How Pepper Moves to Motivate

The effect of robot body language on the user experience of young adults and elderly

Marianne Bossema
Leiden University, The Netherlands
Digital Life Center, University of Applied Sciences Amsterdam, The Netherlands

In this pilot study different body language styles of a Pepper robot were tested with young adults and elderly in the context of a movement exercising game. It was found that an encouraging style had a positive effect on perceived enjoyment of participants, their trust and the intention to use the robot. Test results also suggest that an encouraging body language style has a stronger positive effect on elderly than on young adults regarding their intention to use the robot in the future. Observations showed differences in human-robot interaction concerning vocal communication. For elderly to interact with social robots on their own, speech recognition and production should be optimized and adapted to the target group. In regard to a robot's expressive style, it is suggested to combine persuasiveness with adaptation. In order to prevent 'overtraining', the robot should be able to recognize the users state and adapt exercises, feedback and expressive style.

Introduction

The proportion of elderly in society is growing rapidly, which presents public health challenges and increases the demand for health care services and caregivers (He, Goodkind & Kowal, 2015). Facing those challenges, healthcare technology can support medical staff, caregivers and elderly, and promote wellbeing and independence.

Among the many healthcare services that are needed, robot technology can address and contribute to social interaction and companionship. Next to physical exercising, social interaction and companionship showed to have a positive impact on general mental and physical wellbeing (Moak & Agrawal, 2009), and social robots can support and enhance the activity of older adults (Fasola & Matarić, 2013).

The background of this study is a care home in The Netherlands, with a rehabilitation department for elderly with non-congenital brain injury. The healthcare institution purchased a Softbank Robotics Pepper robot to explore and develop the usefulness and added value of a social robot in their homes. The robot is currently being used to inform and notify residents on activities and therapy sessions. The objective is to leave caregivers more time for other tasks, and empower residents by allowing them to be better prepared for appointments, and giving them more control over their daily lives. In the near future the robot will also be used to practice therapy movements with patients, so that the treatments can be more effective and facilitate faster recovery.

However, the success and effectiveness of such an application will largely depend on the robot's capability to be persuasive. Several studies have shown that emotions can be motivating, and can guide perception and attention (Izard, 1993) (Picard 2003). Evolution has shaped the human body and senses around social interaction; we have developed mirror neurons that improve our awareness of others (Rizzolatti & Craighero, 2004), that enable us to share feelings (Iacoboni, 2009) and learn through imitation (Frith & Frith, 2007). In line with this, emotions and moods are found to be contagious in human-human interaction (Neumann & Strack, 2000) as well as in human-robot interaction (Xu et al., 2014). Mood contagion is called one of the psychological mechanisms by which leaders influence followers by expressing positive emotions, and nonverbal behavior has been found to play a vital role in persuasiveness (Bono & Ilies, 2006).

A considerable amount of research has explored the role of non-verbal cues in human-robot interaction (e.g., Sidner et al., 2004; Breazeal et al., 2005), as well as the effect of these cues on persuasiveness (Chidambaram et al., 2012). Nevertheless, the impact of these dynamics on older adults has not yet been investigated. Seniors are a specific target group; they are often less familiar with new technologies and are more reluctant or even afraid to use it. Furthermore, older adults generally have less control over their lives and may experience persuasive behavior differently than younger adults. The presented study investigates the impact of a human-like robot's body language on its effectiveness to motivate and activate elderly.

The remainder of the paper will describe related works and discussions that lead to the research question of this study. Subsequently, the approach will be explained, covering hypotheses, measurements and procedures. Next, we will reports on the results, formulate a conclusion and describe topics for discussion and further work.

Related work and research question

To investigate the process of acceptance of intelligent systems for elderly care, Heerink et al. (2010) developed and tested an adaptation and theoretical extension of the Unified Theory of Acceptance and Use of Technology. The study presents interrelated factors, showing that Social Presence -the feeling of being in the company of a social entity (Lombard & Ditton, 1997) - has an effect on Perceived Enjoyment and ultimately on the decision to actually use a system. Results show that Perceived Enjoyment influences Ease of Use rather than being an aftereffect; when seniors experience pleasure in interacting with a robot, it has a direct impact on the Intention of Use. Trust and sense of presence increase when a robot has more social abilities, which makes it relevant to explore the robot's behavioural factors that may have an impact on older adults.

A study by (Chidambaram et al., 2012) focuses on persuasiveness: How can a social assistive robot encourage and convince people using nonverbal behaviour? The team investigated how bodily cues (e.g., proximity, gaze, gestures) and vocal cues (e.g., tone and expressions) can reinforce mutual understanding between people and a human-like robot. Different sets of nonverbal behaviours were compared: only bodily cues, only vocal cues, combined cues and no cues. Results showed that the presence of nonverbal behaviours significantly improves agreement between people and robot. Bodily cues alone turned out to be more effective than vocal cues alone, which has implications for the design of persuasive behaviour of human-like robots. Comparing high-level sets of nonverbal behaviours (e.g., only bodily cues vs. only vocal cues) has its limitations and Chidambaram et al.(2012) recommend further research on the effects of individual cues.

Nomura et al. (2010) studied age dependencies in identification of affective body expressions by a robot. They found differences in emotion recognition, perceived attributes like motion, speed and magnitude, and body parts paid attention to. For example when the robot expressed sadness, older participants more often looked at lower body parts that were not informative, which prevented them from identifying the emotion. Although the study was performed in specific settings, it showed that even small differences in a robot's non-verbal behaviour influence perception and interaction. It was also found that age dependencies should be considered and investigated when designing human robot interactions for older adults.

A robotic elderly assistant can promote a healthier lifestyle and increase well-being by providing concrete recommendations. This should be done with care; elderly may be less familiar with new technology, and may find themselves in a situation where they cannot take care of themselves and need help to accomplish everyday tasks. The verbalizations of the robot should consider possible feelings of embarrassment and frustration, and be as polite as possible while keeping a certain persuasiveness. In this perspective, Hammer et al. (2016) investigated linguistic variations of the robot's recommendations. They found that some verbalizations were considered to be polite and persuasive, and could be generally used, while others were found impolite but convincing and might be used in more critical situations. Some results may be influenced by limitations of the elderly; a follow-up study will select recommendations based on participant's preferences and abilities.

Xu et al. (2014) addressed the question whether robot mood displayed during an imitation game can be perceived by participants and produce contagion effects. They found that participants were able to correctly identify the mood expressed by the robot, and showed evidence for mood contagion. In addition, they found that an expected effect of negative mood on task performance could be replicated: in the negative mood condition participants performed better on difficult tasks than in the positive mood condition.

Research question

When adapted to specific age groups, the behavior of a robot can be more persuasive and effective. This makes it relevant to compare the impact of different expressive styles on young adults and the elderly. Affective bodily cues have an impact on the expressive style, and are in particular relevant for a Pepper robot that lacks facial features. In this paper we focus on the question:

"What is the effect of a human-like robot's body language on perceived enjoyment and trust of elderly and young adults, and their intention to use the system in the future?"

Approach

Within 6 weeks (fall 2019), a between-groups experiment is carried out, where 8 students and 8 elderly are asked to interact with the Pepper robot. Students were recruited from different faculties at the university and were all under 30 years old, 5 males and 3 females. Elderly were recruited in an elderly home in Amsterdam, were all above 65, 2 males and 6 females. The tests with students have been taken at the university, the tests with seniors were executed with 4 participants in the elderly home and 4 at the university (Fig.1).

The interaction is a simple movement exercise executed in a wizard-of-oz setup; because the robot is not programmed to react on voice input and validate the movements, the dialogue is controlled via a keyboard. Participants are invited to practice arm movements together with the robot. For each correct movement, a star is added on the robot's tablet display. Every 5 stars, participants are asked if they want to stop or continue, until a maximum of 15 stars is reached. This star reward system is to designed to give positive feedback and motivate participants, not to measure task performance.

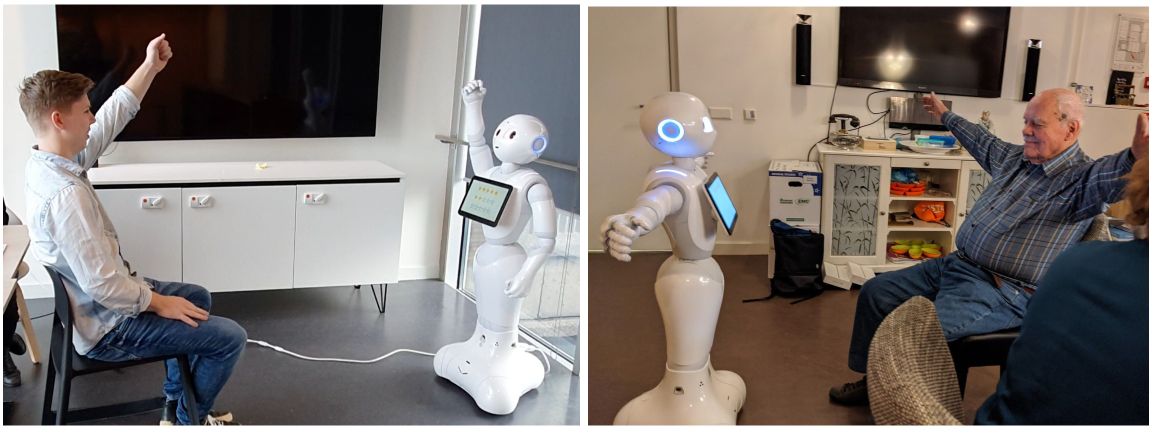

***Figure 1**. Experiments carried out with University students, and with seniors in an elderly home.*

Scenarios

To compare the effects of body language, two scenarios are used that differ only in body language style. To compare differences between age groups, these scenarios are be presented randomly to young adults and seniors.

In scenario A, the robot will display a positive, encouraging style, in scenario B it will display a neutral, formal style. Scenarios A and B will further be called 'Encouraging' and 'Neutral'.

In the opening and closing phases of scenario Encouraging, the robot greets the user by waving, thus applying robot etiquette as an emblems gesture (Ekman, 2004), while expressing an enthusiastic style. The design of the positive expression style is further based on a study by Marmpena et al. (2018), who classified and validated animations from the libraries for Pepper, pre-designed by professional animators from Aldebaran/SoftBank Robotics. They explored how humans perceive their affect content, and annotated the animations with reliable labels of valence and arousal. In scenario A we used 5 animations that were classified and perceived as positive/calm and positive/excited. In scenario Neutral the robot expresses a formal style, without any movements aimed at affective communication. A description of the scenarios (in Dutch) can be found in Scenario-Encouraging and Scenario-Neutral

Questionnaires

Before and after the interaction, participants answered likert-scale questions. The questions asked beforehand were about attitude and expectations of the robot. The questions asked afterwards concerned the user experience. Questions on the attitude towards robot technology were asked both before and after the interaction. An overview of questions (in Dutch) plus references can be found in questionnaires. Arguments for the selected questions are described under 'Measurements'.

Hypotheses

When comparing a positive and a neutral robot body language style we hypothesize that:

1. A positive body language style increases perceived enjoyment.
2. A positive body language style increases trust.
3. A positive body language style increases the intention to use the robot.
4. A positive body language style increases the score for attitude towards robot technology.
5. There are significant differences between young adults and elderly

Measurements

We performed a two-factor factorial experiment, where the factors are 1) robot body language style and 2) age group. Robot body language style is defined in the two scenarios 'Encouraging' and

‘Neutral’. Age is defined in the two groups ‘Young adults’ and ‘Elderly’. Table 1 illustrates the distribution of 16 participants in groups and subgroups.

		Scenarios		
		Encouraging	Neutral	Total
Age groups	Young adults	4	4	8
	Elderly	4	4	8
	Total	8	8	16

Table 1. *Distribution of participants in groups and subgroups.*

The dependent variables attitude (ATT), perceived enjoyment (PENJ) and trust (TRUST) are based on the suggested toolkit for measuring robot acceptance described by Heerink et al. (2009). Heerink’s questions on Intention to Use (ITU) related to short term use, e.g. “I think I’ll use the robot during the next few days”. Because that was not relevant in our case, those questions were replaced by the similarly named scale from Moon & Kim (2000).

For hypotheses testing we measured PENJ, TRUST and ITU. To explore a possible change in attitude towards robot technology, ATT is measured before and after the interaction. Questions on the robot’s body language (RBL) were added to measure specific effects on the user experience. Dependent variables are listed in Table 2.

Code	**Description**
ATT-1	Positive or negative feelings about the appliance of the technology - before the interaction.
PENJ	Feelings of joy or pleasure associated by the user with the use of the system.
RBL	Feelings of joy or pleasure associated by the user with the robot’s body language
TRUST	The belief that the system performs with personal integrity and reliability.
ITU	The outspoken intention to use the system over a longer period in time.
ATT-2	Positive or negative feelings about the appliance of the technology - after the interaction

Table 2. *Abbreviation and descriptions of dependent variables*

Results

For each dependent variable, percentages of the highest possible score were calculated. Table 3 gives an overview of the means, standard deviations and results of Shapiro-Wilk tests. In the case were the data was not normally distributed, a Mann-Whitney test was executed instead of an independent sample t-tests.

				ATT-1	PENJ	RBL	TRUST	ITU	ATT-2
Scenarios	Encouraging	n=8	Mean	0.75	0.83	0.77	0.76	0.73	0.77
			SD	0.12	0.12	0.12	0.17	0.19	0.13
			Normal distr.	yes	yes	yes	no	yes	yes
	Neutral	n=8	Mean	0.66	0.70	0.78	0.51	0.53	0.68
			SD	0.18	0.13	0.06	0.20	0.13	0.17
			Normal distr.	yes	yes	yes	yes	yes	yes
	ind. sample t-test	df=14	t	1.2	2.19	-0.27		2.45	1.11
			p	0.25	0.046*	0.79		0.03*	0.28
	Mann-whitney		Statistic				54.5		
			p				0.02*		
Age groups	Elderly	n=8	Mean	0.69	0.76	0.81	0.70	0.70	0.74
			SD	0.16	0.14	0.09	0.21	0.18	0.14
			Normal distr.	yes	yes	yes	yes	yes	yes
	Young adults	n=8	Mean	0.72	0.78	0.74	0.58	0.55	0.71
			SD	0.16	0.14	0.09	0.23	0.18	0.18
			Normal distr.	yes	yes	yes	yes	yes	yes
	ind. sample t-test	df=14	t	-0.33	-0.21	1.41	1.16	1.7	0.42
			p	0.75	0.84	0.18	0.27	0.11	0.68
	Mann-whitney		Statistic				41.5		
			p				0.33		

Table 3. *Statistical descriptions and test results per dependent variable in different conditions. Values are calculated based on percentage scores. *results are statically significant: p < 0.05*

Outcomes (Tabe 3) indicate significant differences for PENJ, TRUST and ITU between the two scenarios. No significant differences were found for the other dependent variables in the same conditions, and when comparing the dependent variables in the age group conditions.

To find out whether the interaction between both factors is statistically significant we performed a two-way ANOVA test. We compared scores of PENJ and ITU using the scenarios and age groups as fixed factors. Because the data for TRUST was not normally distributed, that variable was excluded from the test.

ANOVA - ITU - PENJ

Tested factor	p-value	
	ITU	PENJ
Scenarios	0.02*	0.06
Age groups	0.07	0.82
Scenarios * Age groups	0.48	0.6

Table 4. *Results of a two-way ANOVA test. *p-value indicates statistical significance.*

Outcomes (Table 4) tell us that scenario conditions have an effect on the intention to use the robot, but the interaction between both factors is not statistically significant. However, a post hoc test reveals a distinctive difference between elderly that were presented the Encouraging scenario and young adults that were presented the Neutral scenario (Table 5). This suggests that an Encouraging scenario has a stronger positive effect on elderly than on young adults regarding their intention to use the robot in the future.

Post Hoc Comparisons - age groups * scenarios - ITU			
		t	p Tukey
Encouraging, elderly	Neutral, elderly	2.4	0.13
	Encouraging, students	1.94	0.26
	Neutral, students	3.31	0.03*
Neutral, elderly	Encouraging, students	-0.47	0.97
	Neutral, students	0.91	0.8
Encouraging, students	Neutral, students	1.38	0.54

Table 5. *Results of a Tukey HSD post hoc test. * p-value indicates statistical significance.*

Observations

In general, participants were curious about the robot and enjoyed meeting Pepper and communicating with it. They often reacted enthusiastic to the robot's gaze: the way the robot looked around and followed them with its head. Some people said the eyes made them feel as if the robot was really looking at them.

Some of the participants (n=5) were shown both scenarios after the experiment, and were asked which one would be more motivating. They all opted for Scenario-Encouraging, and found this to be more 'activating'.

Participation in the movement exercises worked well for everyone, and participants seemed to enjoy being 'guided' by the robot. People tended to either follow the robot or stop exercising, and the majority collected more than 10 stars, regardless of the scenario. The reactions to the movement exercises were generally positive.

Some remarks that were made: "The robot can't see how you feel while doing exercises. You may be forcing something, especially when the robot is persuasive." "Getting custom feedback while exercising is important for motivation." "There should be an option for imitating the movements next to mirroring them, for short-term memory training." and "It is hard to hear the robot because the sound is coming from its ears."

Especially for the elderly, it was difficult to hear and understand what the robot was saying. This had an impact on the human-robot dialogue and made it much harder for those people to interact with the robot on their own. After a first tests with students at the university, the speaking speed had been adjusted, but later experiments showed that it was still difficult to listen to spoken text. Often, elderly participants did not respond spontaneously to the robot asked and then questions had to be repeated.

Conclusion

From the hypotheses described before, hypothesis one, two and three are found to be true; when the robot displayed a positive, encouraging style, this had a positive effect on perceived enjoyment of participants, their trust and the intention to use the robot. Concerning hypothesis four, the attitude towards robot technology before and after interaction did not significantly change.

Results of an ANOVA and additional post hoc test suggest that the Encouraging scenario had a stronger positive effect on elderly than on young adults in relation to their intention to use the robot in the future. As to the other dependent variables, we found no significant differences between young adults and elderly.

Discussion and further work

In this study, the effect of a robot's body language was investigated in the context of movement exercises. Because the two scenarios only differed in robot body language, this was a discerning factor that could explain effects. However, questions on the robot's body language might be misunderstood by participants, and interpreted as questions on the movements related to the exercises. This especially applies to participants presented Scenario-Neutral, because in that scenario the robot's body language was more subtle . This can explain the fact that no significant differences were found for the dependent factor of Robot Body Language. In further work, scenarios should be tested and compared in different contexts.
Filling in questionnaires can be difficult for the target group; some elderly struggled reading and remembering the questions. A number of participants at the elderly home therefore completed the questions with assistance of a caretaker. This had an impact on the number of experiments that were carried out.
In regard to questions on the attitude of participants towards robot technology, those were rather general and thereby unclear. For example, some participants asked how to interpret the question "I think it's a good idea to use the robot" and argued that their answer would depend on use case and conditions. In further research these questions could be complemented by questions on the attitude towards robot technology based on concrete examples.
In general, the method and approach chosen for this pilot study can be discussed. Especially with elderly participants we had conversations about robot technology, about usefulness, usability and ethical aspects of social robots. Preparing and structuring these conversations into interviews, and asking input and feedback on alternative implementations, would have contributed strongly to the study.
Although measurements of the dependent variables for young adults and elderly did not show significant differences, observations showed that the way of interacting with a robot is different. This was related in particular to vocal communication with the robot. We suggest that for elderly to interact with social robots on their own, vocal communication (dialogues, sound, prosodic cues) should be optimized and adapted to the target group. This is a point of attention for further work.
As to the robot's expressive style, it is important to combine persuasiveness with adaptation. As appeared from observations of the movement exercises, people tend to either follow the robot or stop exercising. In order to prevent 'overtraining', the robot should be able to recognize the users state and adapt exercises, feedback and expressive style.

References

Bono, J. E., & Ilies, R. (2006). Charisma, positive emotions and mood contagion. *The Leadership Quarterly*, *17*(4), 317-334.

Breazeal, C., Kidd, C. D., Thomaz, A. L., Hoffman, G., & Berlin, M. (2005, August). Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. In *2005 IEEE/RSJ international conference on intelligent robots and systems* (pp. 708-713). IEEE.

Chidambaram, V., Chiang, Y. H., & Mutlu, B. (2012, March). Designing persuasive robots: how robots might persuade people using vocal and nonverbal cues. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction* (pp. 293-300). ACM.

Ekman, P. (2004). Emotional and conversational nonverbal signals. In *Language, knowledge, and representation* (pp. 39-50). Springer, Dordrecht.

Fasola, J., & Matarić, M. J. (2013). A socially assistive robot exercise coach for the elderly. *Journal of Human-Robot Interaction*, *2*(2), 3-32.

Frith, C. D., & Frith, U. (2007). Social cognition in humans. *Current Biology*, *17*(16), R724-R732.

Hammer, S., Lugrin, B., Bogomolov, S., Janowski, K., & André, E. (2016, April). Investigating politeness strategies and their persuasiveness for a robotic elderly assistant. In *International Conference on Persuasive Technology* (pp. 315-326). Springer, Cham.

He, W., Goodkind, D., & Kowal, P. R. (2016). An aging world: 2015.

Heerink, M., Krose, B., Evers, V., & Wielinga, B. (2009, September). Measuring acceptance of an assistive social robot: a suggested toolkit. In *RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication* (pp. 528-533). IEEE.

Heerink, M., Kröse, B., Evers, V., & Wielinga, B. (2010). Assessing acceptance of assistive social agent technology by older adults: the almere model. *International journal of social robotics*, *2*(4), 361-375.

Izard, C. E. (1993). Four systems for emotion activation: Cognitive and noncognitive processes. *Psychological Review, 100*(1), 68-90.

Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of computer-mediated communication*, *3*(2), JCMC321.

Marmpena, M., Lim, A., & Dahl, T. S. (2018). How does the robot feel? Perception of valence and arousal in emotional body language. *Paladyn, Journal of Behavioral Robotics*, *9*(1), 168-182.

Moak, Z. B., & Agrawal, A. (2009). The association between perceived interpersonal social support and physical and mental health: results from the National Epidemiological Survey on Alcohol and Related Conditions. *Journal of public health*, *32*(2), 191-201.

Moon, J. W., & Kim, Y. G. (2000). Beyond the Usefulness and Ease of Use: Extending the TAM for a World-Wide-Web Context. In *International conference of the KMIS, 2000* (pp. 379-386). The Korea Society of Management Information Systems.

Neumann, R., & Strack, F. (2000). " Mood contagion": the automatic transfer of mood between persons. *Journal of personality and social psychology*, *79*(2), 211.

Nomura, T., & Nakao, A. (2010). Comparison on identification of affective body motions by robots between elder people and university students: A case study in japan. *International Journal of Social Robotics*, *2*(2), 147-157.

Picard, R. W. (2003). Affective computing: challenges. *International Journal of Human-Computer Studies*, *59*(1-2), 55-64.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.*, *27*, 169-192.

Sidner, C. L., Kidd, C. D., Lee, C., & Lesh, N. (2004, January). Where to look: a study of human-robot engagement. In *Proceedings of the 9th international conference on Intelligent user interfaces* (pp. 78-84). ACM.

Xu, J., Broekens, J., Hindriks, K., & Neerincx, M. A. (2014, May). Robot mood is contagious: effects of robot body language in the imitation game. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems* (pp. 973-980). International Foundation for Autonomous Agents and Multiagent Systems.