, 5(3):229–252.

Cesta, A., Cortellessa, G., Rasconi, R., Pecora, F., Scopelliti, M., and Tiberio, L. (2011). Monitoring elderly people with the ROBOCARE Domestic Environment: Interaction synthesis and user evaluation. *Computational Intelligence*, 27(1):60–82.

Cortellessa, G., Scopelliti, M., Tiberio, L., Koch Svedberg, G., Loutfi, A., and Pecora, F. (2008). A Cross-Cultural Evaluation of Domestic Assistive Robots. In *Proceedings of AAAI Fall Symposium on AI in Eldercare: New Solutions to Old Problems*.

Demiris, G., Rantz, M., Aud, M., Marek, K., Tyrer, H., Skubic, M., and Hussam, A. (2004). Older adults attitudes towards and perceptions of smart home technologies: a pilot study. *Medical Informatics and the Internet in Medicine*, 29(2):87–94.

Heerink, M., Kröse, B. J. A., Evers, V., and Wielinga, B. J. (2010). Assessing acceptance of assistive social agent technology by older adults: the almere model. *I. J. Social Robotics*, 2(4):361–375.

Heeter, C. (1992). Being there: the subjective experience of presence. *Presence: Teleoperators and Virtual Environments*, 1(2):262–271.

Helal, A. and Abdulrazak, B. (2006). TeCaRob: Tele-Care using Telepresence and Robotic Technology for Assisting People with Special Needs. *International Journal of Assistive Robotics and Mechatronics*, 7(3):46–53.

Hutchins, E. (1995). *Cognition in the Wild*. MIT Press.

Ijsselsteijn, W. A., de Ridder, H., Freeman, J., and Avons, S. (2000). Presence: Concept, determinants and measurement. In *Proceedings of SPIE*.

InTouch (2011). InTouch Health Comprehensive Solutions. http://www.intouchhealth.com/products.html. Last checked: October, 2011.

Kanda, T., Miyashita, T., Osada, T., Haikawa, Y., and Ishiguro, H. (2005). Analysis of humanoid appearances in human-robot interaction. In *IROS 2005. IEEE/RSJ International Conference on Intelligent Robots and Systems*.

Khatib, O. (1986). Real-time obstacle avoidance for manipulators and mobile robots. *Int. J. Rob. Res.*, 5:90–98.

Kristoffersson, A., Coradeschi, S., Loutfi, A., and Severinson Eklundh, K. (2011). Towards Evaluation of Social Robotic Telepresence based on Measures of Social and Spatial Presence. In *Proceedings on HRI 2011 Workshop on Social Robotic Telepresence, Lausanne, March*, pages 43–49.

Lee, M. K. and Takayama, L. (2011). Now, I Have a Body: Uses and Social Norms for Mobile Remote Presence in the Workplace. In *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems*, CHI'11, pages 33–42, New York, NY, USA. ACM.

Lombard, M., Ditton, T., and Weinstein, L. (2009). Measuring telepresence: The temple presence inventory. In *Proceedings of the Twelfth International Workshop on Presence, Los Angeles, California (USA).*, San Francisco.

Pineau, J., Montemerlo, M., Pollack, M., Roy, N., and Thrun, S. (2003). Towards Robotic Assistants in Nursing Homes: Challenges and Results. *Robotics and Autonomous Systems*, 42(3–4):271–281.

Russell, D., Peplau, L. A., and Cutrona, C. E. (1980). The revised ucla loneliness scale: Concurrent and discriminant validity evidence. *Journal of Personality and Social Psychology*, 39:472–480.

Sabanovic, S., Michalowski, M., and Simmons, R. (2006). Robots in the wild: Observing human-robot social interaction outside the lab. In *Proceedings of the International Workshop on Advanced Motion Control*, Istanbul, Turkey. ACM.

Saffiotti, A. (2009). The Concept of Peis-Ecology: Integrating Robots in Smart Environments. *Acta Futura*, 3:35–42.

Schloerb, D. (1995). A quantitative measure of telepresence. *Presence: Teleoperators and Virtual Environments*, 4:64—80.

Sharit, J., Czaja, S. J., Perdomo, D., and Lee, C. C. (2004). A cost-benefit analysis methodology for assessing product adoption by older user populations. *Applied Ergonomics*, 35(2):81 – 92.

Smith, R., Self, M., and Cheeseman, P. (1990). *Estimating uncertain spatial relationships in robotics*, pages 167–193. Springer-Verlag New York, Inc., New York, NY, USA.

Sumi (2011). Software Usability Measurement Inventory, University College Cork. http://sumi.ucc.ie/. Last checked: November 2011.

Terracciano, A., McCrae, R. R., and Costa, P. T. (2003). Factorial and construct validity of the italian positive and negative affect schedule (panas). *European journal of psychological assessment official organ of the European Association of Psychological Assessment*, 19(2):131–141.

Thacker, P. (2005). Physician-robot makes the rounds. *Journal of Telemedicine and Telecare*, 293(2):150.

Tiberio, L., Padua, L., Pellegrino, A., Aprile, I., Cortellessa, G., and Cesta, A. (2011). Assessing the tolerance of a telepresence robot in users with Mild Cognitive Impairment – A protocol for studying users' physiological response. In *Proceedings on HRI 2011 Workshop on Social Robotic Telepresence, Lausanne, March*, pages 23–28.

Tsai, T. C., Hsu, Y. L., Ma, A. I., King, T., and Wu., C. H. (2007). Developing a telepresence robot for interpersonal communication with the elderly in a home environment. *Telemedicine and e-Health*, 13(4):407–424.

Tsui, K. M., Desai, M., Yanco, H. A., and Uhlik, C. (2011a). Exploring use cases for telepresence robots. In *Proceedings of the 6th international conference on Human-robot interaction*, HRI '11, pages 11–18, New York, NY, USA. ACM.

Tsui, K. M., Norton, A., Brooks, D., Yanco, H. A., and Kontak, D. (2011b). Designing Telepresence Robot Systems for Use by People with Special Needs. In *Proceedings of the International Symposium on Quality of Life Technologies 2011: Intelligent Systems for Better Living (held in conjunction with RESNA 2011)*.

VGo (2011). Introducing VGo. From anywhere. Go anywhere. http://vgocom.com/. Last checked: March, 2011.

Ware, J. E. J., Kosinski, M., and Keller, S. D. (1996). A 12-item short-form health survey: Construction of scales and preliminary tests of reliability and validity. *Medical Care*, 34(3).

WillowGarage (2011). Texai remote presence system. http://www.willowgarage.com/pages/texai/overview. Last checked: March, 2011.

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., and Leirer, V. O. (1983). Development and validation of a geriatric depression screening scale: a preliminary report. *Journal of Psychiatric Research*, 17(1):37–49.

Zimet, G. D., Dahlem, N. W., Zimet, S. G., and Farley, G. K. (1988). The multidimensional scale of perceived social support. *Journal of Personality Assessment*, 52(1):30–41.