A Robotic Shower System - Evaluation of multimodal Human-Robot Interaction for the Elderly

Inga Schlömer, Barbara Klein, Holger Roßberg

Research Group Assistive Technologies, Frankfurt University of Applied Sciences

Abstract

Due to the demographic change most European countries face an increase of older people and a lack of (nursing) care staff. The development of technological systems supporting activities of daily living seems to be increasingly relevant as these might contribute to tackle problems associated with these developments. Within the interdisciplinary European HORIZON 2020 project "I-SUPPORTED Bath Robots" a robotic shower system is being developed that uses a multimodal human-robot interaction approach to support frail and mobility impaired persons. Because of limited physical and cognitive abilities of the intended primary user group of elderly people, formative evaluations of verbal commands and gestures were conducted to ensure acceptance, applicability, and usability. Main contribution of the evaluation was the identification of verbal commands and gestures that were regarded as too difficult or superfluous by elderly users. Those insights will inform the design and development of further prototypes. Furthermore, general suggestions for the design of multimodal HRI for elderly users could be derived.

1 Introcuction

The aging of society results in a growing need in the healthcare sector. Elderly people increasingly require support for activities of daily living (ADL) such as bathing and showering (Katz, 1963). The prevalence of bathing limitations increases with age and is significantly associated with moving into a care home (Gill et al. 2006). To enable elderly people a longer independent life at home, a robotic shower system is being researched and developed by a consortium of nine European partners within the interdisciplinary project "I-SUPPORTED Bath Robots", funded within the framework of HORIZON 2020, Activity PHC-19-2014, grant agreement n°: 643666. The service robot consists of three devices: A motorized shower chair; two soft robotic arms for pouring water, soaping, and washing; and

Veröffentlicht durch die Gesellschaft für Informatik e. V. 2017 in
M. Burghardt, R. Wimmer, C. Wolff, C. Womser-Hacker (Hrsg.):
Mensch und Computer 2017 – Workshopband, 10.–13. September 2017, Regensburg.
 https://doi.org/10.18420/muc2017-ws17-0415

different sensors for human-robot interaction (HRI) and action recognition. Primary users (PU) are elderly people with bathing limitations. Professional caregivers represent secondary users (SU). According to the specific abilities, and requirements of PUs, it is necessary to provide a natural HRI that is easy to understand (Chen, 2014). For this purpose, a multimodal interaction control, with originally 33 iconic gestures and corresponding verbal commands, was designed by project partners from the Institute of Communication and Computer Systems from the National Technical University of Athens (I-SUPPORT Progress Report, 2016). In cooperation with another German project partner, the Karlsruhe Institute of Technology (KIT), German translations were prepared that served as a basis for this evaluation. The aim of this evaluation was to identify gestures and verbal commands that are not applicable for the user groups and to derive suggestions for the modification and/or deletion of commands and gestures.

2 Methods

2.1 Research Design

For evaluating the multimodal HRI a mixed methods approach was applied. For an objective evaluation of the applicability of proposed gestures and verbal commands, participants were observed and video recorded while performing them. In order to collect participants' subjective assessment of the usefulness of HRI, the think-aloud protocol was applied (Boren, 2000). Furthermore, the following questionnaires were adopted: System Usability Scale (SUS) for usability evaluation; Senior Technology Acceptance Model (STAM), and the extended technology acceptance model (TAM2) for evaluating technology acceptance (Brooke, 1996; Chen, 2014; Venkatesh & Davis, 2000). The applicability of long and short versions of spoken commands was evaluated with a five point Likert scale (1-Short version is most suitable, 5-Long version is most suitable). Subsequently, focus group discussions were conducted to gather information on personal opinions and ideas for further and/or less gestures and commands. Evaluated HRI focused on emergency situations and core processes for washing the back. This focus resulted in following ten German short and long, more polite verbal commands which were evaluated (cf. Table 1). Associated twelve gestures covered the same activities.

Table 1: Verbal commands

German verbal command	**English verbal command**
Roberta, dusche (brause) den Rücken. + Roberta, bitte dusche (brause) meinen Rücken.	Wash / rinse my back
Roberta, gut (genug) + Roberta, danke, es ist gut (genug).	OK, thank you / Yes, I am ok! / Good, I am done
Roberta, wasche den Rücken. + Roberta, bitte wasche meinen Rücken.	Scrub my back
Roberta, wärmer. + Roberta, bitte das Wasser	Temperature up / Raise the temperature

German verbal command	English verbal command
wärmer stellen.	
Roberta, kühler. + Roberta, bitte das Wasser kühler stellen.	Temperature down/decrease the temperature
Roberta, noch einmal. + Roberta, bitte noch einmal von vorne.	Repeat / Repeat this / Please start again
Roberta, stopp (halt) + Roberta, bitte stoppe das Duschen.	Stop/Stop immediately / stop right now
Roberta, heiß. + Roberta, das Wasser ist zu heiß.	It's too hot.
Roberta, ich friere.	I'm freezing.
Roberta, zu kalt. + Roberta, das Wasser ist zu kalt.	I'm too cold.

After a demonstration of either gestures or spoken commands participants performed them. Camera and an audio recorder were used for data collection. The evaluation was split into four rounds, because it was not possible to schedule a common appointment for all participants. Each evaluation approximately took two hours. For analyzing the collected data, video and audio recordings were transcribed, coded and examined in a qualitative data analysis (Kuckartz, 2014). In addition, the applied questionnaires were evaluated.

2.2 Participants

In total 22 persons, 13 elderly users and 9 caregivers evaluated both modalities (cf. Table 2). Because of their deeper understanding of abilities and limitations of PUs, SUs were asked to evaluate gesture interaction and comment on it from the elderly peoples' perspective. If a questionnaire was filled out incompletely, it was not included in the analysis. This resulted in divergent number of participants and analyzed questionnaires.

Table 2: Distribution of participants and their characteristics

	Evaluation 1 Gestures, verbal commands	**Evaluation 2** Verbal commands	**Evaluation 3** Gestures	**Evaluation 4** Gestures	
	SUs	PUs	PUs	PUs	SUs
Male	1	3	-	1	3
Female	3	2	3	4	2
Average age	36	70	71	72	34
Age span	27-44	64-73	60-81	67-80	19-51

3 Results

The following section 3.1 presents the results of the analysis of selected items from the STAM and TAM2. In section 3.2 the results of evaluating verbal commands are presented briefly. The results of evaluating gestures are summed up in section 3.3.

3.1 Attitude and Intention towards using the Robotic Shower System

Attitude and intention to use were obtained with items borrowed from the STAM and TAM2 (PUs, n=9; SUs, n=8). STAM uses a ten point Likert scale. Answers within the range of one to four were considered as negative, from five to six as neutral, from seven to ten as positive. TAM2 uses a seven point Likert scale. Answers within the range of one to three were considered as negative, from seven to ten as positive, a rating of four was neutral. **PU**s stated a positive attitude towards using the robotic shower. Being asked if they would like the idea of using the system, the tendency remained unchanged. Six participants were positive, two were neutral, and one was negative. PUs evaluated the perceived usefulness also positively (seven out of nine). This was also true for the expected convenience the system would have for their life and their effectiveness in life. **SU**s' intention to use the service robotics system was positive. When asked about an assumed increase of work performance, SUs' opinions were divided: Two SUs gave a positive answer, two answers were neutral, and four negative. Regarding an expected productivity increase, three SUs evaluated the system positively, five negatively. In summary, it could be deduced, that PUs had a more positive attitude towards the robotic shower system. SUs seemed to be more skeptical, especially when it came to effects on their work.

3.2 Verbal Commands

Overall usability evaluation: Five PUs and three SUs filled out the SUS questionnaire based on a five point Likert scale, ranging from 0='strongly disagree' to 4='strongly agree'. SUS scores have a range from 0 to 100; from a score of 68 the usability is considered as good (Schröder, 2014). The individual SUS score by PUs ranked from 40 to 100, with a total SUS score of 70.5, indicating a good usability. The individual SUS score by SUs ranked from 45 to 67.5, with a total SUS score of 60, indicating a below-average usability. Differences between PUs' and SUs' evaluation can be found in the assessment of ease of use. For SUs the voice control was rather hard to use. In addition, the voice control had too much inconsistency for SUs. Although PUs felt more confident in using the voice control compared to SUs, they thought they needed to learn a lot to use the voice control.

Evaluation of certain verbal commands: Participants found it difficult to distinguish between different commands (e.g. "showering" and "washing"). For some participants "showering" had a wider meaning and included "washing". Some alternatives were suggested such as direct commands: "wet" or "water". This was in line with a predominant preference for short commands. However, even some short commands were considered as too long and cumbersome, e.g. "Roberta, scrub my back". In contrast, for one PU the

interaction with long verbal commands, especially using phrases like “please” and “thanks” was very important. It was pointed out, that the verbal interaction with the robot might be the only communication elderly people have throughout the day. Some individual choices of German words and sentence structure were discussed throughout the different focus groups, e.g “dusche” instead of “brause”, “kälter“ instead of “kühler”. Furthermore, some processes were considered as redundant. With regard to the following emergency commands: “Roberta, heiß” (It’s too hot) and “Roberta, zu kalt” (I’m too cold) participants expected an automatism so that the water temperature would remain within appropriate levels. If emergency commands would be necessary, shorter versions such as “heiß” (hot) or “zu heiß” (too hot) as well as “kalt” (cold) instead of “Roberta, zu kalt” (I’m too cold) were preferred.

3.3 Gestures

Overall usability evaluation: Six PUs and nine SUs participated in the usability evaluation of gestures. The SUS score by SUs ranked from 20 to 77.5, with a total SUS score of 45.3. The SUS score by PUs (n=6) ranked from 30 to 75, with a total SUS score of 50.4. PUs were more likely to use the gesture control. They also felt much more confident in using it. Still, PUs thought they needed to learn a lot to use the gesture control. However, PUs rather than SUs thought that most people would learn to use the gesture control quickly. Overall SUS scores of 45.3 and 50.4 indicated a below-average usability.

Evaluation of certain gestures: Some participants found the movement of the arm towards the upper back for initiating “wash my back” difficult. Arthrosis of the shoulder joint would restrict elderly users’ ability to move their arms up. In addition, one user received the arm movement as unnatural. Alternatives were suggested such as pointing at the back with a raised thumb and performing gestures in front of the upper body without big movements in the shoulder joint. Furthermore, participants noted, that contractures of the hand resulted in a decreased flexibility and functionality. In this context, participants stated that the gesture for adjusting the water temperature - a three-step movement that included both arms and hands – was much too complex and could not be remembered. Participants suggested, e.g. moving one hand or one hand with an outstretched index finger up and down to indicate a temperature adjustment. Particular importance was laid on users who had suffered a stroke followed by one-sided hemiparesis or hemiplegia. Participants explained that also for disabled users, gestures that only require one arm are desirable. Participants explained that it should be noted, that gesture recognition should not misinterpret hyperkinesia or shrugging for a gesture. As already noted in the evaluation of verbal gestures, emergency gestures that indicated too hot water were again considered as unnecessary.

4 Preliminary Conclusions and Outlook

The evaluations identified verbal commands, e.g. long versions, and gestures, e.g. adjusting the temperature, that were regarded as too difficult or superfluous. In addition, general suggestions for the design of multimodal HRI for elderly users could be derived. Because of

their specific limitations, gestures should be limited to the movement of one arm and hand, avoiding complex motions. In addition, movements should be performed in the upper body area without too much movement in the shoulder joint. Complex gestures should be avoided, because the intended user group might not be able to memorize them. Furthermore, gestures should avoid similarities to characteristic symptoms of diseases that are related to aging. More importantly, results from the four HRI evaluations indicated that a large majority of PUs welcomed voice and gesture control. In contrast, SUs were rather skeptical. In addition, they also assumed that PUs' acceptance would be rather low. SU seem to have a biased perception of PU and underestimate their abilities and interest in technology. To avoid a paternalistic design of future care arrangements, intended elderly user groups should be integrated into design decisions more frequently and at an earlier stage. Although the findings of our research had limited statistical validity, they provided useful qualitative information and guidance for further design and development. For further investigations, tests with a functioning prototype are planned for obtaining realistic evaluations.

References

Boren, T.; Ramey, J. (2000): *Thinking aloud: reconciling theory and practice*. In: IEEE Trans. Profess. Commun. 43 (3), (p. 261–278)

Brooke, John (1996): *SUS-A quick and dirty usability scale*. In: Usability evaluation in industry 189 (194), (p. 4–7)

Chen, Ke; Chan, Alan Hoi Shou (2014): Gerontechnology acceptance by elderly Hong Kong Chinese: a senior technology acceptance model (STAM). In: *Ergonomics* 57 (5), (p. 635–652).

Gill, Thomas M.; Allore, Heather G.; Han, Ling (2006): Bathing disability and the risk of long-term admission to a nursing home. In: *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 61 (8), (p. 821–825).

Katz, Sidney; Ford, Amasa B.; Moskowitz, Roland W.; Jackson, Beverly A.; Jaffe, Marjorie W. (1963): Studies of illness in the aged: the index of ADL: a standardized measure of biological and psychosocial function. In: *Jama* 185 (12), (p. 914–919).

Kuckartz, U. (2014): *Qualitative text analysis: A guide to methods, practice and using software*. Sage.

Lopez, R. (2016): I-SUPPORT: ICT-Supported Bath Robots, *Deliverable: Progress Report II*

Schröder, M. (2014): *Die System Usability Scale* In: UseTree - Berliner Kompetenzzentrum für Usability Maßnahmen. Onlone: http://usetree.de/die-system-usability-scale/, Requested on 26.02.2017

Venkatesh, Viswanath; Davis, Fred D. (2000): A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. In: *Management Science* 46 (2), (p. 186–204).