Human-Robot User Studies in Eldercare: Lessons Learned

Marcel HEERINK [a, 1], Ben KRÖSE [a, b], Bob WIELINGA [b], Vanessa EVERS [b]
[a] *Institute for Information Engineering, Hogeschool van Amsterdam, The Netherlands*
[b] *University of Amsterdam, Informatics Institute, The Netherlands*

Abstract. This paper describes our experiences in collecting user data on human-robot interaction in nursing homes for the elderly. Learnings from two experiments were used to develop guidelines to support human-robot user studies with elderly users, in particular for experiments in an eldercare institution. Our experiences show that this demands a very strict organization, full cooperation by nursing personnel and extreme attention to informing the participants both before and during the experiment. Furthermore, first analysis of data from the studies suggests that social abilities in a robotic interface contribute to feeling comfortable talking to it and invite elders to be more expressive.

Keywords: Human–robot interaction, eldercare, social interaction, technology acceptance

Introduction

In the last few years, the expected growth in the elderly population and the labor shortages in the healthcare sector have inspired a number of researchers to explore the applicability of intelligent systems in general and robotic products in particular to be used in assisted-living environments [16, 19]. For robots, the functionalities are related to supporting independent living [9] by supporting basic activities (eating, bathing, toileting, getting dressed) and mobility, providing household maintenance, monitoring of those who need continuous attention and maintaining safety [1, 13]. Some studies also focus on the companionship a robot might provide [21, 19], or on the environment where they can be used and on the factors that influences user acceptance [9, 10].

Recent studies on interaction with robots stress the importance of social intelligence [8, 2, 3, 4, 10] even more so in a healthcare/eldercare environment. Our study focuses on the influence of perceived social intelligence on acceptance. A more social intelligent robot should be more effective in its communication, more pleasant to interact with and could therefore be accepted easier.

Much of the findings in recent research are based on either theoretical considerations or on small sample size experiments. We have conducted an experiment to collect a large amount of structured interaction data to investigate the influence of perceived social skills on acceptance of a robot interface by elders. The objective of this paper is to describe methods, experiences and lessons learned from these experiments.

---

[1] Corresponding author: Institute for Information Engineering, P.J. Oudweg 25, 1314 CH Almere, The Netherlands; E-mail: m.heerink@hva.nl

1. Related work

Research involving explicit tests of robots or agents with elderly users has been carried out by Wada et al. [21] and Shibata et al. [17]. These studies concerned a seal shaped robot named *Paro* that was positioned in a group of elders where they could interact with it, mainly by caressing and talking to it. The aim of this study was to observe the use of a robot in a setting described as 'robot assisted activity' and to prove that elders felt more positive after a few sessions. This was done by measuring the moods of the participants, both with a face scale form and the Profile of Mood States (POMS) questionnaire.

Another experiment that took place in an eldercare institution concerned a robot named *Pearl* as described by Pollack [15] and Pineau et al. [14]. The robot was used in open-ended interactions, delivering sweeties and used to guide elders through the building to the location of a physiotherapy department.

The experiments with Paro and Pearl both registered a high level of positive excitement on the side of elders, suggesting that a robotic aid would be accepted. However, these studies were not directed towards collecting quantitative data on acceptance of robotic technology by elders and it is not clear what aspects of the robot interface caused the users' positive attitude and whether such a robotic aid would ensure actual use on a longer term basis.

Related research in which acceptance did play a significant role is described by De Ruyter et al. [7]. It concerned a robotic interface (the iCat made by Philips), which was tested in a Wizard of Oz experiment where the robot was controlled remotely by an experimenter. The participants were asked to program a dvd-recorder and to participate in an online auction, by using the iCat interface. They were exposed to an introvert and an extravert version of the iCat interface to see whether this difference in interaction would lead to different scores in degree of acceptance. To measure acceptance, the UTAUT questionnaire (Unified Theory of Acceptance and the Use of Technology, [20]) was used. UTAUT is a model that incorporates several influences on acceptance of technology, usually in the workplace. It covers the following constructs: performance expectancy, effort expectancy, attitude toward using technology, self-efficacy, anxiety and behavioral intention to use. The aim of the study was to find out to what extent participants would use the iCat at home after having experienced it. To see whether participants would perceive the extravert iCat to be more socially intelligent, a social behavior questionnaire (SBQ) was developed and used. The results showed that the extravert iCat was indeed perceived to be more socially intelligent and that this version also was more likely to be accepted by the user.

This experiment was done in a laboratory setting, with adult, but not elderly participants. It resembles the experiment we want to do, but our focus is on elderly participants (aged 65 and older) that experience a robot in the familiar environment of their nursing home.

2. Methods and instruments

3.1. Selecting social abilities

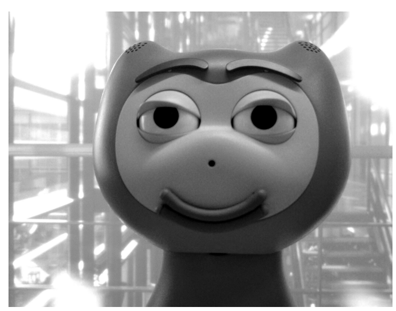
A widely used tool to evaluate social abilities is Gresham & Elliott's Social Abilities Rating System (SSRS) [12]. Although this tool usually is applied in social research, the five basic features Cooperation, Empathy, Assertion, Self-Control and Responsibility match the aspects found in Human-Robot Interaction literature on social (or sociable) robots and agents [2, 5]. Besides, these five also appear to be relevant abilities in De Ruyter et al. [7].

We decided to add Trust and Competence to this list. Not only do they appear relevant in the experiments by De Ruyter et al., they also appear as a very relevant item in research done by Shinozawa et al. [18].

This would lead to the following list of social abilities: (1) cooperate,,(2) express empathy, (3) show assertion, (4) exhibit self control, (5) show responsibility, (6) gain trust, (7) show competence. To translate these into programmable features, analyzed the list of social behaviors, set up in the experiments by De Ruyter et al. [7] and Markopoulos et al [14] and selected the following behavioral features to be programmed into our robot's character (the numbers refer to the above listed abilities):

- listening attentively, for example by looking at the participant and nodding (1, 2);
- being nice and pleasant to interact with, for example by smiling and being helpful) (1, 2, 7);
- remembering little personal details about people, for example by using their names (6, 7);
- being expressive, for example by using facial expressions (2, 3);
- admitting mistakes (5, 6).

3.2. The iCat



The particular robot we used in our experiment is the iCat ("interactive cat"), developed by Philips, also used in the experiments by De Ruyter et al. [7].The iCat is a research platform for studying social robotic user-interfaces. It is a 38 cm tall immobile robot with movable lips, eyes, eyelids and eyebrows. It is capable of displaying many different facial expressions in order to express different states of mind. There is a camera installed in the iCat's nose which can be used for different computer vision capabilities, such as recognizing objects and faces. The iCat's base contains two microphones to record the sounds it hears and a loudspeaker is built in for sound and speech. The iCat can be connected to a home network supporting the control of various in-home devices and to access the Internet.

3.3. Experimental setup

In our study, a specific interaction context was created where the iCat could be used in a Wizard of Oz fashion, which guaranteed a similar pattern for all sessions. Elders were exposed to the iCat in groups (8 participants per group). After a short introduction, the robot told them what its possibilities were: an interface to domestic applications, monitoring, companionship, information providing, agenda-keeping and memorizing medication data. After this, they were invited one by one to have a conversation with the robot, while the other group members were waiting in a different section of the room. The conversation was standardized as much as possible and we gave the participants a few simple tasks for the robot: setting an alarm, asking the way to the nearest supermarket and asking the weather forecast. While being engaged in conversation, the participants' behavior was observed and recorded by camera. The group session and the individual session were both about 5 minutes, so the maximum time spent with the robot was 10 minutes.

The experiment was executed with several groups of elders living more or less independently and elders needing daily care. Half of the groups were exposed to a more sociable version of the iCat and the other half to a less sociable one. The sociable version showed the abilities listed earlier: it was listening attentively, was more expressive, friendlier, remembered the participant's name, and admitted mistakes.

3.4. Instruments

After the conversation the participants were interviewed, using the questionnaire related to the previously mentioned UTAUT model [20], which we adapted in a few ways because of this specific context. First, UTAUT features a list of statements, related to the earlier mentioned construct, that participants had to respond to using a five point scale. Some elders that we pretested the list on could not handle this concept and responded far better to questions than to statements. Besides, also because some of them had trouble reading, it turned out to be much easier for most of them if they were asked the questions by an interviewer, who could clarify the question if necessary. Furthermore since UTAUT is developed for using technology at work, the statements/questions needed to be adapted to a domestic user environment. This meant we had to omit statements/questions that could not be adapted.

To measure the perceived sociability we also translated the Social Behavior Questionnaire (SBQ) as used by De Ruyter et al. [7] into Dutch, to be used after the UTAUT questionnaire. Besides, we added five questions to the UTAUT questionnaire on trust and sociability.

Since this experiment was about accepting a robot that is not only a peace of technology, but also a conversational partner, we wanted to complemented the technology acceptance of the UTAUT model with instruments concerning conversational acceptance. We did this by adding a question on feeling comfortable talking to a robot and by analyzing observations of conversational expressiveness by the participants.

3. Experiences and observations

We were able to do this experiment in two eldercare institutions in the Dutch cities of Almere and Lelystad, in November and December 2005. The first experiment, which was in Almere, was meant as a pilot, with a relatively small group of 28 participants. The second experiment in Lelystad featured 40 participants. In this section we will describe these experiments and briefly discuss their outcome.

3.1. First (pilot) experiment

Our pilot experiment made it very clear that we had a lot to learn. We received usable data of only 11 of the 28 participants.

First, there were organizational issues, due to our inexperience with setting up an experiment in cooperation with the nursing staff. For example, as soon as we were ready to let the participants into the testing room, there appeared to be no one waiting. We had to pick them up at their apartments ourselves, which took a lot of time, also because some participants were not dressed yet. Also, a lot of participants came during an earlier or later session than the one they were invited to.

Secondly, there were issues concerning the mental state of the participants that we took too little into account: about half of the participants had forgotten about the experiment and many of the remaining half had forgotten what it was about. Besides, some participants forgot during the experiment what it was about, just a few minutes after we had explained. Also, for many participants the questionnaire was longer than their memory of the session lasted.

Third, there were behavioral issues that we didn't take into account due to inexperience with dealing with groups of elders: some participants refused to work on the given task with the robot; they simply started a conversation with it, ignoring all instructions. Also, some participants walked away as soon as it was time for the questionnaire, because they didn't find it a necessary thing.

Finally, we found many participants thought we were trying to sell the robot, even after we explained that this was not a sales presentation. Later, we learned that the room we used was indeed often used for sales presentations. Some participants left because of this, because the robot was too expensive for them. We could not convince them that it was not our intention to sell anything.

3.2. Second experiment

Our second experiment featured 40 participants, divided into 4 groups of 8 and 2 groups of 4. Half of the participants (2 groups of 8, 1 group of 4) were exposed to the more sociable version and the other half to the less sociable one. We had asked the nursing home staff to select participants who's memory would last long enough to be able to complete the questionnaire. The experiment was prepared much more thoroughly and we asked more assistance from the caregivers at the eldercare institution. They made sure that everyone arrived on time, appropriately dressed, at the right session. We used more explicit flyers explaining the purpose and set-up of the experiment and we had extra people to keep the elders informed and entertained while they were waiting for their encounter with the iCat or in line for the questionnaire after the encounter.

After their sessions, the participants were interviewed using the UTAUT related questionnaire, expanded with the questions on perceived social abilities and conversational acceptance. We decided that it would be too much to add the SBQ.

Again, many participants had a conversation with the robot that was not only beyond the given tasks but also far beyond the presented possible functionalities of the robot. This was either because the they found it difficult to understand the limitations of the iCat's possibilities or because (perhaps due to the excitement about being observed in an experiment) they felt like making a joke to make the researchers laugh or the robot confused.

3.3. Results of the second experiment

The second experiment was more successful. Of the 40 participants, 36 sessions resulted in usable data (4 participants were omitted because they were obviously disturbed by external factors). There were no participants who walked away or refused to answer the questionnaire.

An analysis of the data showed that none of the UTAUT constructs showed a significant difference for the two conditions (sig. > 0.5 for all constructs). A significant difference between the two conditions was found on the question on feeling uncomfortable talking to a robot. Of the participants who met the more sociable version of the robot, no one reported to feel uncomfortable talking to a robot while many of the ones who met the less social condition felt more or less uncomfortable (Table 1).

Table 1. T score on feeling uncomfortable talking to a robot regarding the more and less socially communicative conditions

Condition	*N*	*Mean*	*t*	*Sig. (2-tailed)*
more social	17	1,00		
less social	19	1,53	-3,7500	0,0015

Another result concerned our observations of conversational expressiveness. We counted expressions like nodding, waving, smiling and laughing of participants during their individual expressions and found those who met the more social condition to be significantly more expressive (Table 2).

Table 2. Totals and t scores on observations concerning conversational expressiveness

	more social	*less social*	*t*	*Sig. (2-tailed)*
Positive	10,0526	7,0588	2,450	0,020
Negative	0,8947	1,2353	-0,986	0,333
All items	11,0526	8,2941	2,063	0,047

Although many participants tended to be much more enthusiastic about the possibilities of the robot, this did generally did not result in a high acceptance rate for either the more social version or the less social version. A reoccurring remark made that indicated that they would not want to use the robot if it would be available was, that they generally would not want any technology that would help them too much doing and remembering things. They would prefer to try to remember and do as much as possible without any help until there would really be no way out but to have this piece of technology.

4. Discussion and conclusions

Considering our experiences we found the following challenges ar to be faced when setting up an experiment in an eldercare environment to gather user experience data.

- A very strict organization is necessary and including the participation of caregivers who are dedicated to contribute to the success of the experiment is essential. They are the ones who know the different participants and how to ensure their participation. We needed them not only to bring the participants to the experiment and prepare them to take part in it, but also to stay with them while they were waiting.

- Elders who are suffering dementia can in many cases participate in an experiment like ours, but if they have forgotten their experiences by the time they are questioned about it, this might lead to unreliable data. If these participants are identified before the experiment, it remains possible to use other methods to gather data on their experiences. If the questionnaire is essential, like in our case, only participants that will remember their experiences long enough should be selected
- Participants have to be well informed about the purpose and procedures both before and during the experiment. They have to be aware that they are participating in an experiment and that a questionnaire is part of the protocol.
- There appears to be a limit to the length of a questionnaire elders have patience for. Of course there are individual differences, but a questionnaire containing up to 30 questions is generally about as much as elders can take.
- Many participants may express demands that are not appropriate to a robot's functionalities. This could be anticipated by having standard replies like 'I am sorry, but I am not programmed to do this'.

In the experiments, we programmed behavior into the robotic interface that concerned some abilities that could be applied in a quite brief encounter (about 5 minutes collectively and 5 minutes individually) of each participant with the iCat. Experiments concerning elders working more intensively with robots, within a setting in which these robots are for a longer time a part of their environment, might provide some relevant data that cannot be collected in a setting like ours and that might show different results.

The UTAUT model and the way we used it could be subject to discussion. It has been developed to be used for acceptance of technology in a working environment. It is not developed for elders and not for a technology that performs as a conversational partner as a robot does.

Also a Wizard of Oz setting as we used it, could also be subject to discussion (see [11] for arguments against it). One could say it is a way of cheating participants and it gives an unreal impression of the possibilities of the used technology. Nevertheless, it guarantees an experiment in which all sessions produce a very similar interaction.

Regarding physical en mental fitness, we recognized three categories of elders: (1) those who were in a good mental and physical condition and understood well what was expected from them during the experiment, (2) those who suffered physical disabilities that influenced their communication (mostly not hearing or seeing well) and (3) those who suffered mental weakness or a bad short term memory. In our second experiment, we asked the nursing home staff to select participants who's memory would last long enough to be able to complete the questionnaire. This selection contributed to the success of the experiment, but there would have been other ways to deal with participants that suffered from bad memory.

We find the data and remarks inviting to do research on the influence of perceived adaptability and to further investigate the relationship between perceived social abilities and technology acceptance in the particular context of using robotic products in an eldercare environment. In order to do this, we intend to develop a more sophisticated model of social abilities that can be applied to robots and that allows to measure the influential differences between the particular abilities.

Acknowledgments

This work was supported in part by the Hogeschool van Amsterdam and in part by the European Commission Division FP6-IST Future and Emerging Technologies under Contract FP6-002020 (Cogniron). We like to thank Hatice Çal, Wouter van Gils and Suzanne Roelofs for their contribution to the project. Furthermore we are very grateful for the hospitality and cooperation by De Archipel and De Ankerplaats. Finally, we thank Philips for enabling us to work with their iCat.

References

[1] Bickmore, T., and Picard R. W., *Towards Caring Machines*, Proceedings of CHI, Vienna, Austria, April 2004.
[2] Breazeal, C., *Towards sociable robots,* Robotics and Autonomous Systems 42.3-4: 167-175, 2003
[3] Breazeal, C., *Socially intelligent robots*, Interactions, Volume 12 Issue 2: 19 – 22, 2005
[4] Dautenhahn, K., Ogden, B., and Quick, T., *From embodied to socially embedded agents—implications for interactionaware robots*, Cognitive Systems Research (3), 2002
[5] Dautenhahn, K. *Roles and functions of robots in human society: implications from research in autism therapy*. Robotica, Volume 21 Issue 4 august 2003
[6] Dautenhahn, K. *Robots We Like to Live With?! - A Developmental Perspective on a Personalized, Life-Long Robot Companion*. Invited paper, Proc. IEEE Ro-man 2004, 13th IEEE International Workshop on Robot and Human Interactive Communication 2004.
[7] De Ruyter, B., Saini, P., Markopoulos., P. and Van Breemen, A.J.N. *Assessing the Effects of Building Social Intelligence in a Robotic Interface for the Home*. Interacting with Computers, Volume 17, Issue 5, 1 September 2005, 522-541. 2005
[8] Duffy, B.R. *Anthropomorphism and The Social Robot*. Robotics and Autonomous Systems, march (2003): 170-190.
[9] Forlizzi, J., DiSalvo, C., and Gemperle, F. *Assistive Robotics and an Ecology of Elderly Living Independently in Their Homes.* Journal of HCI Special Issue on Human-Robot Interaction, V19 N1/2, January, 2004.
[10] Forlizzi, J. *Robotic products to assist the aging population* . Interactions, volume 12 Issue 2, 2005
[11] Fraser, N. M. and Gilbert, G.N. *Simulating Speech Systems*. Computer Speech and Language 5 (1991)
[12] Gresham, F. M., and Elliot, S. N. *Social abilities rating system. Manua*. Circle Pines: American Guidance Service, 1990
[13] Mynatt, E.D., Essa, I., and Rogers, W. *Increasing the opportunities for aging in place*. Proceedings of the CUU 2000 Conference on Universal Usability New York: ACM. 65-71. 2000
[14] Pineau, J., Montemerlo, M., Pollack, M., Roy, N. and Thrun, S. *Towards robotic assistants in nursing homes: Challenges and results*. Robotics and Autonomous Systems 42 (2003): 271-281. 2003
[15] Pollack, M., Brown, L., Colbry, D., Orosz,. C., Peintner, B., Ramakrishnan, S., Engberg, S., Matthews, J., Dunbar-Jacob, J. and McCarthy, C. *Pearl: A Mobile Robotic Assistant for the Elderly*. AAAI Workshop on Automation as Eldercare. 2002.
[16] Pollack, M. *Intelligent Technology for an Aging Population: The Use of AI to Assist Elderly with Cognitive Impairment* . AI Magazine, Summer (2005)
[17] Shibata, T, Wada, K., and Tanie, K., *Statistical Analysis and Comparison of Questionnaire Resultsof Subjective Evaluations of Seal Robot in Japan and U.K.*. Proceedings of the 2003 IEEE International Conference on Robotics & Automation 2003.
[18] Shinozawa, K, Naya, F., Yamato, J. and Kogure, K., *Differences in Effect of Robot and Screen Agent Recommendations on Human Decision-Making*, IJHCS Vol 62/2, pp 267-279, 2005.
[19] Taggart, W., Turkle, S., and Kidd, C., *An interactive robot in a nursing home: Preliminary remarks.* In "Towards Social Mechanisms of Android Science", Cognitive Science Society, Stresa, Italy, July 2005.
[20] Venkatesh, V,. Morris, M. G., Davis, G. B., and Davis, F. D. *User Acceptance of Information Technology: Towards a Unified View*. MIS Quaterly, 27(3), 425-478, 2003
[21] Wada, K., Shibata, T., Saito, T. and Tanie, K., *Effects of Robot Assisted Activity to Elderly People who Stay at a Health Service Facility for the Aged.* Proceedings of the 2003 IEEE/RSJ Intl. Conference on Intelligent Robots and Systems, Las Vegas, Nevada, October 2003.