



Using socially assistive robots for monitoring and preventing frailty among older adults: a study on usability and user experience challenges

Richelle A. C. M. Olde Keizer¹ · Lex van Velsen^{1,2} · Mathieu Moncharmont³ · Brigitte Riche³ · Nadir Ammour³ · Susanna Del Signore⁴ · Gianluca Zia⁵ · Hermie Hermens^{1,2} · Aurèle N'Dja³

Received: 12 April 2018 / Accepted: 27 March 2019 / Published online: 9 April 2019

© The Author(s) 2019

Abstract

Socially assistive robots can play an important role in the monitoring and training of health of older adults. But before their benefits can be reaped, proper usability and a positive user experience need to be ensured. In this study, we tested the usability and user experience of a socially assistive robot (the NAO humanoid robot) to monitor and train the health of frail older adults. They were asked to complete a set of health monitoring and physical training tasks, once provided by the NAO robot, and once provided by a Tablet PC application (as a reference technology). After using each technology, they completed the System Usability Scale for usability, and a set of rating scales for perceived usefulness, enjoyment, and control. Finally, we questioned the participants' preference for one of the technologies. All interactions were recorded on video and scrutinized for usability issues. Twenty older adults participated. They awarded both technologies 'average' usability scores. Perceived usefulness and enjoyment were rated as very positive for both modalities; control was scored positively. Main usability issues for NAO for these tasks were related to speech interaction (e.g., NAO's limited speech library, NAO's difficulty to cope with Dutch dialect), older adults' difficulties with taking their proper role in human-robot interaction, and a lack of affordances of NAO. Seven participants preferred NAO: it was easier to use and more personal. Social robots have the potential to monitor and train the health of frail older adults, but some critical usability challenges need to be overcome first.

Keywords Socially assistive robots · Usability · User experience · Older adults · Healthcare

1 Introduction

The first law of robotics, as taken from Asimov's famous novel *I, Robot*, states that "a robot may not injure a human being or, through inaction, allow a human being to come to harm" [1]. Within the healthcare context, social robots are usually expected to do a bit more and

are used to increase the health of human beings. Social robots, in this context, are referred to as socially assistive robots: robots that aim to foster "close and effective interaction with a human user for the purpose of giving assistance and achieving measurable progress in convalescence, rehabilitation, learning, etc" [2]. Such robots can be used in a wide range of tasks. For elderly care, they are used to, for example, bathe people, provide companionship, monitor health, and monitor falls [3]. And the first studies that delved into the effectiveness of using social robots for these goals show positive effects [4, 5]. However, a myriad of factors act as prerequisite for successful acceptance of socially assistive robots among older adults and need to be accounted for during design and implementation, including ease of use, enjoyment, and controllability [6, 7].

Frailty, a situation in which a person (most often an older adult) is "at increased risk for future poor clinical outcomes, such as development of disability, dementia,

✉ Lex van Velsen
l.vanvelsen@rdd.nl

¹ eHealth Group, Roessingh Research and Development, Roessinghsbleekweg 33b, 7522AH, Enschede, The Netherlands

² Biomedical Signals and Systems Group, University of Twente, Enschede, The Netherlands

³ Sanofi, Paris, France

⁴ Bluecompanion LTD, London, UK

⁵ Caretek s.r.l., Torino, Italy

falls, hospitalization, institutionalization, or increased mortality” [8], may lend itself excellently for using socially assistive robots. As frailty is made up of many dimensions (such as decline on the physical or cognitive condition, or malnutrition [9]) which deteriorate gradually, and whereby the older adult does not notice the development of an unhealthy situation, identification of frailty is important to prevent the negative consequences of being frail. And in a situation where frailty or the first signs thereof have become a reality, a person’s health needs to be closely monitored, and health training and education need to be provided. Since frailty is a quite recently discovered but highly prevalent phenomenon (the percentage of community-dwelling adults showing the first signs of frailty range between 30.4% to 44.9% in ten European countries while the frail group ranged from 1.3% to 5.9% [10]), social robots may be an engaging, cost-effective means to monitor and train the health of older adults who live in caring homes (and where frailty can be considered to be highly prevalent [11]).

Before a socially assistive robot can be helpful in frailty care, proper usability and a positive user experience need to be ensured. In this article, we report on a study that aimed to uncover the usability and user experience issues that socially assistive robot design needs to overcome in order to be an effective and well-accepted means among older adults for identifying and monitoring frailty and for providing health training. We also determined older adults’ acceptance of both technologies. The results of this study will allow us to understand what hinders effective human-robot interaction among older adults that use this technology for health purposes, and what influences their decision to use them. For policy makers and robot designers, such information is crucial when deciding whether or not to use social robots for frailty screening, monitoring and prevention, and how to design such technology in order to optimize usability and the user experience.

2 Theoretical background

Human-robot interaction can be perceived from either the robot’s or human point of view. One can perceive the robot to be an entity by itself, and human-robot interaction serves to fulfil the robot’s needs (in which case needs are pre-programmed by the design team, and can be, for example, the need to extract knowledge from a human in order to complete a user model). When considered from the human point of view, human-robot interaction focuses on how a robot can complete a task in an acceptable and comfortable manner [12]. The second interpretation includes two aspects. On the one hand, it highlights the concept of acceptance, which is often studied by means of the Technology Acceptance Model [e.g., 13] or the Unified Theory of Acceptance and Use of Technology [e.g., 14], and

deals with the identification of factors that explain the intention to use a robot, such as the aforementioned ease of use, enjoyment, and controllability. On the other hand, it highlights the concept of usability.

Usability is defined as “the effectiveness, efficiency and satisfaction with which specified users achieve specified goals in particular environments” [15]. Usability engineering has a long tradition in product design and human-computer interaction. This tradition has resulted in a rich methodological toolkit that supports researchers in identifying issues that hinder good usability, such as thinking aloud (whereby an end-user is asked to interact with a technology while constantly voicing his/her thoughts out loud) [16], heuristic evaluation (whereby evaluators are asked to judge an interface with the aid of a set of design guidelines) [17], and cognitive walkthroughs (in which experts are asked to ‘walk through’ a computer system or website as a normal user and report any shortcomings) [18]. Besides methods to elicit usability issues, there is also a range of tools to benchmark the usability of a given technology, of which the System Usability Scale (a ten item questionnaire) [19] is the most widely used one. Usability tests of social robots or socially assistive robots are quite hard to find in the scientific literature. Fischinger and colleagues [20] tested the usability of a socially assistive robot that aims to prevent falls, and to detect and handle emergencies among older adults and identified several improvements that needed to be made so as to improve switching between input modalities. Other studies were mainly designed to assess usability metrics (like task completion time and number of errors) [21]. However, it has been stated that it is difficult to define a set of common metrics and instruments to assess the quality of human-robot interaction, as the range of robots with which people can interact is incredibly diverse [22, 23].

The user experience is a concept that has recently gained a lot of attention in research and design. It deals with the cognitive, socio-cognitive and affective aspects a person experiences while interacting with a product or technology, like enjoyment, aesthetics, and a desire for repeated use [24]. So, where usability focuses on the pragmatic qualities of a technology or product, the user experience is concerned with its hedonic qualities and people’s reactions after a period of usage [25, 26]. A variety of factors that potentially contribute to the user experience of interacting with a social robot have been explored in previous studies. Examples include the congruence between robot and end-user personality [27], empathy [28], and appearance [29]. However, a well-researched and widely accepted model of the user experience of social robots is currently lacking. The experience of users while interacting with a social robot can best be assessed by using a combination of qualitative data collection during interaction (such as thinking-aloud) and a quantitative post-interaction data collection method (such as interviews or questionnaires) [30].

3 Method

We assessed the usability and user experience of a socially assistive robot by means of observing older adults interacting with a robot, programmed to question their frailty status and to explain physical exercises. Afterwards, they were questioned about acceptance, usability and a set of user experience factors. In order to have a reference point for interpreting the results, we asked them to perform the same screening and exercise tasks by means of a tablet PC application.

3.1 Participant recruitment

We recruited 20 older adults, aged 70 years or older via an organization for elderly care in the region of Twente in the Netherlands. Participants needed to speak Dutch fluently and they were excluded from participation when they had physical impairments that posed a safety risk for doing the exercises as was being instructed by the socially assistive robot or Tablet PC application. Finally, a participant needed to be either frail or pre-frail (a state in which the first signs of frailty are present), so that administering the screening instrument and providing physical exercises would make sense. This verdict was given by the care team of a potential participant.

3.2 Procedure

Each individual test was started by explaining the goal of the evaluation to the participant, after which basic demographics were assessed. Then, a participant was asked to interact with the socially assistive robot or the tablet PC application, which were offered in random order. Both technologies provided a module to monitor the frailty status of the older adult, as well as a module to instruct older adults in doing physical exercises (as deterioration of the physical condition is an important aspect of frailty, and physical exercises are an important part of frailty treatment [31]). The monitoring module consisted of the SARC-F, a questionnaire to screen for sarcopenia [32]. The exercising module consisted of four physical exercises, taken from the OTAGO program (which aims at preventing falls by improving physical strength) [33]. They are: (1) Stretching for shoulder, (2) Walking and turning around, (3) One leg stand (no support), (4) Knee bends (hold support). After completion of both modules, the participant completed a set of usability and user experience questionnaires. Then, the participant interacted with the other technology, performed the same tasks, and completed the same questionnaires (albeit for the different technology). At the end of the session, we asked whether the participant preferred the socially assistive robot or the tablet PC application, as well as their rationale for this choice.

3.3 Technology

The socially assistive robot we used during the tests was the NAO humanoid robot (SoftBank Robotics). NAO was programmed to read aloud and interpret the monitoring questions belonging to the SARC-F questionnaire, and could perform the four different physical exercises, after which the robot asked the participant to repeat each exercise. See Fig. 1 for a photo of the NAO robot. The tablet PC which we used was a Samsung TabPRO (SM-T520), with a screen diameter of 10.1 inch.. The questionnaire and exercises were provided in a web-environment, opened within Firefox. This web-environment was optimized for use for older adults (e.g., large buttons and fonts were used), see Fig. 2.

3.4 Data collection


At the start of each session, we interviewed participants about demographics (gender, birth date, living situation, cognitive and physical impairments). After interacting with each technology we assessed the usability of the technology (by means of the System Usability Scale, the preferred method for assessing this factor (SUS) [26, 34–36], perceived usefulness via three statements and a five-point Likert scale [37], and two user experience factors: Enjoyment and control. Enjoyment can be defined as “the extent to which the activity of using the device is perceived to be enjoyable in its own right, apart from any performance consequences that may be anticipated”



Fig. 1 NAO robot

voortgang: oefening 1 van 4

Rekken van het bovenlichaam 10 sec



Starttijd: 14:34:04 | Duur: 00:03:00

Inleiding

Deze oefening heeft als doel de lenigheid te vergroten en uw spieren soepeler te maken.

Uitvoering

- Ga met uw gezicht naar de muur staan iets verder dan armlengte van de muur
- Houd uw voeten op schouderbreedte van elkaar
- Leun voorzichtig met uw lichaam naar voren en plaats uw handen tegen de muur op schouderhoogte
- Houd uw rug recht en loop langzaam met uw handen tegen de muur op, totdat uw handen boven uw hoofd zijn
- Houdt deze positie 10 seconden vast
- Loop vervolgens voorzichtig met uw handen weer naar schouderhoogte

Aantal herhalingen

Herhaal deze oefening 3 keer.

✔ Oefening gedaan

Fig. 2 Frailty exercise provided via Tablet PC application

[38], while control is “the extent to which a user can bring about or prevent particular actions or states of the system if she has the goal of doing so” [39]. The Enjoyment scale, was assessed by means of a five-point semantic differential scale, was based upon van der Heijden [40], while the Control scale, assessed via a five-point Likert scale, was based upon van Velsen et al. [41].

Via a short interview, a participant’s preference for either NAO or the tablet PC and their reasons for this preference was questioned. A voice recorder was used to record the conversation. All tasks, performed with NAO and the tablet PC application were recorded on video.

3.5 Data analysis

Demographics were analyzed in descriptives. For the remainder of the analyses, results were split in results for NAO and results for the Tablet. The SUS was analyzed following the standard method. The perceived usefulness, enjoyment and control scales, were analyzed on a per-item and scale basis (mean score and standard deviation). Responses on the enjoyment scale were recoded for easier interpretation: higher

scores now denote a positive evaluation of enjoyment while using the Tablet or NAO. Paired-samples t-tests were used to test for significant differences between the scale averages. Video and audio recordings of the participants interacting with NAO and the Tablet were scrutinized for usability issues (issues that hinder effective use, efficient use, and/or user satisfaction). A rehabilitation physician assessed whether or not the physical exercises that were performed during the test were done correctly. Based upon van Velsen et al. [42], issues were provided a severity rating by means of the following rules:

- **Critical problems** prevented participants from completing tasks and/or recurred across all participants;
- **Serious problems** severely increased the task completion time and/or recurred frequently across participants. However, a serious problem did not prevent a participant from completing the task eventually;
- **Minor problems** increased task completion time slightly and/or recurred infrequently across the evaluation participants. Finally, a minor problem did not prevent the evaluation participants from completing a test task easily.

The amount of critical, serious, and minor problems were compared between NAO and the Tablet. Finally, short interviews recordings were transcribed, preferences were determined, and reasons for preferences were thematically grouped.

3.6 Ethics

As this study was conducted to identify problems with different technologies for monitoring frailty and providing health training to reverse or postpone the development of frailty, but the participants volunteered to take part and were not forced to answer questions or conduct exercises, approval from a medical ethical committee was not necessary [43]. Before participation, participants were sent an information package by the research team. Before starting a test, the participant completed an informed consent form.

4 Results

4.1 Participants

Twenty older adults (12 males, 8 females) participated. They had a mean age of 78.5 ± 7.1 years. Fifty percent of them lived alone. No one suffered from cognitive impairments. Some participants suffered from small physical impairments but this did not pose a safety risk for doing the exercises.

4.2 Quantitative measures

After interacting with each technology, participants completed a survey with the System Usability Scale (SUS), and rating scales that assessed perceived usefulness, control, and enjoyment. Results of this survey can be found in Table 1. For the user experience scales (perceived usefulness, control, and enjoyment), scores range from 1 (lowest) to 5 (highest).

The numbers show that, following the interpretation of SUS scores by Sauro and Lewis [44], the usability of both NAO and the Tablet PC application score below average (with scores that represent a grade D). Figure 3 discloses however, that opinions about the usability of both technologies differed. Some had a positive, and some a negative opinion on this point, with one participant being extremely negative about NAOs usability. The scores for perceived usefulness and enjoyment are very positive for both technologies, while the score for control is positive for both NAO and the Tablet PC application. In these cases, opinions also differed, but the majority of the participants was positive about these factors for both technologies (see Fig. 4). T-tests showed that there were no differences between the Tablet PC application and NAO with respect to the SUS score ($t(19) = .585, p = .566$), and the average score for perceived usefulness ($t(19) = .64, p = .53$), control ($t(19) = 1.01, p = .33$), and enjoyment ($t(19) = .27, p = .79$).

Table 1 Survey results

	Tablet	NAO	<i>p</i> value
Usability			
System Usability Scale	62.50 ± 16.50	60.50 ± 18.18	0.566
Perceived usefulness			
I think that using [the tablet/the robot] makes it easier to do exercises correctly	4.80 ± 0.70	4.35 ± 1.27	
I think that [the tablet/the robot] is useful for remaining healthy	3.70 ± 1.78	3.95 ± 1.50	
I think that using [the tablet/the robot] provides me with good insights in my health status	4.15 ± 1.27	3.75 ± 1.41	
<i>Average score</i>	4.21 ± 0.97	4.02 ± 1.16	0.528
Control			
I have a lot control over what I can do with [the tablet/the robot]	3.40 ± 1.57	3.00 ± 1.45	
The [tablet/robot] always listens to what I what I want it to do.	3.75 ± 1.48	3.55 ± 1.47	
I can determine for myself what happens on [the tablet/the robot]	3.90 ± 1.59	3.35 ± 1.66	
<i>Average score</i>	3.68 ± 1.44	3.30 ± 1.33	0.326
Enjoyment			
Using [the tablet/the robot] was enjoyable/disgusting	4.20 ± 1.01	4.05 ± 0.89	
Using [the tablet/the robot] was exciting/dull	3.95 ± 1.00	4.00 ± 1.21	
Using [the tablet/the robot] was pleasant/unpleasant	4.40 ± 1.00	4.45 ± 0.95	
Using [the tablet/the robot] was interesting/boring	4.65 ± 0.75	4.90 ± 0.31	
<i>Average score</i>	4.30 ± 0.80	4.35 ± 0.68	0.794

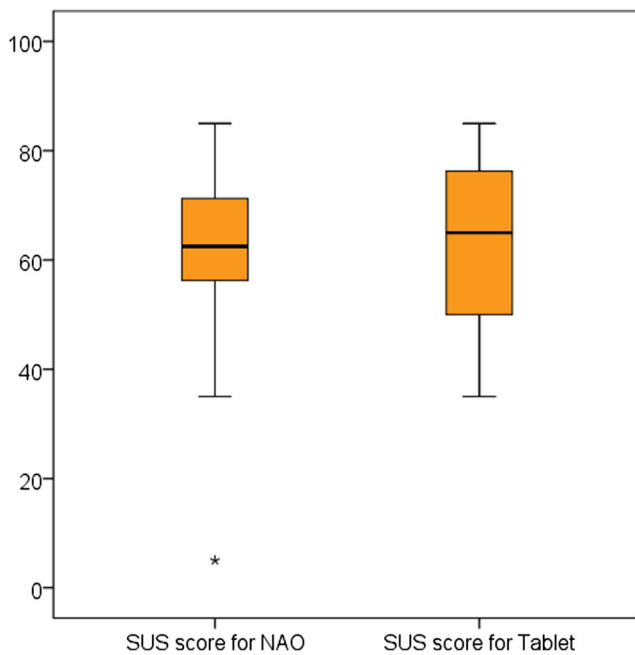
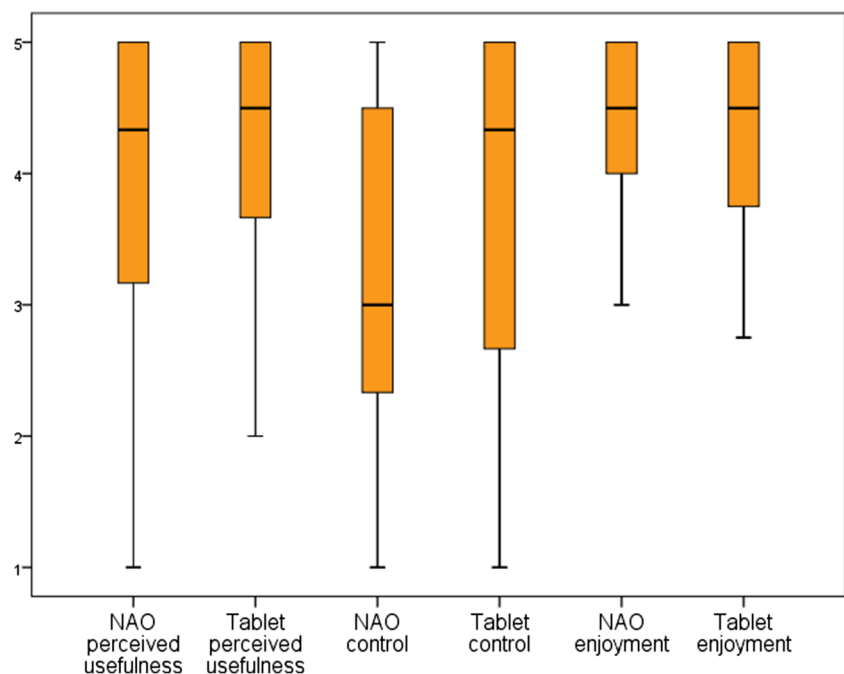


Fig. 3 SUS scores for NAO and the Tablet PC application

4.3 Usability issues

The usability issues that we identified can be divided over three categories: issues that hinder proper monitoring, issues that hinder proper exercising, and finally, general issues (i.e., usability issues that occur in both modules).

Fig. 4 User experience scores for NAO and the Tablet PC application



4.3.1 General issues

Table 2 shows usability issues we identified during our viewing of participants using of NAO and the Tablet PC application, and that are not specific towards monitoring or exercising. For every problem, we noted how often it occurred during all interactions, how many participants experienced this issue, and what priority level we assigned to it. As you can see, an issue could occur multiple times within a single session.

There were three general, major usability issues that occurred when participants interacted with NAO. All three were related to the human-robot interaction using speech. First and foremost, they answered too soft for NAO to hear them (correctly). Second, the participants were unable to hear NAO, and third, they answered too soon (i.e., before NAO completed its speech, so that the participant's answer could not be rightly interpreted by the robot). With regard to interacting with the Tablet PC, we identified one general, critical usability issue, namely that participants had problems with touching the tablet's touchscreen correctly.

4.3.2 Issues hindering proper monitoring

A specific set of usability issues was identified when observing participants interact with NAO or the Tablet PC application for the purpose of monitoring health (for which they completed the SARC-F questionnaire). An overview of these issues is presented in Table 3.

Table 2 General usability issues for NAO and the Tablet PC application

Problem	NAO			Tablet PC		
	Occurrence	n	Severity	Occurrence	n	Severity
<i>The older adult...</i>						
...answered too soon	25	12	Critical			
...answering too softly	47	15	Critical			
...was unable to hear NAO	34	10	Critical			
...did not answer at all	6	4	Serious			
...answered too late	5	4	Serious			
...was unable to touch touchscreen correctly				28	18	Critical
...was unsure how to use the tablet				3	3	Serious
...was unable to read the text on the screen				1	1	Serious

The most frequently encountered problem when interacting with NAO was that the robot was unable to understand the answer given to it. This could be due to the fact that the answer the older adult provided could not be interpreted by NAO as it was not included in its speech library or that the older adult spoke with a Dutch, regional accent that NAO could not understand. On the other side, older adults could often not understand NAO or did not know what to answer (for example, when, for them, NAO talked too quickly after which they could not remember the different answering possibilities). When interacting with the Tablet PC, only one critical issue surfaced. Here, a part of the older adults was unable to select the right answering possibility. They simply did not know they could press an answering option on the screen, pressed too hard or too long (which caused the tablet to malfunction), or pressed an answering option with their fingernail, which the tablet did not recognize as the selection of an option.

4.3.3 Issues hindering proper exercising

The problems that occurred during exercising (as instructed by NAO or the Tablet PC) can be divided into two categories: problems that occurred over all exercises and exercise-related problems. Table 4 displays the general problems that occurred during exercising. It shows that problems related to a proper

distribution of roles (between instructor and the one being instructed) arose quite frequently. This was mainly the case for NAO, but was also observed while older adults interacted with the Tablet PC. Older adults did not understand that NAO first instructed an exercise, did not understand they should repeat an exercise (observed for both NAO and the Tablet PC), or did not understand they should touch NAO’s head when done exercising.

We identified a very wide range of exercise-related problems of which we will only discuss the critical and serious ones. With respect to exercise 1 (Stretching the shoulder), participants did not use a wall (serious issue; 8 participants; NAO only), stood too close to a wall (serious issue; 2 participants for NAO; 5 participants for the Tablet PC), did not lift their arms high enough (serious issue; 1 participant; Tablet PC only), or did not keep their arms above their head for the full 10 s (serious issue; 2 participants; Tablet PC only). For the case of exercise 2 (Walking and turning around), in which participants were instructed to walk in the shape of an 8, we observed many erroneous executions. Participants exactly copied NAO and walked with very small steps (critical issue; 11 participants; NAO only) or made a zigzagging movement (critical issue; 8 participants; NAO only). Other erroneous executions of this exercise included making only a few steps (critical issue; 3 participants; NAO only), walking from left to

Table 3 Usability issues hindering proper monitoring

Problem	NAO			Tablet PC		
	Occurrence	n	Severity	Occurrence	n	Severity
<i>The older adult...</i>						
...was unable to understand NAO	16	11	Critical			
...did not know what to answer	12	8	Critical			
...did not provide an answer NAO could understand	60	16	Critical			
...did not understand how to answer questions				8	5	Critical
...did not understand instructions				1	1	Minor

Table 4 General problems occurred during exercising with NAO and the tablet

Problem	NAO			Tablet PC		
	Occurrence	n	Severity	Occurrence	n	Severity
<i>The older adult...</i>						
...does not understand that NAO demonstrates an exercise first	49	18	Critical			
...does not understand that s/he should repeat the exercise	18	14	Critical	19	10	Critical
...does not understand that s/he has to touch NAOs head when finished	10	6	Critical			
...has difficulty with exercising and watching NAO simultaneously	3	2	Serious			
...does not understand instructions	2	1	Serious	10	5	Serious
...touches NAOs head at the wrong place	1	1	Minor			
...doubts whether to watch NAO or exercise	1	1	Minor			
...starts exercising too soon	1	1	Minor			
...is unwilling willing to watch the video and starts exercising				13	9	Minor
...is unwilling to read the instructions				1	1	Minor

right (critical issue; 1 participant for NAO; 3 participants for the Tablet PC), walking in an unidentifiable pattern (critical issue; 2 participants; Tablet PC only), walking in very small circles (serious issue; 2 participants; Tablet PC only), or walking in two separate circles (serious issue; 6 participants for NAO; 6 participants for the Tablet PC). Finally, 3 participants decided to stop and stand still in the middle of the exercise (serious issue; NAO only). For exercise 3 (standing on one leg (without support)) and 4 (bending the knee (with support)), we identified only one critical or serious issue per exercise. During exercise three we observed that participants copied the movements of NAO exactly and dragged their feet over the ground (serious issue; 3 participants; NAO only), and during exercise 4, we observed that people did not place their hands on the table (serious issue; 5 participants; NAO only).

4.3.4 Preference for NAO or tablet PC application

At the end of each session, participants indicated whether they preferred using NAO or the tablet for monitoring or training their health. Overall, 13 participants preferred the tablet and 7 participants preferred NAO. Reasons that participants gave for preferring the tablet were:

- The Tablet PC is easier to use (mentioned 5 times).
- The videos in the Tablet PC application make exercising easier, as they are explained more clearly (mentioned 4 times).
- Practicing with NAO is hard if you have a hearing impairment (mentioned 3 times).
- The tablet is smaller than the robot (mentioned 2 times).
- I am already very used to working with a tablet (mentioned 2 times).
- I am not used to working with NAO (mentioned 1 time).
- NAO is something for the future (mentioned 1 time).

Participants who preferred using NAO supplied the following arguments for their preference:

- It is easier to use the robot (mentioned 5 times).
- NAO is more personal, it can be a buddy (mentioned 2 times).
- NAO has no loading time, it is faster (mentioned 1 time).
- NAO is interesting (mentioned 1 time).
- NAO shows how to do the exercises (mentioned 1 time).
- NAO tells you how to do the exercises (mentioned 1 time).

5 Discussion

In this study, we assessed the different usability and user experience challenges that social robots face when being used to monitor and train the health of frail older adults. With respect to usability issues, we found that both the social robot that was being used (NAO), as well as the Tablet PC application (which was tested as a reference technology) suffered from several usability issues. Participants had difficulty with interacting with the social robot as a) they experienced problems while talking with the robot, b) they found it difficult to identify their role during human-robot interaction, and c) did not have a clear image of the relation between the possible interaction options the social robot provided and the consequences of using one of these options. From observing the interaction between the social robot and older adult, it became clear that the social robot was not technically capable of having a full and rewarding conversation with an older adult. Older adults answered too soon, too softly, too late, or were not able to hear the social robot. The social robot, on the other hand, had difficulty with interpreting the speech of the older adults,

which was often in Dutch dialect or included words NAO could not process. Before social robots can be successfully used to monitor and train the health of pre-frail older adults, their speech library, speech, and flexibility for coping with answers that are provided before or after the social robot expects it, should be thoroughly improved. This need was also identified by [45] who found that the hardware of the NAO robot and its speech library are not of such a quality that they provide good speech interaction. Next, older adults found it difficult to find their role while interacting with the robot: what is expected of him or her? During our tests this manifested itself as copying exactly what the social robot does (even if this leads to unnatural movements that hinder the effectiveness of physical exercises) and unsuccessful switching between roles (i.e., the robot first instructs after which the older adult performs an exercise). Social robots are not a common conversation partner or instructional agent for older adults, which makes it difficult for them to determine how they should interact with a robot and how they should interpret their instructions. Introduction of social robots to support health monitoring and physical exercising should therefore always be preceded by a period of thorough instruction and a trial period, so that older adults can get accustomed to the interacting with robots and know ‘how to play the game’. Finally, older adults found it difficult to identify interaction options and to understand the consequences of using such an action (i.e., the affordances [46] of the social robot). This manifested itself most prominently in problems when older adults needed to touch the head of the social robot to continue the physical training program: it was an unnatural action and not something they immediately associated with completing an exercise. Social robot design can therefore best refrain from using such interactions methods and focus on a properly working speech interaction or interaction via a touch-screen (as, for example, the Aido Interactive Personal Home Robot by InGen Dynamics allows). This claim is strengthened by the findings of Hebesberger and colleagues [47] who found that when introducing an autonomous robot, its functionalities should be self-explanatory so that the robot can be used without help from a care professional or support staff. The presence of these usability issues was also reflected in the score the robots received on usability (assessed via the System Usability Scale (SUS)), which could be interpreted as unsatisfactory. This score was similar to the SUS score of the Tablet PC application that was used as a reference technology. These results also suggest that the SUS is indicative of social robot usability, even though the instrument was not developed for assessing this type of technology initially.

The user experience scores that were awarded to the social robot (split out into the factors perceived usefulness, control, and enjoyment) were very positive and equal to the scores that were awarded to the Tablet PC application that was provided as a reference technology. Moreover, a third of the participants

in our study (7 out of 20) preferred the social robot over the Tablet PC application for monitoring and training their health. Mostly, they stated that the robot was easier to use and is a more personal technology. We think these results are highly encouraging. If, with all the flaws that were present in the currently used social robot, our study still showed signs that social robots are an engaging monitoring and training technology for a group of older adults, this acceptance can only increase once the quality of social robots for effective and efficient human-robot interaction further improves.

6 Limitations

The social robot used in this study was NAO, which is only one example of the wide range of social robots that are currently available. As we used only one type of robot, the identified usability issues might not be fully representative of the generic usability issues that social robots currently face. However, NAO is a popular and widely used robot and displays many characteristics that can be found among the current generation of social robots. We therefore think that our inventory gives a good indication of the current state of the art. For a full and definite picture, future research should confirm this claim and should report usability challenges that other social robots face. We focused this study on using social robots for monitoring and preventing frailty among older adults. The results should therefore be taken with caution if one wants to generalize them towards other (geriatric) conditions. For example, a social robot might not be very well suited to support older adults in performing cognitive training exercises. Finally, our study included 20 participants. Although this number is very acceptable for conducting a usability test [48], it might, from a scientific point of view, limit the potential for generalization of our results. Similarly, our local recruitment approach might have introduced some bias in our participant sample (Dutch frail or pre-frail older adults).

7 Concluding remarks

Social robots are becoming increasingly popular in healthcare for monitoring and health training purposes. Especially as they can be an engaging and cost-effective means for doing so. Our study showed that social robots have potential when used for monitoring and training the health of frail older adults, but that there are still some critical usability challenges that need to be overcome first. We are therefore looking forward to studies and technological innovations that tackle the critical usability issues we identified.

Funding This work is conducted within the context of the IMI SPRINTT (IMI-JU 115621) project.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does contain a study with human participants. This study was conducted to identify problems with different technologies for monitoring frailty and providing health training to reverse or postpone the development of frailty. Participants volunteered to take part and were not forced to answer questions or conduct exercises, so, according to Dutch law, approval from a medical ethical committee was not necessary (Central Committee on Research Involving Human Subjects. Manual for the review of medical research involving human subjects. <http://www.ccmo-online.nl>.)

Informed consent Informed consent was obtained from all individual participants included in the study.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Asimov I. I, robot. New York: Gnome Press; 1950.
- Feil-Seifer D, Mataric MJ. Defining socially assistive robotics. Paper presented at the 9th International Conference on Rehabilitation Robotics, 28 June–1 July 2005. <https://doi.org/10.1109/icorr.2005.1501143>.
- Robinson H, MacDonald B, Broadbent E. The role of healthcare robots for older people at home: a review. *Int J Soc Robot.* 2014;6: 575–91. <https://doi.org/10.1007/s12369-014-0242-2>.
- Bemelmans R, Gelderblom GJ, Jonker P, de Witte L. Socially assistive robots in elderly care: a systematic review into effects and effectiveness. *J Am Med Dir Assoc.* 2012;13:114–120.e111. <https://doi.org/10.1016/j.jamda.2010.10.002>.
- Kachouie R, Sedighadeli S, Khosla R, Chu M-T. Socially assistive robots in elderly care: a mixed-method systematic literature review. *International Journal of Human–Computer Interaction.* 2014;30: 369–93. <https://doi.org/10.1080/10447318.2013.873278>.
- Heerink M, Kröse B, Evers V, Wielinga B. Assessing acceptance of assistive social agent technology by older adults: the Almere model. *Int J Soc Robot.* 2010;2:361–75. <https://doi.org/10.1007/s12369-010-0068-5>.
- de Graaf MMA, Ben Allouch S. Exploring influencing variables for the acceptance of social robots. *Robot Auton Syst.* 2013;61:1476–86. <https://doi.org/10.1016/j.robot.2013.07.007>.
- European Innovation Partnership on Active and Healthy Ageing. Action plan on prevention and early diagnosis of frailty and functional decline, both physical and cognitive in older people. Brussels: European Union; 2012.
- Gomez MI, García-Sánchez I, Carta A, Antunes JP. A collection of good practices that support the prevention and early diagnosis of frailty and functional decline, both physically and cognitive, in older people. Brussels: European Commission; 2013.
- Santos-Eggimann B, Cuénoud P, Spagnoli J, Junod J. Prevalence of frailty in middle-aged and older community-dwelling Europeans living in 10 countries. *J Gerontol Ser A Biol Med Sci.* 2009;64A: 675–81. <https://doi.org/10.1093/gerona/glp012>.
- Kojima G. Prevalence of frailty in nursing homes: a systematic review and meta-analysis. *J Am Med Dir Assoc.* 2015;16:940–5. <https://doi.org/10.1016/j.jamda.2015.06.025>.
- Dautenhahn K. Socially intelligent robots: dimensions of human–robot interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences.* 2007;362:679–704. <https://doi.org/10.1098/rstb.2006.2004>.
- Ezer N, Fisk AD, Rogers WA. Attitudinal and intentional acceptance of domestic robots by younger and older adults. Universal access in human–computer interaction: 5th international conference, held as part of HCI International, San Diego, CA, USA, July 19–24, 2009. https://doi.org/10.1007/978-3-642-02710-9_5.
- Heerink M, Kröse B, Evers V, Wielinga B. Measuring acceptance of an assistive social robot: a suggested toolkit. Paper presented at RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication, Sept 27 - Oct 2. 2009. <https://doi.org/10.1109/roman.2009.5326320>.
- International Organization for Standardization ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDTs) — Part 11 Guidance on usability. 1998.
- Jaspers MWM, Steen T, Cvd B, Geenen M. The think aloud method: a guide to user interface design. *Int J Med Inform.* 2004;73: 781–95. <https://doi.org/10.1016/j.ijmedinf.2004.08.003>.
- Nielsen J, Molich R. Heuristic evaluation of user interfaces. Paper presented at the SIGCHI Conference on Human Factors in Computing Systems, Seattle, Washington, USA. 1990.
- Polson PG, Lewis C, Rieman J, Wharton C. Cognitive walkthroughs: a method for theory-based evaluation of user interfaces. *International Journal of Man-Machine Studies.* 1992;36: 741–73. [https://doi.org/10.1016/0020-7373\(92\)90039-N](https://doi.org/10.1016/0020-7373(92)90039-N).
- Brooke J. SUS – a quick and dirty usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland AL, editors. Usability evaluation in industry. London: Taylor and Francis; 1996. p. 189–94.
- Fischinger D, Einramhof P, Papoutsakis K, Wohlkinger W, Mayer P, Panek P, et al. Hobbit, a care robot supporting independent living at home: first prototype and lessons learned. *Robot Auton Syst.* 2016;75:60–78. <https://doi.org/10.1016/j.robot.2014.09.029>.
- Granata C, Pino M, Legouverneur G, Vidal JS, Bidaud P, Rigaud AS. Robot services for elderly with cognitive impairment: testing usability of graphical user interfaces. *Technol Health Care.* 2013;21(3):217–31. <https://doi.org/10.3233/THC-130718>.
- Steinfeld A, Fong T, Kaber D, Lewis M, Scholtz J, Schultz A, Goodrich M. Common metrics for human-robot interaction. Paper presented at the 1st ACM SIGCHI/SIGART conference on Human-robot interaction, Salt Lake City, Utah, USA. 2006.
- Begum M, Wang R, Huq R, Mihailidis A. Performance of daily activities by older adults with dementia: The role of an assistive robot. Paper presented at the 13th IEEE International Conference on Rehabilitation Robotics, 24–26 June. 2013. doi:<https://doi.org/10.1109/icorr.2013.6650405>
- Law ELC, van Schaik P. Modelling user experience – an agenda for research and practice. *Interact Comput.* 2010;22:313–22. <https://doi.org/10.1016/j.intcom.2010.04.006>.
- Hassenzahl M. The interplay of beauty, goodness, and usability in interactive products. *Human-Computer Interaction.* 2008;19:319–49. https://doi.org/10.1207/s15327051hci1904_2.
- Borsci S, Federici s BS, Gnaldi M, Bartolucci F. Assessing user satisfaction in the era of user experience: comparison of the SUS, UMUX, and UMUX-LITE as a function of product experience. *International Journal of Human–Computer Interaction.* 2015;31: 484–95. <https://doi.org/10.1080/10447318.2015.1064648>.
- Lee KM, Peng W, Jin S-A, Yan C. Can robots manifest personality? An empirical test of personality recognition, social responses, and social presence in human–robot interaction. *J Commun.* 2006;56: 754–72. <https://doi.org/10.1111/j.1460-2466.2006.00318.x>.

28. Cramer H, Goddijn J, Wielinga B, Evers V. Effects of (in)accurate empathy and situational valence on attitudes towards robots. Paper presented at the 5th ACM/IEEE International Conference on Human-Robot Interaction, 2–5 March. 2010. doi:<https://doi.org/10.1109/hri.2010.5453224>.
29. Wu Y-H, Fassett C, Rigaud A-S. Designing robots for the elderly: appearance issue and beyond. *Arch Gerontol Geriatr*. 2012;54:121–6. <https://doi.org/10.1016/j.archger.2011.02.003>.
30. Weiss A, Bernhaupt R, Tscheligi M, Yoshida E. Addressing user experience and societal impact in a user study with a humanoid robot. Paper presented at the Symposium on New Frontiers in Human-Robot Interaction Edinburgh, Scotland, 8–9 April. 2009.
31. Langlois F, Vu TTM, Chassé K, Dupuis G, Kergoat M-J, Bherer L. Benefits of physical exercise training on cognition and quality of life in frail older adults. *The Journals of Gerontology: Series B*. 2013;68:400–4. <https://doi.org/10.1093/geronb/gbs069>.
32. Malmstrom TK, Miller DK, Simonsick EM, Ferrucci L, Morley JE. SARC-F: a symptom score to predict persons with sarcopenia at risk for poor functional outcomes. *J Cachexia Sarcopenia Muscle*. 2016;7:28–36.
33. Campbell AJ, Robertson MC, Gardner MM, Norton RN, Tilyard MW, Buchner DM. Randomised controlled trial of a general practice programme of home based exercise to prevent falls in elderly women. *Br Med J*. 1997;315:1065–9. <https://doi.org/10.1136/bmj.315.7115.1065>.
34. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction*. 2008;24:574–94. <https://doi.org/10.1080/10447310802205776>.
35. Lewis JR, Sauro J. The factor structure of the system usability scale. In: Korusu M, editor. *Human centered design*. Berlin: Springer; 2009. p. 94–103.
36. Borsci S, Federici S, Lauriola M. On the dimensionality of the system usability scale: a test of alternative measurement models. *Cogn Process*. 2009;10:193–7. <https://doi.org/10.1007/s10339-009-0268-9>.
37. Segars AH, Grover V. Re-examining perceived ease of use and usefulness: a confirmatory factor analysis. *MIS Q*. 1993;17:517–25. <https://doi.org/10.2307/249590>.
38. Davis FD, Bagozzi RP, Warshaw PR. Extrinsic and intrinsic motivation to use computers in the workplace. *J Appl Soc Psychol*. 1992;22:1111–32. <https://doi.org/10.1111/j.1559-1816.1992.tb00945.x>.
39. Jameson A. Adaptive interfaces and agents. In: Jacko JA, Sears A, editors. *Human-computer interaction handbook*. Mahwah: Lawrence Erlbaum; 2007. p. 433–58.
40. van der Heijden H. User acceptance of hedonic information systems. *MIS Q*. 2004;28:695–704.
41. van Velsen L, van der Geest T, van de Wijngaert L, van den Berg S, Steehouder M. Personalization has a Price, controllability is the currency: predictors for the intention to use personalized eGovernment websites. *J Organ Comput Electron Commer*. 2015;25:76–97. <https://doi.org/10.1080/10919392.2015.990782>.
42. van Velsen L, van der Geest T, Klaassen R. Identifying usability issues for personalization during formative evaluations: a comparison of three methods. *International Journal of Human-Computer Interaction*. 2011;27:670–98. <https://doi.org/10.1080/10447318.2011.555304>.
43. Central Committee on Research involving Human Subjects Manual for the review of medical research involving human subjects. <http://www.ccmo-online.nl>.
44. Sauro J, Lewis JR. *Quantifying the user experience: practical statistics for user research*. 2nd ed. Burlington: Morgan Kaufmann; 2016.
45. Mubin O, Henderson J, Bartneck C. You just do not understand me! Speech Recognition in Human Robot Interaction. Paper presented at the 23rd IEEE International Symposium on Robot and Human Interactive Communication, 25–29 Aug. 2014. doi:<https://doi.org/10.1109/roman.2014.6926324>
46. Gaver W. Technology affordances. Paper presented at the SIGCHI conference on human factors in computing systems, April 27 – May 2, New Orleans, LA. 1991. doi: <https://doi.org/10.1145/108844.108856>.
47. Hebesberger D, Koertner T, Gisinger C, Pripfl J. A long-term autonomous robot at a care hospital: a mixed methods study on social acceptance and experiences of staff and older adults. *Int J Soc Robot*. 2017;9:1–13. <https://doi.org/10.1007/s12369-016-0391-6>.
48. Hwang W, Salvendy G. Number of people required for usability evaluation: the 10±2 rule. *Commun ACM*. 2010;53:130–3. <https://doi.org/10.1145/1735223.1735255>.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.