Observing conversational expressiveness of elderly users interacting with a robot and screen agent

Marcel Heerink, Ben Kröse, Vanessa Evers and Bob Wielinga

Abstract — **While expressiveness in human communication is a natural and widely observed phenomenon, in studies of humans interacting with robots and screen agents it is relatively unexplored. If it occurs however, this could mean that the artificial personality is accepted as a conversation partner by the user. An experiment with a robot and a screen agent in an eldercare institution both in a more and less expressive condition shows that it occurs: participants showed indeed more expressiveness with a more expressive robot or agent. The effect seemed to be stronger for the robot. Although the robot differed in more ways from the agent, this could be an indication of agent embodiment being a moderating factor.**

I. INTRODUCTION

THE last few years, a growing number of HRI research projects concern themselves with eldercare [1-3]. Indeed, the future of eldercare could be that of elders living longer independently, supported by technology. Robotics could be an essential part of this, also because robots and screen agents with social abilities could function both as assistive technology and social company [4]. But will elders be willing to accept all this assistive technology, especially when it concerns interactive systems that could be perceived as autonomous and intelligent such as robots and screen agents [5]? These systems differ from other technologies, because they concern technologies that are not always perceived just as such: a robot or screen agent can be (partly) perceived as a social actor and it could be that interaction with it follows the same principles as inter-human communication rather than those of human-machine interaction and this should show in the behavior of people interacting with robots or screen agents[6].

Recent research with robots in an eldercare environment shows that elders can get excited about robots and that robots can have a comforting effect that is comparable to the effect pets have [7-11]. Experiments focusing on the effects of social behavior of robots and agents, show that a more social or more caring condition does have an effect that is comparable to that of humans behaving more sociable or more caring [12-14].

The research presented here is part of a project on developing a methodology for predicting and explaining the acceptance of robots and screen agents by elderly users after a (short) test and denote the different factors that influence acceptance of robots and onscreen agents. We intend to use observation of user behavior as well as user feedback in our research and we are particularly interested in behavior that indicates acceptance of a robot or screen agent as a conversational partner.

While earlier publications on our research reported on the results of the experiments with the robotic agent in eldercare institutions[13, 15], in this paper we present, compare and discuss data from experiments with both a robotic agent and a screen agent, focusing on non verbal user behavior analysis. After a short review of related research we will describe the set up and instruments used, next we will present and interpret the results.

II. ROBOTS IN ELDERCARE

There have been several projects testing the response of elderly users towards different types of robots that could serve different purposes, varying from just being good company to physical support and giving advice. An example of a pet-like robot with no other functionalities than being good company is Paro. Since 2002 a number of experiments with this seal shaped robot have been carried out [8, 9, 16]. In early studies, it 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. More recently, research with Paro focuses on collecting physical data on elders that have been exposed to the robot to measure its effect on their wellbeing.

An example of a robot with more functionalities that was subject to experiments in an eldercare institution is Pearl [7, 17, 18]. This robot was used in open-ended interactions, delivering candies and used to guide elders through the

Manuscript received May 1, 2007. 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).

Marcel Heerink is with the Hogeschool van Amsterdam, Instituut voor Information Engineering, Almere, Netherlands (+31-36-5480771; e-mail: m.heerink@hva.nl).

Ben Kröse is with the Institute of Informatics, University of Amsterdam, Amsterdam, the Netherlands, (e-mail: Krose@science.uva.nl).

Vanessa Evers is with the Centre for Study of Language and Information, Stanford University, on leave from the Human Computer Studies Lab at the University of Amsterdam, the Netherlands (e-mail: evers@science.uva.nl).

Bob Wielinga is with the Human Computer Studies Lab at the University of Amsterdam, the Netherlands (email: wielinga@science.uva.nl)

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 robot would be accepted. In case of Paro it would merely be beneficial as a pet (a study by Libin and Cohen-Mansfield shows that a robotic pet is preferred over a plush toy cat [19]) and in case of Pearl it would be used as an actual assistant.

Research concerning experiments with screen agents for elders is reported by Bickmore and Picard [12, 20, 21]. The study focuses on the acceptance of a relational agent (a screen agent that simulates a personal interest in the user) appearing on a computer screen and functioning as a health advisor for older adults. Findings (scores on questions related to affection, trust and acceptance) indicate that the agent was accepted by the participants as a conversational partner on health and health behavior issues and rated high on trust and friendliness. It was also found to be successful as a health advisor. Other research with the same agent [22] is focused on the ability to function in long term relationships in which social abilities also appear essential. It is linked to the notion of social presence [23, 24] that people feel in interaction with systems and although it is not measured in the experiments presented in this paper, it can play a role in interpreting the responses of participants when they apparently perceive social abilities.

Research comparing robots and agents generally shows that people respond to them in a similar way. However, findings show there can be differences in trust and it might be that the embodiment of a robot is more appealing and therefore people will invest more effort in communicating with it [25, 26].

We could divide research on robot and agent acceptance into two areas: acceptance of the robot in terms of usefulness and ease of use (functional acceptance) and acceptance of the robot as a conversational partner with which a human or pet like relationship is possible (social acceptance). The experiments with Paro could be seen as a good example of research focused on social acceptance while the experiments with Pearl focused more on the acceptance of the robot regarding its functionalities. When considering behavior an indication of acceptance, in general it could be appropriate tot state we are researching the social side of acceptance.

III. Experiments

By analyzing data from two similar experiments with elderly participants, one with a robotic agents and one with a screen agent, we want to find out whether there would be differences in conversational expressiveness between (a) the robotic agent and the screen agent and (b) a more expressive and less expressive condition for each agent. For each experiment the participants where 40 elderly citizens, living in an eldercare institution. For both agents, we expected the more social condition to evoke more conversational expressiveness by the participants. Between the two types of agents we expected some difference in this effect between the robotic agent and the screen agent.

A. *Experimental design*

For both experiments a specific interaction context was created where the system (robotic agent or screen agent) was used in a Wizard of Oz fashion which made it possible to have a similar pattern for all sessions. A Wizard of Oz setup means the agent is to be perceived as being autonomous, while it is connected to a hidden operator who is controlling its behavior.

For both agents we created two different conditions: a more social one (showing more expressiveness) and a less social one. They were realized with the following behavioral features:

1) The agent in the more social condition would gaze straight at the conversation partner, the agent in the less social condition would look past the participant.
2) The agent made mistakes such as saying good morning in the afternoon or the other way round. When this would be made clear, the agent in the more social condition would apologize for the mistake, the agent in the less social condition would not.
3) The agent in the more social condition would smile when appropriate and express cheerfulness in its facial expression, the agent in the other condition did not.
4) The agent in the more social condition remembered the participant's name and use it – the agent in the less social condition did not.
5) The agent in the more social condition would support the conversation by nodding and blinking, the agent in the less social condition did not do this.
6) The agent in the more social condition was better in turn taking by waiting until the conversation partner finished speaking, the agent in the less social condition was less polite.

The experiment with the robotic agent was executed a few months before the experiment with the screen agent. The participants were principally the same for both experiments.

B. *Procedure*

Participants were elderly people (13 male, 27 female) between 65 and 96 years old, living in eldercare institutions in the cities of Almere and Lelystad, in the Netherlands. They were divided among the two conditions as equally as possible (the social condition featured one more male and one less female).

The participants were first exposed to the agent in groups (two groups of 8 participants and one group of 4 participants for each condition). After a short introduction by one of the researchers the robot told them what its possibilities were: an interface to domestic applications, monitoring, companionship, information providing, agenda-keeping and memorizing medication times and dates. They were told that for today's experiment, the agent was only programmed to perform three tasks: setting an alarm, give directions to the nearest supermarket and giving the weather forecast for tomorrow. The experimenter subsequently demonstrated how to have a conversation with the robot in which it performed these tasks.

After this group session, the participants 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 asked the participants to have the robot perform the three simple tasks.

While being engaged in conversation, the participants' behavior was observed by a researcher 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 for each participant.

C. Behavior analysis methodology

Although participants were observed during the experiment, we based our analysis on observations of the video's afterwards.

During the analysis verbal, and non-verbal forms of conversational expressiveness were counted for each participant such as greeting the agent (with or without words) nodding or shaking the head, smiling, looking surprised or irritated (frowning), and moving towards or away from the robot. This list of items considering conversational expressiveness was generated by listing classical feedback gestures (see [27-31]) without categorizing them to specific communicative functions.

We added the behavior of verbal greeting to it, because we considered this also a sign of relational feedback.

The observers where not made aware of the different conditions of the agents. We had three observers for each video and if their counts differed we mediated the numbers.

Where observer differences occurred, we counted a particular behavior if it was recognized similarly by two of the three observers.

IV. Agents

The robotic agent we used in our experiment is the iCat ("interactive cat"), developed by Philips. 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 to display different facial expressions to simulate emotional behavior. 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 output. We used the iCat with a female voice, simply because this was the voice that was the one three pretest subjects felt most comfortable with.

Fig. 1 The iCat

The screen agent was developed for our tests by students of the Instituut voor Information Engineering in Almere, Netherlands. It features a female, humanoid character (because the robotic agent was also given a female identity) being able to move the same facial parts as the iCat.

It was used on a 17 inch lcd screen in combination with a webcam (attached to the screen), a microphone and two speakers. We named it 'Annie'

Fig. 2 Screen agent Annie

V. Results

The different types of expressive behavior by participants during their interaction with the agent were counted for each participant, added for each condition and analyzed to measure conversational expressiveness.

TABLE I
TWO CONDITIONS OF THE ROBOTIC AGENT - TOTAL COUNTS AND T SCORES ON CONVERSATIONAL EXPRESSIVENESS

Totals for all participants:	more social (N=17)	less social (N=19)	t	Sig. (2-tailed)
Nodding head	66	54	0,3946	0,6958
Shaking head	16	15	-0,1261	0,9005
non-verbal greeting	2	0	1,4552	0,1628
'don't know' gesture	3	0	1,0000	0,3306
move away	0	4	-1,7253	0,1037
approach robot	17	7	1,6170	0,1152
Smile	42	30	1,1380	0,2631
Laugh	26	9	1,8477	0,0775
Surprise	2	0	1,4552	0,1628
Show irritation (frown)	1	2	-0,5045	0,6189
Verbal greeting	36	21	1,9004	0,0672

Tables I and II show that there is a pattern of more conversational expressiveness for the more social condition: the participants, with a higher frequency for almost all types of behavior.

TABLE II
TWO CONDITIONS OF THE SCREEN AGENT - TOTAL COUNTS AND T SCORES ON CONVERSATIONAL EXPRESSIVENESS

Totals for all participants:	more social (N=17)	less social (N=19)	t	Sig. (2-tailed)
Nodding head	83	50	2,526	0,016
Shaking head	9	10	0,015	0,988
non-verbal greeting	3	2	0,603	0,551
'don't know' gesture	2	10	-1,576	0,124
move away	5	6	-0,137	0,892
approach robot	6	17	-2,251	0,031
Smile	47	32	1,915	0,064
Laugh	16	17	0,157	0,876
Surprise	1	4	-1,309	0,199
Show irritation (frown)	11	11	0,293	0,771
Verbal greeting	23	21	0,822	0,417

We categorized the behavior types by them being positive or negative and looked at the total number of times a type of behavior (positive/negative) occurred for the different conditions. We considered the behaviors shaking head, move away and show irritation negative and all others positive.

Table III shows that for both agents there is a difference between the more social and less social condition both in total expressions and in the total amount of expressions that can be categorized as positive expressions, but it is stronger for the robotic agent than for the screen agent.

TABLE III
TWO CONDITIONS OF THE AGENTS - T SCORES ON CATEGORIZED BEHAVIORAL OBSERVATIONS

Agent:	Robotic agent		Screen agent		Combined	
	t	Sig.	t	Sig.	t	Sig.
Positive	2,450	0,020	2,017	0,052	2,902	0,005
Negative	-0,986	0,333	0,457	0,650	-0,471	0,639
All items	2,063	0,047	2,024	0,051	2,607	0,011

VI. Discussion and Conclusions

For both agents there is a clear pattern of more conversational expressiveness, a higher frequency of non-verbal behaviors, of participants that were in conversation with an agent in a more expressive condition. However, the effect is much stronger for the robotic agent. This could indicate that embodiment has a modifying influence on this effect. Nevertheless, we have to be careful drawing conclusions on this influence, since the agent and robot differed in more than one way – besides the screen agent being two-dimensional and the screen agent three dimensional, the screen agent had a more humanoid appearance than the robot. Further study, comparing agents with just one difference in embodiment, seems essential. Using a so called ‘virtual iCat’, a screen version of the robotic agent, would be an appropriate next step.

Another reason for not jumping to conclusions would be the sequential setup of the two experiments and the use of roughly the same participants. This means the participants that were exposed to the screen agent had already met with the robotic agent a few months before and could simply have been less enthusiastic. We have to note though, that they did not report to be less enthusiastic about it and questionnaire scores did certainly not suggest it [15] (in fact, the score on intention to use was slightly higher for the screen agent).

When we look at the differences for particular behaviors of the two agents, it seems there is a third fact we have to consider. The difference is in fact merely caused by the times peopled laughed and by their attempts to approach the robot. Actually, the screen agent in the non social condition was approached more often in the non social condition. Perhaps this has to do with the nature of its embodiment in combination with the less social version being less communicative and therefore less clear. People (not necessarily elderly people) could tend to move closer to the screen if they find it harder to understand the ‘person’ they’re talking to, but approach less for this particular reason if they communicate with a three dimensional entity.

Another item for further research could be the question whether conversational expressions occurred as in response to the same expressions by the agent (a smile in response to a smile, a frown in response to a frown). In that case we would be speaking of imitative behavior. This would be the occurrence of a well known phenomenon in psychology called the chameleon effect [32]. It concerns imitative behavior between humans, which seems to occur naturally unless two people do not like each other. The occurrence of this behavior could even very well be interpreted as a sign of acceptance [33]. But during behavior analysis the observers just counted the number of behaviors, without looking at the behavior of the agent that evoked it - the camera was always directed towards the participant. In future research this possibility of imitative behavior could be something to observe, also when comparing agents with different embodiments, since it could add interesting viewpoints to HRI theory on this aspect [34, 35].

Earlier in this paper we introduced the notion of social presence. This could be a crucial factor here, since difference in embodiment could very well be related to a different sense of social presence [23, 24] and this could even explain the differences between the conditions and (even more) the difference between two dimensional and three dimensional embodiment.

Finally, we like to view the results of our behavior analysis within the context of developing a methodology to predict and explain acceptance. In technology acceptance methodology traditionally measurement is done with verbal user feedback [36]. Our results demonstrate how a behavior analyses can be a complementary instrument in such methodology. This is especially the case when dealing with elderly participants, because many of them are difficult to interview, either because of difficulty remembering what happened moments ago or because of difficulty focusing on anything longer than few minutes. To our experience [37], a questionnaire with 27 items is about the maximum.

ACKNOWLEDGMENT

We like to thank Nicolai Sequeira Geraldes, Tormo Antonides, 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] M. Pollack, "Intelligent Technology for an Aging Population: The Use of AI to Assist Elders with Cognitive Impairment," *AI Magazine*, vol. Summer, 2005.

[2] W. Taggart, S. Turkle, and C. Kidd, "An interactive robot in a nursing home: Preliminary remarks." In *Towards Social Mechanisms of Android Science*, vol. July. Stresa, Italy: Cognitive Science Society, 2005.

[3] E. D. Mynatt, I. Essa, and W. Rogers, "Increasing the opportunities for aging in place." In Proceedings of theCUU 2000 Conference on Universal Usability, New York, 2000.

[4] J. Forlizzi, "Robotic products to assist the aging population," *Interactions, Volume 12 Issue 2*, vol. march, 2005.

[5] J. Forlizzi, DiSalvo, C. and Gemperle, F., "Assistive Robotics and an Ecology of Elders Living Independently in Their Homes," *Journal of HCI Special Issue on Human-Robot Interaction*, vol. January, 2004.

[6] C. Bartneck, Forlizzi, J., "A design-centred framework for social human-robot interaction." In Proceedings IEEE International Workshop on Robot and Human Interactive Communication, 2005.

[7] J. Pineau, M. Montemerlo, M. Pollack, N. Roy, and S. Thrun, "Towards robotic assistants in nursing homes: Challenges and results," *Robotics and Autonomous Systems*, vol. 42, pp. 271-281, 2003.

[8] K. Wada, T. Shibata, T. Saito, and K. Tanie, "Psychological and Social Effects of Robot Assisted Activity to Elderly People who Stay at a Health Service Facility for the Aged." In Proceedings of the 2003 IEEE/RSJ Intl. Conference on Intelligent Robots and Systems, Las Vegas, Nevada, USA, 2003.

[9] K. a. S. Wada, T., "Robot Therapy in a Care House - Results of Case Studies." In Proceedings RO-MAN, 2006.

[10] A. Beck, Edwards, N., Friedman, B., Khan, P., "Robotic Pets and the Elderly," 2003.

[11] A. Beck and A. Katcher, *Between Pets and People*. West Lafayette: Purdue University Press, 1996.

[12] T. Bickmore, Picard R. W., "Towards Caring Machines." In Proceedings of CHI, Vienna, 2004.

[13] M. Heerink, B. J. A. Kröse, B. J. Wielinga, and V. Evers, "The Influence of a Robot's Social Abilities on Acceptance by Elderly Users." In Proceedings RO-MAN, Hertfordshire, 2006.

[14] D. Heylen, I. van Es, A. Nijholt, and B. van Dijk, "Experimenting with the Gaze of a Conversational Agent." In Proceedings International CLASS Workshop on Natural, Intelligent and Effective Interaction in Multimodal Dialogue Systems, 2002.

[15] M. Heerink, B. J. A. Kröse, B. J. Wielinga, and V. Evers, "Studying the acceptance of a robotic agent by elderly users," *International Journal of Assistive Robotics and Mechatronics*, vol. 7, 2006.

[16] T. Shibata, K. Wada, and K. Tanie, "Statistical Analysis and Comparison of Questionnaire Results of Subjective Evaluations of Seal Robot in Japan and U.K." In Proceedings of the 2003 IEEE International Conference on Robotics & Automation, 2003.

[17] M. Montemerlo, J. Pineau, N. Roy, S. Thrun, and V. Verma, "Experiences with a Mobile Robotic Guide for the Elderly." In Proceedings of the AAAI National Conference on Artificial Intelligence, 2002.

[18] M. E. Pollack, L. Brown, D. Colbry, C. Orosz, B. Peintner, S. Ramakrishnan, S. Engberg, J. T. Matthews, J. Dunbar-Jacob, C. E. McCarthy, S. Thrun, M. Montemerlo, J. Pineau, and N. Roy, "Pearl: A Mobile Robotic Assistant for the Elderly," in *AAAI Workshop on Automation as Eldercare*, 2002.

[19] A. Libin and C.-M. J., "Therapeutic robocat for nursing home residents with dementia: preliminary inquiry," *American Journal of Alzheimer's Disease and Other Dementias*, vol. 19, 2004.

[20] T. Bickmore, Caruso, L., and Clough-Gorr, K., "Acceptance and Usability of a Relational Agent Interface by Urban Older Adults." In Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, 2005.

[21] T. W. Bickmore, L. Caruso, K. Clough-Gorr, and T. Heeren, "'It's just like you talk to a friend' relational agents for older adults," *Interacting with Computers*, vol. 17, pp. 711-735, 2005.

[22] T. Bickmore and P. R. W., "Establishing and Maintaining Long-Term Human-Computer Relationships," *ACM Transactions on Computer-Human Interaction*, vol. 12, 2005.

[23] M. Lombard and T. B. Ditton, "At the heart of it all: The concept of presence," *Journal of Computer-Mediated-Communication*, vol. 3, 1997.

[24] K. M. Lee and C. Nass, "Designing social presence of social actors in human computer interaction." In Proceedings of the SIGCHI conference on Human factors in computing systems, 2003.

[25] C. Bartneck, J. Reichenbach, and A. J. N. van Breemen, "In your face, robot! The influence of a character's embodiment on how users perceive its emotional expressions, Design and

Emotion." In Proceedings of the Design and Emotion 2004 Conference, Ankara, Turkey, 2004.
[26] K. Shinozawa, F. Naya, J. Yamato, and K. Kogure, "Differences in Effect of Robot and Screen Agent Recommendations on Human Decision-Making," *IJHCS*, vol. 62, pp. 267-279, 2005.
[27] L. Axelrod and K. Hone, "Identifying Affectemes: Transcribing Conversational Behaviour." In Proceedings of the Symposium on Conversational Informatics for Supporting Social Intelligence and Interaction, 2005.
[28] L. A. Cerrato, "Comparison between Feedback Strategies in Human-to-Human and Human-Machine Communication." In Proceedings of ICSLP, 2002.
[29] D. Heylen, A. Nijholt, and D. Reidsma, "Determining what people feel and think when interacting with humans and machines: notes on corpus collection and annotation." In Proceedings 1st California Conference on Recent Advances in Engineering Mechanics, 2006.
[30] C. L. Sidner and C. Lee, "Engagement During Dialogues with Robots." In AAAI Spring Symposia, 2005.
[31] K. R. Scherer, "Toward a dynamic theory of emotion: The component process model of affective states.," *Geneva Studies in Emotion and Communication*, 1987.
[32] T. L. Chartrand and J. A. Bargh, "The chameleon effect: the perception-behavior link and social interaction," *Journal of Personality and Social Psycholology*, vol. 76, pp. 893-910, 1999.
[33] P. H. Kahn, H. Ishiguro, B. Friedman, and T. Kanda, "What is a Human? – Toward Psychological Benchmarks in the Field of Human-Robot Interaction." In Proceedings RO-MAN, 2006.
[34] K. Dautenhahn, "Trying to imitate - a step towards releasing robots from social isolation." In Proceedings From Perception to Action Conference, 1994.
[35] K. Dautenhahn and C. L. Nehaniv, "An agent-based perspective on imitation." In *Imitation in Animals and Artifacts*, K. Dautenhahn and C. L. Nehaniv, Eds.: MIT Press, 2002, pp. 1-40.
[36] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User Acceptance of Information Technology: Toward a Unified View," *MIS Quarterly*, vol. 27, pp. 425-478, 2003.
[37] M. Heerink, B. J. A. Kröse, B. J. Wielinga, and V. Evers, "Human-Robot User Studies in Eldercare: Lessons Learned." In Proceedings ICOST, Belfast, 2006.