




# An User-Centered Evaluation of Two Socially Assistive Robots Integrated in a Retirement Home

Antonio Jerez<sup>1</sup> · Ana Iglesias<sup>2</sup> · José Manuel Pérez-Lorenzo<sup>3</sup> · Alberto Tudela<sup>1</sup> · Alejandro Cruces<sup>1</sup> · Juan Pedro Bandera<sup>1</sup> 

Accepted: 9 September 2024  
© The Author(s) 2024

## Abstract

Socially assistive robots are receiving a growing interest in the health and social care sectors. They are considered a promising technology to add value to the work of caregivers, and relieve them of simple and repetitive tasks. However, these robots currently face significant difficulties when deployed in everyday scenarios due to a number of factors. Most of these factors are related to insufficient consideration of the user perspective and incorrect evaluation procedures. This paper aims to address these issues. Its objective is to analyze the long-term accessibility, usability, social acceptance and user experience for two different socially assistive robots performing the same tasks in a retirement home. The evaluation procedure is based on a framework specifically designed to consider these criteria. Collected results show that both robots received an overall positive feedback. Although the number of users participating in the evaluation was not very high, due to the chosen recruitment criteria and the period of activity of this research project, during the COVID19 pandemic, these results allow to extract relevant insights towards a meaningful use of social robots in shared social care contexts.

**Keywords** Socially assistive robot (SAR) · Long-term evaluation · User studies · Human–robot interaction (HRI)

## 1 Introduction

Population worldwide is aging significantly. This demographic change particularly affects European (EU) countries, Japan or the United States, among others. According to current projections, by 2050, one in four people living in Europe and North America will be over 65 [1]. Moreover, certain regions within these countries are more sensitive to this change. For instance, in the South of Spain, the proportion of Andalusians aging 64 or more will rise from 19.6% in 2015, to 28.8% in 2035 [2]. This *silver society* will inevitably lead to a higher number of people requiring health and social care services either at home or in retirement homes. Moreover, as a consequence of this demographic shift, the evolution of the Caregiver Support Ratio (CSR) suggests that in the near future there will be very limited availability of carers to meet this increasing demand [3].

Technological tools are emerging as a key asset for these caregivers. Ambient Assisted Living (AAL) environments can integrate autonomous monitoring and communication systems in the home or in health and social care facilities. They can also include mechanisms that can remind or encourage residents to perform rehabilitation or main-

---

✉ Juan Pedro Bandera  
jpbandera@uma.es

Antonio Jerez  
antoniojcor@uma.es

Ana Iglesias  
aiglesia@inf.uc3m.es

José Manuel Pérez-Lorenzo  
jmperez@ujaen.es

Alberto Tudela  
ajtudela@uma.es

Alejandro Cruces  
acruces@uma.es

<sup>1</sup> Department of Electronic Technology, Universidad de Málaga, ETSI de Telecomunicación, Campus de Teatinos, s/n, 29071 Málaga, Spain

<sup>2</sup> Department of Computer Science and Engineering, Universidad Carlos III de Madrid, Av. de la Universidad, 30, 28911 Leganés, Madrid, Spain

<sup>3</sup> Department of Telecommunication Engineering, Universidad de Jaén, E.P.S. de Linares, Campus Científico Tecnológico, s/n, 23700 Linares, Jaén, Spain

tenance exercises, and provide remote assistance. Within AAL ecosystems, the use of *Socially Assistive Robots* (SAR), designed to provide assistance through social interaction [4], may become an interesting option due to the proactivity, autonomy, adaptability and potential acceptability of these agents [5]. They can also become efficient social facilitators [6].

However, the development of SAR is currently facing major difficulties, that go beyond the expected technical ones. The main constraints that currently prevent the widespread use of these robots include: (i) the lack of long term experimentation in real environments [7]; (ii) the acceptability, utility and ethical issues that usually arise from designs that did not sufficiently incorporate the user perspective [8]; and (iii) the difficulties in properly describing, evaluating and regulating SAR [9, 10].

This paper presents the results of an experiment conducted at Vitalia Teatinos retirement home, located in Southern Spain, over several months. The objective of the experiment is to analyze the long-term accessibility, usability, social acceptance and user experience, for two different socially assistive robots doing the same tasks in a retirement home.

The experiment was performed in the context of two regional research projects and one European project. The starting hypothesis for these projects is: *Socially Assistive Robots (SAR) in retirement homes can be accepted by staff, residents and their relatives when their tasks have been defined using user-centered strategies and participatory design.*

This paper focuses on the second part of the validation of this hypothesis. In a previous paper, we focused on the methodology and methods used to carefully design the tasks to be performed by the SAR and the robot interfaces [11]. This paper, on the other hand, focuses on the following research questions: Given a SAR whose functionality has been designed according to user-centered strategies, will this SAR be socially accepted by the residents? Will the user experience of interacting with the robot be positive?

In this paper, Sect. 2 provides an insight into the current research advances both in Socially Assistive Robots and Human–Robot Interaction fields, regarding technology evaluation from both technical and social perspectives. Section 3 describes the methodology and methods employed in the evaluation of this paper, the robotic platforms, and information about the recruitment of participants at the retirement home.

Two different tasks were defined for this long-term evaluation, described in Sect. 4. The interfaces employed for the robots to interact with the residents in these use cases are detailed in Sect. 5. The evaluation used two different robotic systems with different physical features, to compare how these features influence results. Both implemented the same tasks, following the same protocols and using accessible and



**Fig. 1** The Socially Assistive Robots employed in this study. The GoBe robots are on the left, the CLARA robots on the right

usable robot interfaces according to the users' needs. Figure 1 shows the two types of robots employed. One of them (the GoBe robot) is a commercial prototype of a telepresence robot. The other (the CLARA robot) has been specifically designed as a SAR to be used in previous projects [12].

Section 7 shows the main outcomes related to the interaction between the residents and the robots, providing information on the accessibility and usability of the robot interfaces, as well as the user experience when interacting with the robots and the social acceptance of the robots in the retirement home. Sections 8 and 9 discuss the main results and conclude the paper.

## 2 State of the Art

Since the very first implementations of Socially Assistive Robots in particular, and social robots in general, the consideration of the user perspective has been clearly established as a key requirement for defining the appearance and characteristics of these agents [4, 13]. As a result, SAR today, inspired by user opinion in the best implementations and, at least, following well-known acceptable cultural formulas taken from cartoons, films or others, usually attract positive attention. They are perceived as promising devices, and tend to make good or very good first impressions when exposed to untrained users. In fact, it is common for these devices to generate emotional bonds from the user to the robot, an event that the SAR can foster by mimicking emotional responses [14]. These bonds can make the SAR become a useful tool in certain tasks, such as rehabilitation procedures [15], therapies for autism spectrum disorder [16, 17], or for people with dementia [18]. The appearance and consolidation of these bonds also raises important ethical discussions [8, 10, 14]. These discussions are not always able to keep pace with

the rapid spread of SAR, although strong warnings against this dissociation continue to be published [10, 19].

Among the application contexts available for SAR, retirement homes and day care centers are receiving a growing interest in recent years [20–22]. SAR in these scenarios aim at assisting caregivers with simple and repetitive tasks, so that these professionals have more time to provide personalized meaningful care. In fact, SAR have been considered for years as a promising technology for assisting elderly and dependant people [7]. However, these robots still face important issues that prevent them from fulfilling these perspectives [21, 23].

From a technical point of view, SARs struggle to adapt their behaviour to their interlocutors, even if their perception systems are able to robustly acquire data from typically cluttered, noisy environments [24]. These adaptation problems limit their ability to exhibit social behaviour, which is the most important feature these agents need to incorporate. Furthermore, in these assistive contexts, Human–Robot Interaction (HRI) processes have to consider the variable, usually reduced, interaction capabilities of elders and their relatives. Cultural changes between regions also force an adaptation of the HRI protocols to each specific case [25]. Finally, retirement homes and day care centers are a challenging scenario for an autonomous or semi-autonomous agent, in terms of user safety and comfort [21].

From a user perspective, recent studies show that when SARs move from controlled environments to everyday unconstrained contexts, initial positive impressions usually do not evolve into long term acceptance and utility [10, 26], even if the robot behaves technically correctly [24]. On the contrary, once the novelty effect fades out, the false expectations created, the lack of real utility, or the absence of any added value compared to other solutions may dramatically affect the user perception of the SAR and prevent its use [7]. Some of the reasons for these problems are related to the fact that SAR experiments are traditionally affected by short duration, biased selection of participants, and incorrect experimentation procedures [7, 19, 23]. But another key factor is that the user centered approach followed to design the appearance of the robot usually does not extend to the definition of its functional properties, nor to the tasks it should perform [10].

Recent work aims to overcome these issues, mainly by (i) approaching the research, design and development on socially assistive robots from a multidisciplinary point of view; (ii) involving users in the design process from the very beginning; and (iii) testing SAR in everyday settings, in long-term experiments where user feedback allows adapting the final solution to their needs [27–30]. To meet these requirements in SAR experimentation, it is necessary to consider user centered design approaches [31] and co-creative and participatory methodologies [32–35], that are able to

incorporate the social dimension along with the technical one. Moreover, development and evaluation frameworks [36] become an interesting asset to evaluate the use of SAR in everyday environments. One example of these frameworks is the recently proposed AUSUS evaluation framework [22], employed in this paper.

## 3 Methodology and Methods

### 3.1 User Centered and Participatory Design

The methodology and methods used during the research work presented in this article follow an extended User Centered Design (UCD) approach. This approach combines iterative, user-centred design [31] and participatory design [34] to perform the design of the use cases, and the evaluation of the results. The methodology and methods of the extended approach followed are detailed in [31] and the details of how this approach was used to design the use case and interfaces for the evaluation described in this paper are detailed in [11]. Once the use cases and interfaces have been implemented in the robotic platforms, this paper focuses on another phase of the design process: the evaluation of the residents' acceptance and their experience interacting with the robot over a long period of time in a retirement home.

The residents provide during the evaluation procedure detailed in this paper a valuable feedback, that will be used to improve the robots and use cases in the future, and in short, the future experience of residents if SAR were fully integrated into retirement homes. They are therefore participating in the final design of the interfaces and services provided by the robots, and are empowered by having a say in the important decisions about the future technology in their retirement home, in contrast to the common situation of not taking part in the design activities, which can lead to an exclusion feeling [37].

### 3.2 AUSUS Framework

The AUSUS framework [22] (Fig. 2) was used to evaluate the performance of the robotic platforms in the retirement home. AUSUS takes into account different factors related to Human–Robot Interaction (HRI), such as accessibility and usability, but it also includes factors related to user experience while interacting with the robot, and social factors such as the social acceptance and societal impact. The study presented in this paper focuses on Accessibility, Usability, Social acceptance and User Experience.

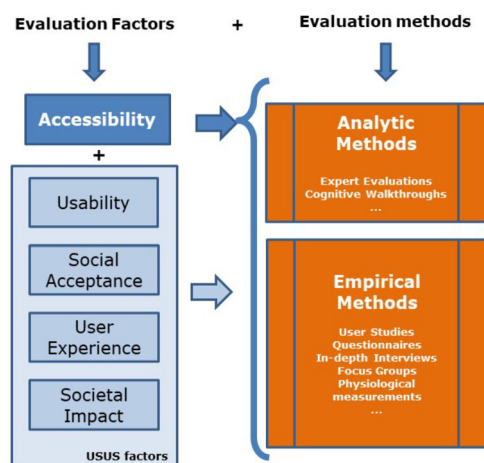


Fig. 2 AUSUS framework's evaluation factors and methods [22]

### 3.3 Robotic Platforms

Figure 1 shows the two type of robots employed in the paper. Their appearance and features are very different, as they have been designed for different purposes. However, they run the same cognitive architecture. They also execute the same use cases, with only minor implementation differences that only affect how the technicians control them.

CLARA robot [38] is a socially assistive robot initially developed within the framework of the CLARC EU project ECHORD++ (FP7-ICT-601116) to perform Comprehensive Geriatric Assessment (CGA) procedures [39]. It is equipped with a differential base, a tactile screen, speakers, a Sick 2D lidar, and one RGB-D camera. This camera complements the lidar sensor for navigation purposes, and allows the detection of objects and people in the environment. The robot also has a safety bumper that stops it if the robot contacts with an obstacle.

CLARA robot uses a navigation stack, programmed on the Robot Operating System (ROS) framework [40], that allows it to navigate autonomously in unconstrained environments, safely avoiding obstacles. Thus, during the use cases the robot simply sets random way points located within predefined areas (e.g. 'announcing area in room 1') to move around the retirement home.

CLARA robot has been provided with an *ad-hoc* video call system specifically implemented for this robot using web technologies. This system uses two external servers but it does not depend on third-party providers. As a result, relatives only need a web browser in order to reserve time slots and establish the video call on their endpoint. Relatives provide an email address during the reservation procedure, so the robot can send a reminder email containing the reservation time, the instructions needed to make the video call

connection, and a security code generated individually for each video call.

The GoBe robot, developed by Blue Ocean Robotics, is a telepresence robot.<sup>1</sup> Its main features are a large tactile screen, a pair of loudspeakers specifically designed for human–robot voice interaction, and a wide-angle camera that gives the user a better view of the environment. These features define GoBe as a platform specifically suited to the video calling use case. The robot is also equipped with a laser sensor that allows it to avoid collisions, by gradually reducing its speed as it approaches an obstacle.

The GoBe robot does not move autonomously, but it is designed to be controlled remotely. The remote user employs a web based interface to move and maneuver the robot in different environments. The same interface provides video and audio communication between the remote user and the people around the robot (i.e. telepresence). For the use cases studied in this paper, this telepresence interface is used to make the video calls. These interfaces are provided by the manufacturer and have not been modified for this study.

Both robots use an instance of the CORTEX [41] cognitive architecture. The technical details about these architectures lie beyond the scope of this paper. Briefly, they contain the information of the context as an inner representation, that the different components of the robot use to share data and synchronize during the execution of the use case. These components also allow the architecture to interchange data with motors, sensors, screens and external servers. As detailed above, the only component of this architecture that differs from CLARA to GoBe robot is the navigation stack.

### 3.4 Inclusion and Exclusion Criteria

The residents included in the study must have a certain cognitive capacity. The threshold was determined following both the Montreal Cognitive Assessment (MoCA) and the Lobo's Minimental scale, which are the ones used by the psychology department of Vitalia Home retirement home. According to the Lobo's Minimental scale, for geriatric people, score equal or higher to 23 is considered normal [42].

The MoCA is brief, simple, and reliable as a screening test for Alzheimer's disease, and is used along other test (in this case the Lobo's Minimental scale). A score of 26 or higher of 30 is considered normal [43].

Regarding other capabilities (sensory, emotional or motor), the residents interacting with the robot were only required to be able to hear and/or see the displayed texts. No more constraints were imposed. The diversity that brings this broad inclusion criteria aligns with the following goal: to design a SAR able to adapt to as many users as possible. As pointed

<sup>1</sup> <https://gobe.blue-ocean-robotics.com/>.

out in [37, 44], older people are as different as everyone else, or even more, since ageing meets us in very different ways.

### 3.5 Recruitment of Participants

The two evaluation groups were recruited from residents at the Vitalia Teatinos retirement home. All the participants were elderly citizens (age between 65 and 85) who fulfilled the inclusion criteria previously detailed. Table 1 details the physical and cognitive characteristics of the users. The level of impairment is encoded using the following 4 levels:

- None. The participant does not have any impairment.
- Mild. The participant has some level of impairment, but only needs assistance tools in specific circumstances (e.g. glasses for reading).
- Medium. A notable impairment which requires an assistive tool.
- Severe. Almost total impairment which not only requires an assistive tool, but also depends in other people to do daily activities.

Table 2 describes the previous technological experience of each user. Three levels are employed for this description:

- None. The user is unable to use, or knows nothing, about any technological device.
- Little experience. The user is able to use some technological devices, but always with some assistance.
- Some experience. The user is capable of use the basic functions of any device by herself.

Regarding the professionals of the retirement home, they were not intrinsically participating in the evaluation sessions, that focused on the experience of residents. However, this evaluation can benefit from being supplemented by the experience of others [37, 44]. Hence, the professionals provide their perspective about the use of a SAR in their retirement home, through the arrangement of one formal meeting during the evaluation period and casual conversations throughout the sessions. The feedback collected in these activities allows identifying the "standing up" dilemma discussed in Sect. 8.2.

The general characteristics of the study are summarized in the following list:

- Number of participants: 14
- Timeframe: 11:00–13:00.
- Duration of individual interviews: 10 min.
- Duration of focus group sessions: 45 min.
- Duration of observation (per participant): 5 min.
- Method of data collection: Observation, questionnaires and interviews.

## 4 Use Case Specifications

This study focuses on two different use cases performed by the robots during the evaluation sessions. These use cases were previously co-defined in participatory workshops by the retirement home staff, residents and researchers [11]. The use cases were not modified for or during the evaluation presented in this paper. The use cases are the "town crier" (robot announcer) and the "video call" (videoconferencing with family and friends).

### 4.1 "Town Crier" Use Case

In this use case, the SAR navigates around the retirement home and announces relevant information to the residents, such as the daily menu, upcoming birthdays, or upcoming events. It is also able to display this information on demand, when users tap on the icons on its touch screen. The robot starts this use case from its charging station, and returns to it when the use case finishes.

For the experiments discussed in this paper, the SAR can provide the following information:

- Daily menu.
- Current time and date.
- Current weather forecast and temperature.
- Upcoming events.
- Upcoming birthdays.

In order to adapt to the capabilities and abilities of the residents, being older persons which a great variety of functional disabilities, the interfaces were designed following the principles of Universal Design and according to the accessibility guidelines for Human–Robot Interaction [45]. Therefore, the information is presented in a multimodal way: spoken aloud and displayed as subtitles on the SAR screen. If a resident wants to access any particular information, she can tap on the appropriate area on the screen mounted on the robot (see Sect. 5 for more details). This action triggers the corresponding announcement.

For the tests evaluated in this paper, the robot was programmed to announce the daily menu twice, just before lunch (announcements at 12:00 and 12:30) and dinner (announcements made at 19:00 and 19:30). The tactile screen was not in use during these programmed announcements, but was available at all other times, allowing residents to receive information on demand (Fig. 3).

### 4.2 Video Calling Use Case

The main objective of this use case is to enable residents to communicate with their relatives via video calls. The implementation of this use case was mainly motivated by the

**Table 1** User's characteristics

User ID	Gender	Hearing impairments	Visual impair	Motor impair	Cognitive impair
User 1*	Male	Mild	Mild	Medium	None
User 2**	Male	Medium	Mild	Severe	None
User 3**	Female	Mild	Mild	Mild	None
User 4	Male	Medium	Medium	Medium	None
User 5*	Male	None	Medium	None	None
User 6	Female	None	Medium	Severe	None
User 7	Female	None	None	Medium	None
User 8*	Female	Mild	None	Medium	None
User 9**	Male	None	Medium	Severe	None
User 10	Female	Severe	Mild	Mild	None
User 11**	Male	None	Medium	Mild	Medium
User 12	Female	None	Medium	Mild	None
User 13**	Male	None	Medium	Medium	None
User 14	Female	Medium	Mild	None	None

\*The user did not fill the final comparative test

\*\*The user did not complete the evaluation process

**Table 2** Previous technological experience of the users

User ID	New technology aptitudes	Robot interaction	Video call experience
User 1*	None	None	None
User 2**	None	None	None
User 3**	None	None	None
User 4	Computer (at work)	None	None
User 5*	Cell phone, computer (Social media, news)	None	None
User 6	Cell phone (internet browsing)	None	None
User 7	Computer (e-mail)	None	None
User 8*	Cell phone (social media)	None	Little exp
User 9**	Cell phone (watching videos, social media)	None	Some exp
User 10	None	None	None
User 11**	Computer (writing)	None	None
User 12	None	None	Little exp
User 13**	Computer(at work)	None	Little exp
User 14	Computer(writing)	None	Some exp

\*The user did not fill the final comparative test

\*\*The user did not complete the evaluation process

isolation imposed by the COVID19 pandemic: visits were not allowed in the retirement home. Although the pandemic restrictions are no longer in place (except for the wearing of face masks), this service is still requested by residents who want another way to communicate with relatives and friends living abroad.

In this use case, the robot is able to autonomously start a video call with the resident on a day and time that are previously selected by the resident's relative through a booking process. Once a time slot has been reserved, and a few minutes before the video call is due to start, an email is sent to the relative with instructions on how to make the video call.

The SAR then begins to navigate from its charging station to the designated points in the common areas of the retirement home, where it announces the video call using both voice and text. The announced information can be adjusted to the preferences of the users (i.e. they may prefer the robot not to provide certain details in the announcements). In this particular implementation, following the preferences of the residents, the announce includes:

- The name and surname of the resident.
- The name of the relative, along with the type of relationship between the two.



**Fig. 3** Resident tapping in the robot screen during the town crier use case

- The common room where the video call will set up.
- The time at which the video call starts.

Once the announcements have been made, the robot moves to the location where the video call will take place. When it reaches this position, it waits for a person to be detected in front of the robot. Once the resident is detected, the video call starts. Finally, when the call is over, the robot says goodbye to the resident and returns to its charging station.

## 5 Human–Robot Interfaces

The two robots employ the same interface to interact with residents in the "Town-crier" use case. The only difference is the screen configuration, that is vertical in the GoBe and horizontal in the CLARA. The interface adapts dynamically to this change. The booking and video call interfaces, however, are different. CLARA robot uses the ones specifically designed for this study, while GoBe robot employs the ones provided by its manufacturer. The robots also included teleoperation and supervision interfaces for the technicians. These last interfaces are not evaluated in this paper.

### 5.1 "Town-Crier" Interface

Figure 4 represents the main view in the tactile screen installed in the SAR for the town crier use case. This interface has been implemented following accessibility and usability criteria. The user can select the following options in this interface by tapping on them:

- Calendar. Information about the current day in the calendar.
- Weather information, obtained from an Internet service.



**Fig. 4** "Town-crier" use case interface in SAR screen: **a** Main screen; **b** Showing text while speaking

- Activities. Scheduled activities and events for the current day.
- Birthdays. Upcoming residents' birthdays.

Additionally, this interface will automatically popup subtitles with the text of the sentences that are being spoken at that moment by the SAR (Fig. 4b).

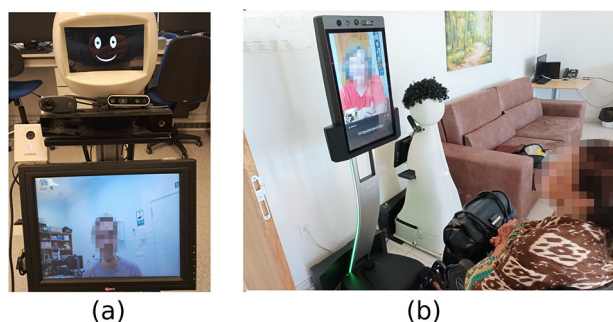
## 5.2 Video Call Interfaces

The two robots use a different interface for the relatives or friends performing the video call. It was not possible during the experiment to properly collect feedback from these users: they did not have time to conduct interviews, and the prototype status of the robots -specially the GoBe robot- required a strong technical supervision for these users. Hence, the interfaces discussed here are only the ones employed by the residents during the video calls. Figure 5 shows these interfaces. CLARA robot uses a full-screen interface within a web browser, while the GoBe robot employs a proprietary interface very similar, but including a space under the video stream to write messages. For the conducted experiment, this space only displayed the name of the relative to avoid evaluating different functionalities.

## 6 Experimental Set-up

### 6.1 Environment

Experiments presented in this paper started in November 15th, 2021. The first robot deployed in the retirement home was the CLARA robot. But after one month, the retirement



**Fig. 5** Video call interfaces for the residents: **a** CLARA robot; **b** GoBe robot

home forbade the entrance to all visitors due to a COVID-19 wave that lasted until March, 2022. Then, a second evaluation period took place during the next four months. The GoBe robot was deployed at the beginning of this second period, in March 2022. The experiments ended in July, 21th, 2022. The robots were used in the retirement home for five months. The agenda of the retirement home and the availability of residents made it possible to schedule 1–2 sessions per week. A total of 24 sessions were carried out. The appearance, features and functionality of the robots were set following an UCD procedure before the evaluation period [11]. They did not change during the experiments described in this paper.

The experiment was set up in the Vitalia Teatinos retirement home, located in Málaga, Spain. Due to access restrictions resulting from the COVID19 pandemic situation, in order to prevent the spread of the virus, the evaluation sessions were finally conducted in just one room (Fig. 6). This multi-purpose room is where residents watch TV, meet for a chat, read, take a nap, or have lunch, among other activities. The room is equipped with furniture such as tables, chairs, armchairs, cabinets and a sofa (Figs. 6 and 7). Some of these items are usually moved to different positions, according to the preferences of residents and the activities performed in the room. The robots had to navigate avoiding (autonomously, in the case of the CLARA robot) all these obstacles. Moreover, there were residents, and occasional workers, who moved around the room walking, using a walker or wheelchair, or pushing trolleys with medication or food.

The charging bases of both robots (Fig. 8) were also located in the room, in a place that avoided interfering with the daily activities in the retirement home.

During the tests, it was mandatory to follow the protocol of the retirement home (wearing masks, checking body temperature, frequent hand cleaning, etc.). Furthermore, all the residents are vulnerable population. Hence, each time there was a rise in COVID-19 cases, all visitors (including the researchers) were forbidden to enter the retirement home. These circumstances negatively affected the number of sessions that were able to be conducted.



**Fig. 6** Common room



**Fig. 7** Common room



**Fig. 8** Common room. SAR's charge base

It was important to avoid any kind of distraction during the evaluation sessions, as the only people staying in the common room during those sessions were the residents who voluntarily enrolled in the evaluation and one of the researchers. This technician was in charge of setting up and monitoring the evaluation session, and guaranteeing safe operation during the experiments. However, as commented above, it was not possible to completely avoid other residents or staff members to occasionally go around the room during the sessions.

## 6.2 Evaluation Procedure

The user-centered evaluation applied to this study was based on A/B testing [46]. We created two different groups and users were randomly assigned to one of these groups. The first group, named Group A, started the testing procedure using the CLARA robot. Group B started testing the GoBe robot. During the evaluation of the "town crier" use case, participants were given a situation where they had to pretend that they needed information provided by the SAR, so they had to interact with it. Regarding the evaluation of the video call use case, participants only needed to sit in front of the robot and have a conversation with another person.

After several weeks of interacting with the SARs, the participants in each group were asked to fill a user questionnaire. There they could share their impressions of the SAR they interacted with. Then, by the end of the evaluation period, both groups switched SARs, so the participants could interact with the other platform. Finally, a comparative questionnaire was filled again by both groups. The goal of this questionnaire was to collect the different impressions about both SARs, and highlight the qualities that made each one distinct from the other in various aspects of the HRI procedure.

Finally, at the request of the participants and in order to facilitate communication during the interviews, the residents gave each robot a name. They chose the name "Felipe" for the CLARA robot, while the GoBe robot was named "Cipri" after a video call with one relative of the same name who was the first to use the video call function on the GoBe robot. Regarding this naming process, the CLARA robot was not only the first one the residents put a name on, but also the one whose naming procedure occurred more naturally, without any intervention of the researchers. In the case of the GoBe robot, the researchers had to ask the residents to name the robot.

The Group A was the first to start the evaluation, in November 15 2021. Due to a lock down imposed by the COVID-19 pandemic, it was not until March when Group B would start the evaluation with the GoBe robot. This detail is important due to the lapse of time between the deployment of both robots (the CLARA robot was in the retirement home 5 months before the GoBe robot).

The sessions were conducted following these steps:

1. The evaluator reminds the residents the itinerary of the robot.
2. The residents stay at the room as they usually do, depending on the hour. Hence, if the test is performed near lunch time, they sit around the tables. At mid-morning, some of them will usually be on the armchairs reading newspapers, while others may be at the sofa watching TV.
3. The socially assistive robot performs its task (town crier or video-conference).
4. After the use case, the evaluator interviews the residents to collect first impressions about the session. Additionally, some guided questions are also asked, related to the socially assistive robot and interaction details (volume, text on screen, robot speed, etc).

## 6.3 Material

Quantitative and qualitative methods were combined, to evaluate and compare both the performance of the robots and the feedback provided by the residents.

1. Questionnaires and structured interviews: User questionnaires and interviews were performed with the residents and staff in the retirement home, before and after interacting with the robotic platforms. Before the evaluation process, sociodemographic variables and technology use and skills (Table 2) were surveyed. After each session, questionnaires and structured interviews were conducted to measure subjective accessibility, usability criteria, user's experience and robot's acceptance. The test, questionnaire, and structured interviews contained questions on a 5-point Likert scale [47]. In this study, there were two main types of questionnaire: (1) a post-session interview, where general data about HRI is collected (Figs. 9, 10 and 11, where the X axis indicates number of residents); and (2) a comparative questionnaire, where general aspects of both robots were asked (Fig. 12), once both groups interacted with both robots. This comparative questionnaire is not a Likert scale. However, it keeps using 5-point items to describe preferences, so it could be easily filled by participants already used to these scales.
2. Observations: Direct observations of the HRI process were performed by the evaluator. The sessions were also recorded so the rest of experts working in this study could analyze these interactions. These observations complemented the data collected in the interviews (e.g. they allow measuring interaction times, or checking how each user tap the screen).

## 7 Evaluation of the Results

This section presents the results collected using the previously described AUSUS framework [22].

Regarding participants, although at the beginning of the study 14 residents volunteered, a decrease in participation occurred during the post-session interviews. The main reasons were:

- The main inclusion criteria was to have an adequate cognitive capacity. Some of the residents that meet this criteria in the retirement home were staying there only temporarily (e.g. because they were in a rehabilitation process). Users 9, 11 and 13 left the retirement home before the study ended. Thus, they were not able to complete the post test interviews.
- User 2 participated in the sessions, but he did not feel like doing the post-sessions interviews.
- Finally, user 3 left the study after one session because she did not agree with the study and stated that "this robot is going to substitute a human job".

Due to these circumstances, it was finally possible to collect the complete post-test data from 9 participants (Users 1, 4, 5, 6, 7 and 8 at Group A and users 10, 12 and 14 at Group B).

Finally, at the end of the evaluation period, both groups switched the robots that were used in the evaluations, and a comparative interview about both robots was done to the six participants who agreed to complete it (Users 4, 6 and 7 from group A and users 10, 12 and 14 from group B). Questions focused on which robot the user preferred depending on certain metrics (sound, image, appearance, utility, etc.).

### 7.1 Accessibility

#### 7.1.1 Perception

Regarding perception, no major issues were found related neither to the sound nor the images and text. However, beyond questionnaire results, subjectively all users positively appreciated the significantly larger screen and the higher quality sound of the GoBe robot.

#### 7.1.2 Operation

While all users could perceive the sounds and images provided by the robot, it was difficult for some of them to operate it. As stated in Table 1, 5 of the 14 participants had never used cell phones, computers, or any kind of technology, which is reported as one of the common challenges that older users face with welfare technology [37]. The rest of them were

familiar with at least some technologies. Only 4 of them use cell phones.

Participants who had no technological background, and those of them who had only used computers, found difficulties to operate the interfaces of both robots. The reason behind this issue is that there are strong similarities between the operation mode of the robots and the use of a cell phone, as both systems rely on a touchscreen as its main interface. As a result, only those participants who are used to cell phones showed a better attitude towards learning to use the robot. This is in line with the study described in [48], where the participants who owned a smartphone and a tablet were also those more interested in the use of modern technologies. Furthermore, this reduced group of participants in our experiments were able to follow the tasks given by the evaluator without any help by May (12 sessions after reopening the retirement home). More details about learnability are provided in its corresponding subsection.

#### 7.1.3 Understanding

GoBe robot is equipped with a larger screen and better quality speakers. However, as Fig. 10 shows, when both robots provide the same information, both robots are clearly understood.

### 7.2 Usability

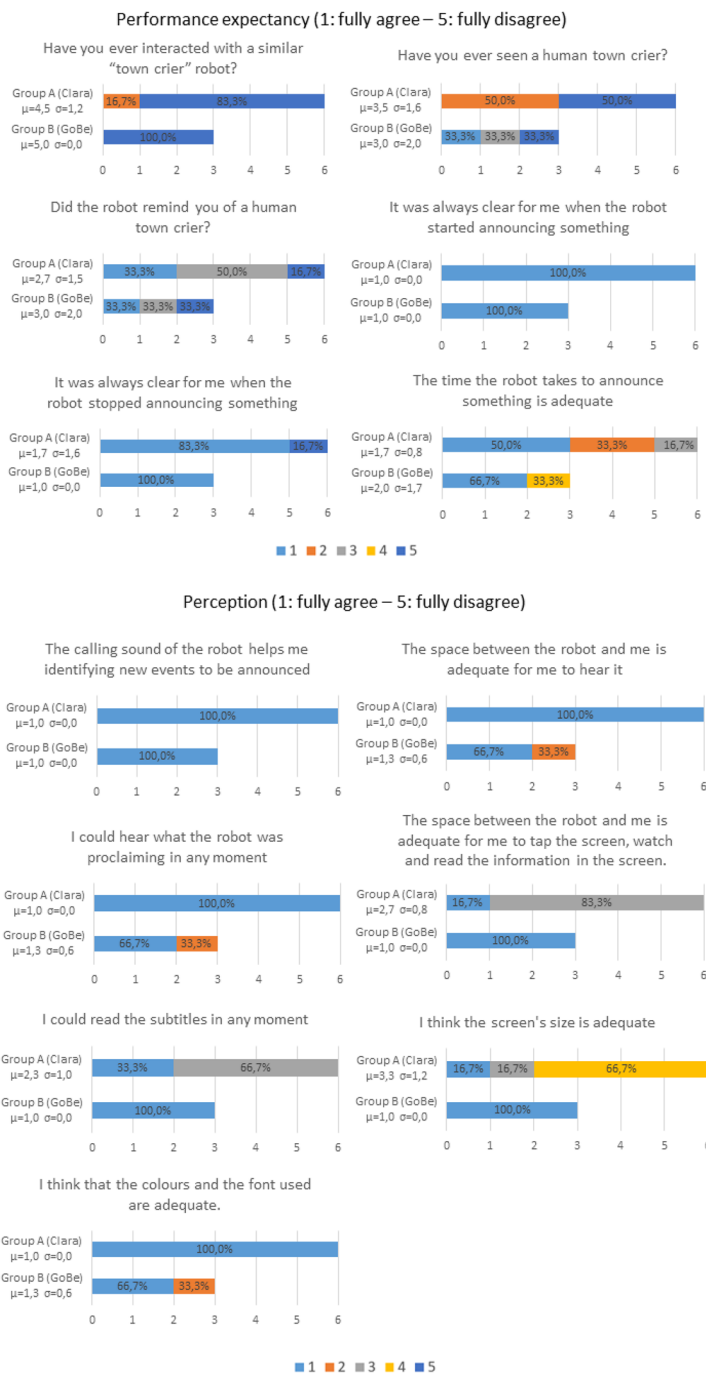
#### 7.2.1 Effectiveness

During the first 10 sessions with the CLARA robot and the first 8 with the GoBe, one evaluator monitored and, when necessary, gave clues to the participants. This was necessary because especially in the first sessions of interaction with the robot, the success rate of the residents in executing a given task was low. Nearly all of them required some clues and help selecting the options in the touch screen. This was because the residents did not initially know how to interact with the robot and very few had previous experience with digital interfaces similar to this one. However, the success rate was measured after each session and increased at the end of the evaluation: 83% of the participants who finished the study were able to interact with the robot, without any clue from the evaluator, at the end of the experiments.

#### 7.2.2 Learnability

The learnability factor measures how easy is for novice users to learn using a system [36]. As mentioned in the previous section, at the beginning of the sessions the residents needed a lot of hints from the technician in order to successfully get the information from the robot. However, as the study continued, these clues were reduced. After these aforemen-

**Fig. 9** Results for the individual evaluation of each robot (1). Questions are grouped by categories

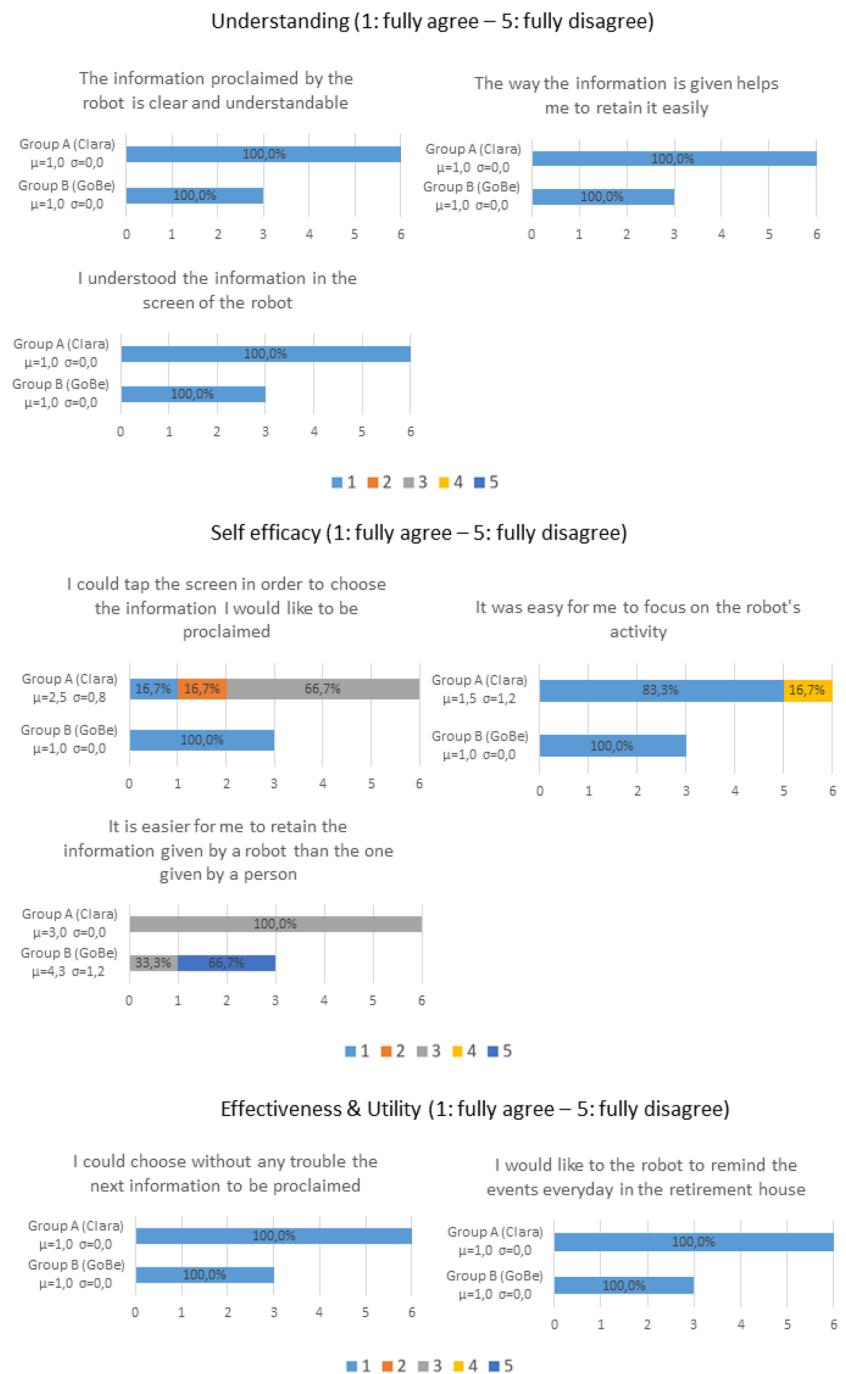


tioned sessions, around two months after the reopening of the retirement home due to COVID-19, 3 of the residents from Group A and 2 from Group B, who were the participants with a previous technological experience, were able to interact with the robot without further help. In the case of the rest of participants, the researcher needed to continue providing some clues in order to successfully conclude the task.

In this study, 9 of the 14 participants have previous experience with technologies (mainly with cell phones or computers), but none of them had interacted with a robot

before. Despite this, these 9 participants showed not only a better attitude towards using the robot, but also to learn about how to use it. This is in line with [48] and consistent with [49], where it is concluded that a positive transfer from a previous experience from learning technology or from systems that are quite similar makes the new learning process easier. Another example can be found in [37] where it is reported an user that, without a previous technology experience, didn't want to test in front of others as she felt that her learning curve didn't evolve at the same pace as others. By May (two

**Fig. 10** Results for the individual evaluation of each robot (2). Questions are grouped by categories

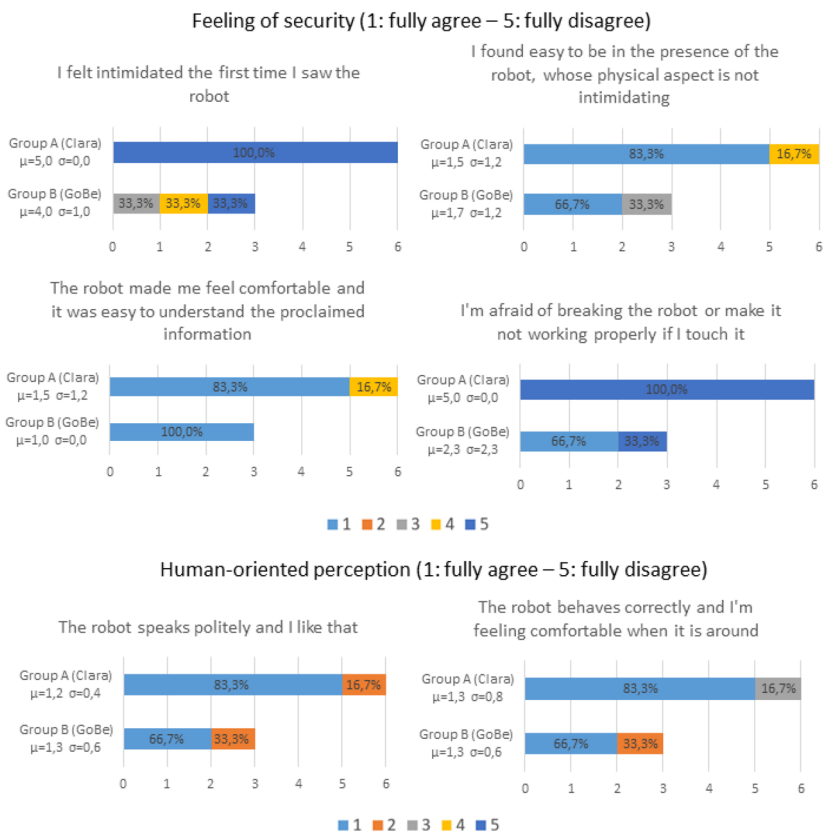


months after the re-open of the retirement home due to the COVID-19 wave), the evaluator did not need to tell these 9 participants how to interact with the robot. Instead, those participants were just asked questions like "What day is today?" or "How is the weather today?", and they got the answer tapping the adequate section on the screen.

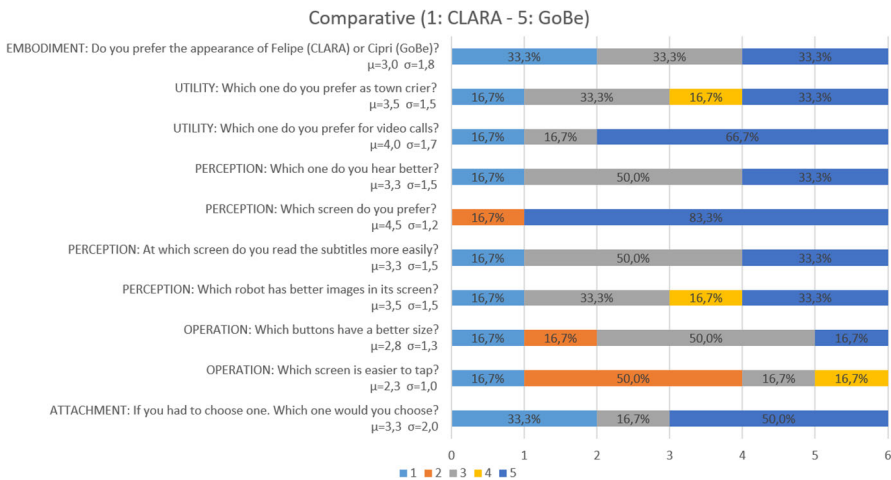
### 7.2.3 Efficiency

Two months after the reopening of the retirement home due to COVID-19, the average time participants took to successfully fulfill the task of the "town crier" use case in Group A was 10s, whereas for Group B was 17.57s. This difference is not caused by the interaction process itself (both robots had the same user interfaces implemented in their screens), but by the different heights of the position of the screen. Most participants were sat on an armchair or their wheelchair when

**Fig. 11** Results for the individual evaluation of each robot (3). Questions are grouped by categories



**Fig. 12** Results for the comparative post-test questionnaire. Each question is preceded by the category it is related to



interacting with the robot, so for the GoBe robot they had to get up to tap the screen.

**7.2.4 Flexibility**

Regarding flexibility, the users had only one input channel with the robot (the touchscreen) to ask for specific information. However, the robot has several ways to communicate with the users: voice, icons and subtitles, adapting to the user’s needs. In the next versions of the robots, Automatic

Speech Recognition is planned to be included to facilitate the interaction with the robot via voice.

**7.2.5 Robustness**

Both robots were able to perform the use cases adequately in the selected locations. While the ability to navigate autonomously became a particularly robust asset for the CLARA robot, in the experiments presented in this paper the autonomous motion followed very similar paths to the pre-

viously teleoperated motion. The residents did not perceive any change in the behaviour of the robot.

Regarding the interface, for some of the residents the tactile screen was difficult to handle. They had to learn how to correctly tap it, considering the sensitivity of the device. This issue caused some discomfort, specially when the residents have to approach or stand up to touch the screen.

Furthermore, the robustness concept is also related to a proper functionality in the interface with the robot and with the possibility of using that interface to ask for help, if required. In the robots employed in this study, the latter function was not included in the interface, mainly because there were always a technician close to the experiment. However, its implementation is totally feasible for future iterations of the evaluation.

### 7.2.6 Utility

In the comparative questionnaire mentioned at the beginning of this section, some of the questions were related to which robot (CLARA or GoBe) was more useful for each use case. For this comparison, a score of 5 means that the user clearly preferred the GoBe robot, while a score of 1 means that the CLARA robot was preferred. In the "town crier" use case, the results were as follows:

- Two of the participants definitely preferred the GoBe robot (score: 5), because it has a larger and easier to use touchscreen.
- One participant slightly preferred the GoBe robot over the CLARA one (score: 4). She mentioned that she would have chosen the 5 value if the GoBe screen was in a lower height.
- Two participants felt indifferent about which robot is more useful (score: 3 in the Likert scale).
- Lastly, one participant preferred the CLARA robot (score: 1). She felt more emotionally attached to that robot due to its appearance, that resembled a person for her.

The general consensus was balanced between both robots, leaning slightly towards the GoBe robot, with an average score of 3.5.

In the "video call" use case, the main results were as follows:

- Four of the participants definitely preferred the GoBe robot (score: 5). The participants explained that "it has a larger screen" and "it is simpler to use".
- One of the participants felt indifferent (score: 3).
- The same participant that preferred the CLARA robot in the "town crier" use case also preferred it in the video call

use case (score: 1). As before, she felt more emotionally attached due to its appearance.

The general consensus leaned towards the GoBe robot, with an average score of 4.0. These results are logical due to the GoBe robot being originally designed for this use case specifically. This is in line with the study in [48] regarding the appearance of a robot versus the functionality, where it is concluded that the participants agreed that functionality is more important than appearance based on participants interviews. However, it must be taken into account that this may depend on the use case as stated in [50], and for example in stressful or complex situations, more human-like robots could be preferred [51].

## 7.3 Social Acceptance

All the participants, except one at Group A, felt that the behaviour of the robots were socially adequate in every moment. Furthermore, all the participants also felt that the robots were polite from a social standpoint.

As mentioned in the Experimental Setup section, participants belonging to Group A showed more initiative naming the robot, while in Group B the researchers needed to suggest that task to the participants.

### 7.3.1 Performance Expectancy

Figure 9 shows that none of the participants (except one belonging to Group A) have seen a robot before. Furthermore, not all participants have ever seen either a real town crier, so the overall performance expectancy of the robot was not high.

Regarding the video call use case, and particularly with the GoBe robot, the perform expectancy were not high due to video calls being more common in the retirement home. However, this led to an interesting result, which was reflected in two comments given by the participants:

- "Thanks to the robot I can see my relatives without masks".
- "Although it is not a personal visit, at least I can see more relatives than usual" (NOTE: Due to COVID-19, personal visit where limited to one relative per resident)

### 7.3.2 Effort Expectancy and Self-efficacy

Due to the absence of any previous experience with robots for all participants, there was a high effort expectancy. Furthermore, during the recruiting period, most of the possible participants who did not enroll the test group argued that "it is very difficult to use the robot", even though they have never

interacted with any. These comments agree with the conclusions reached in [49], where the authors conclude that both the social environment or the self-perception may influence the use of technology, and for example the idea that people are too old to learn will negatively affect it. However, by the end of the evaluation period, most participants agreed that using the robot was not that difficult as they expected.

Figure 10 shows that, although both groups found easy to focus on the robot (except for one person in group A), in Group B, participants believed that they would remember more information if it was given by a human instead of a robot while interacting with the GoBe robot. In the case of Group A, they felt indifferent whether the information was given by a robot or a human when they interacted with the CLARA robot (all of them gave a score of 3 to this question).

### 7.3.3 Attitude Towards Using Technology

This item is related with other results presented throughout this section. Participants who are used to technologies, such as computer and cell phones, showed not only a better understanding of the usage of the robot, but also a better attitude towards using it. Furthermore, these participants were more efficient performing the designated tasks in the use cases, and provided more nurturing feedback.

### 7.3.4 Forms of Grouping

This study did not lasted enough to let the participants share some kind of identity with the robot, so this subsection will be left for future iterations of this research.

### 7.3.5 Attachment

Figure 12 shows the comparative evaluation results. They show that, although both are almost equally appreciated in the retirement home, the GoBe robot was slightly preferred (average score of 3.3) due to its more robust design and better equipment.

As mentioned before, the CLARA robot was the first one which was brought to the retirement home, thus residents had more time to develop a feeling of attachment. The appearance of CLARA robot also helped fostering this feeling. Despite these facts, the result presented here shows that, at the end, a more robust design and higher quality equipment compensated both the time advantage and the "endearing" factor, leading to an equal appreciation for both robots (as stated in Fig. 12).

### 7.3.6 Reciprocity

Regarding the reciprocity, there is only one issue that occurred during several sessions of the CLARA robot. Some

participants wanted to tap one of the buttons of the touch-screen, but the robot did not react. These participants felt frustrated due to this lack of response. This issue is produced by the limited sensitivity of the touch screen of CLARA robot. For future tests either the sensitivity will be adjusted or a new touch screen will be used.

## 7.4 User Experience

### 7.4.1 Embodiment

The analysis of the effects of the embodiment in the perception of the robot was one of the main goals in this study. The two platforms were analyzed in parallel following the A/B testing approach, as stated above. As displayed in Fig. 12, the results for the question about embodiment are equally distributed. Two of the participants preferred the appearance of the CLARA robot because it was "cute" and "endearing", due to its more humanoid form. On the other hand, other two participants preferred the GoBe robot's embodiment, due to its more robust design, giving them a sense of security.

### 7.4.2 Emotion

Regarding the CLARA robot, none of the users were afraid to touch it. The main emotions that produced this robot on the users, on their own words, were: "joy", "satisfaction" (as the robot implies an advance in society) and "excitement". Only one of the six users felt emotionally indifferent toward the robot.

On the other hand, when interacting with the GoBe robot, two of the residents agreed that they felt afraid of breaking the robot at the beginning of the experiments. The third user was not afraid.

In both cases, the emotions of the users after interacting with the robot were overall positive. They used expressions as: "excitement", "admiration" and "I really liked the experience".

### 7.4.3 Human-Oriented Perception

The only outcome for this subsection is the one displayed in Fig. 11, which shows that the behaviour of the two robots was considered socially adequate (with average values of 1.2 and 1.3 in these items). Following social conventions and expectations is an important feature to avoid stressful human robot interaction [52]. While there was a transference of human features to the robots -e.g. they were named-, the perception of them as artificial devices was always present: "the video call has been good, but I would have preferred my daughter to be here".

#### 7.4.4 Feeling of Security

No major issue was found regarding this item. Only one user on Group A and another from Group B felt certain discomfort when both robots were around, stating that they were intimidated by their appearances, although both robots' movements were slow and stop moving as soon as an obstacle was in the way. Nevertheless, these participants did not explained with any details why they did not feel secure around the robots.

#### 7.4.5 Co-experience

These results were not collected from questionnaires, but from field experience (observations). The evaluation period was not long enough to guarantee a complete removal of the novelty factor. However, it is important to highlight one phenomena that occurred during the last month of the evaluation: when the evaluator started the evaluation sessions and proceeded to call the participants, they began to call other residents (even though they did not belong to the proper study group). They wanted them to join the experiment and see the robot working. The experiments started conversations about the robots, that extended well beyond the sessions themselves. Given the limited duration and number of participants involved in the experiments, further studies should be conducted to clarify the reasons under these behaviours. Previous research [53, 54] suggest these results are produced by the SAR acting as a social facilitator.

### 7.5 Societal Impact

The evaluation of the quality of life, working conditions and employment indicators will be reported in a different paper of one of our colleagues [55]. Both Education and Cultural context were not evaluated due to being indicators which were not centered on the objectives of this study.

## 8 Discussion

Although there were not so many volunteers in this project, two interesting queries arose regarding the Human–Robot Interaction process. One of these questions was proposed before the study (*Which robot is easier to interact with?*). The other emerged during one of the focus group sessions with the clinics (*Is it correct that the residents have to stand up to interact with the robots?*).

### 8.1 Which Robot is Easier to Interact with? GoBe vs CLARA

As previously detailed, a comparison questionnaire between the two robots was passed to all participants who con-

cluded the experiment. These participants were interviewed after having tested both robots. Figure 12 shows the results collected in these questionnaires. These results can be complemented with the following qualitative feedback:

- For the town crier use case, 3 participants preferred the GoBe robot, mainly due to the larger screen. Only one participant preferred the CLARA robot, because she felt more attached to it. The average score for this question was 3.5.
- For the video call use case, 4 participants clearly preferred the GoBe robot. Again, the main reason for this decision was its larger screen. Also, one of these participants mentioned that the GoBe robot was simpler to use (GoBe is completely teleoperated). One of the participants felt indifferent, and the other preferred the CLARA robot, again due to her stronger emotional attachment to it.
- 5 of the 6 participants preferred the screen of the GoBe robot (average score of 4.5), mainly because of the larger size. Only one participant slightly preferred the screen of the CLARA robot, because it was better aligned with her line of sight.
- 3 participants did not notice the subtitles.

The last question of the comparative questionnaire made the users choose about which robot would they keep. This question was also asked during the last weeks of the experiments, where both robots were in the same room. At first, none of the robots was preferred over the other. However, at the end of the evaluation period, when the comparative questionnaire was passed, the results slightly leaned towards choosing the GoBe robot, since 3 participants voted it with a 5 score. The reasons behind their decision were its better audio and video interfaces. These participants also stated that it has a prettier design. On the other hand, two participants chose the CLARA robot (voting with a 1 score), because they felt more emotionally attached to it. They also described CLARA as "easier to approach", both physically and from an interaction perspective. Only one participant felt indifferent about which robot to keep.

Furthermore, participants were given the opportunity of commenting freely about any of the robots or the research. The following are representative answers:

- "Felipe could be a little higher"
- "I like them"
- "I like both robots. I would love that they could give us the medicine"
- "Due to my situation, robots are not part of my worries"
- "This represents a major advance"
- "Both robots are really well designed, especially Cipri"

The group of participants was not large enough to jump to any solid conclusions. However, from these results it seems that, in the long term, functionality (i.e. the better interfaces of GoBe robot) and a more robust design, which gave a sense of security in most participants, prevailed against the emotional attachment (that CLARA robot generated more easily) for the designed use cases. However, these results also show that a more humanoid (or "endearing" as some participants described) embodiment of the robot not also raised attachment feelings among the participants, but it also facilitated the interaction in a certain degree. This should be taken into account if the intention were to provide a companion robot. Actually, in [48] it is reported that companion robots such as an AIBO were not interesting enough as they were "nothing to cuddle with".

The GoBe robot gave a sense of security, but its design also overwhelmed the participants, leading to a lack of incentive to interact with it. On the other hand, the CLARA robot became a focus of interaction for participants and non-participants in the retirement home, even when the robot was turned off. These interactions off the tests were usually performed as jokes. Some residents liked to talk to the robot just for fun, even if it was turned off. While this situation had not a real effect in the programmed use cases, it gave a good insight of how the robot can facilitate socialization. However, it will be important to analyze how the functionality of the robot changes this effect against the one produced, for example, by a fixed puppet. This analysis will require a different set of experiments, to be performed in future projects.

## 8.2 The "Standing Up" Dilemma

One of the main discussions that arose during the evaluation was what we would like to provisionally call the "Standing up" dilemma. This predicament consists on whether is correct to: (i) totally facilitate the interaction for the users; or (ii) leave some degree of difficulty to them, in order to encourage a more pro-active HRI.

In this study, the dilemma came up during the focus group session with the clinics held on 14<sup>th</sup> June 2022. One of them mentioned how the height of the GoBe robot was an advantage for them, because it would make residents stand up to interact with it. Hence, interactions with the robot required both a cognitive and a physical activity.

The problem came when some residents complain on the very same reason the clinics were praising the GoBe robot. They did not feel like getting up to interact with the robot (as Fig. 12 shows, for question "Which screen is easier to tap?"). Moreover, there is still the case of those residents which are incapable of standing up.

The importance of this dilemma is directly tied to the design of the socially assistive robot. In this study, one of the main goals that were set was to make the robot adapt to

the user, regardless of their sensory and motor capabilities. However, the present dilemma questions this goal. Hence, it is important to regulate the facilitation lent by the robot to avoid a lack of proactive interaction in the user's side. The dilemma can also be stated as follows: when designing a robot following an user centered design, how should designers proceed if two different user groups have opposite requirements for the robot? The answer to this question is not obvious. It implies exhibiting different behaviours for different situations and users. Hence, the robot should not follow a generic use case, not even when it has been designed following user needs. Instead, individualized responses and behaviours become the mechanism that allows the robot modulate its responses to each particular user and context [56]. This behavioural adaptation may let the robot avoid, at least partially, issues emanating from these dilemmas.

## 8.3 Main Difficulties Met

- The COVID-19 constant hazard both hindered the communication with the residents and significantly affected the number of possible sessions for the study.
- As previously detailed, our target users were usually people able to live independently. Hence, many of them were staying at the retirement home only for a certain time period (e.g. during a rehabilitation process). This circumstance produced losses in the volunteer groups. Next studies will address this issue by determining as exclusion criteria "People who are expected to be in the retirement home for less than 2 months".
- Employed questionnaires are too long. After each session, most participants felt exhausted. After they had to perform the first post test interview, they were no longer interested in doing that again.
- Some of the questions were difficult to answer for some participants (e.g. "Are you familiar with the concept 'smart environment'?").
- Safety distance detectors, that avoided the robot to approach too much to residents, had to be installed in the robots. These same systems prevented the robot to approach the users to a comfortable interaction distance. Hence, the residents had to approach the robot when it stops close to them. This circumstance limits the proactive HRI capabilities of the robots.
- During the video call use case study with relatives, many difficulties were found regarding both the recruiting and the scheduling process (agenda mismatches, retirement home COVID-19 protocol, etc). This issue limited the number of video calls that could be performed during the evaluation period.
- SAR usually raise high expectations. Seeing the robot moving and interacting, even in the simple scenarios employed in this study, made it easy for users to expect more capabilities. Some false expectations inevitably arose between users, specially in the early stages of the experiments. For example, as

soon as the SAR was stationed in the retirement home, many residents started talking to it, expecting a response. Others wanted it to cook for them, or expected the robot to recognize them and remember previous conversations. These false expectations should be avoided as they are one of the main reasons why users disengage from using the robot [10, 24]. In the study presented in this paper, the functionalities of the SARs were defined beforehand counting with the end users in the design process [11]. These participatory strategies reduce expectations, and lead to more constructive feedback and higher participation rates.

- From a technical perspective, these experiments have highlighted some minor and major issues that affect robot performance and will be discussed in a different paper. They also help identifying physical and functional features for the robots to achieve long-term acceptance.

## 9 Conclusion

This paper describes user experience results from a study on the long-term use of SAR in retirement homes. The use cases were not too ambitious, but they were created in a co-creative process to meet the specific needs of the users in the target retirement home. These use cases sufficed to gain insight into the HRI process. The number of participants involved in the experiments was also limited. However, the paper provides some preliminary results on how different factors, related to appearance and functionality, affect the acceptability, utility and accessibility of the robots once the novelty effect fades out.

The presented study has shown that, although an 'endearing' or 'cute' appearance encourages interaction with a SAR, functionality is the quality that prevails over time in the evaluated use cases. On the other hand, as simple as the employed use cases are, a SAR working in a retirement home raises some high expectations it will not fulfill. In order to avoid these false expectations to produce disengagement, constant and fluent communication between researchers and users is necessary during the experiments.

One of the main highlights of this study is the direct relationship between previous experience with technology and (i) the ability to learn to use new devices and; (ii) most importantly, attitudes towards interacting with a robot. Nine of the 14 participants had previous experience with technology (either computers and/or mobile phones). This group not only experienced a better learning curve, but were also more motivated to use the robot and give feedback.

Finally, this study also concludes that the role of a SAR as a social facilitator can have a relevant impact in elderly care. The benefits and drawbacks of this particular effect should be further explored, specially when robots are equipped with

the ability to adapt their behaviour to different contexts and users [56].

**Acknowledgements** This work has been funded by the regional projects AT17-5509-UMA 'ROSI' (Plan Andaluz de Investigación, Desarrollo e Innovación—PAIDI 2020, Junta de Andalucía) and UMA18-FEDERJA-074 'ITERA' (Programa Operativo FEDER Andalucía 2014–2020. Convocatoria 2018), by the FEDER/Ministerio de Ciencia e Innovación—Agencia Estatal de Investigación RTI2018-099522-B-C41 and RTI2018-099522-A-C44 projects, and by the DIH-HERO SUSTAIN project (European Union's Horizon 2020 research and innovation programme, under grant agreement No 825003). The authors would like to thank the residents, relatives and professionals at Vitalia Teatinos, for their kind help during all these experiments. Funding for open access charge: Universidad de Málaga / CBUA.

**Funding** Funding for open access publishing: Universidad Málaga/CBUA

**Data Availability** The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

## Declarations

**Conflict of interest** People affiliated to the same university might know about the work (Universidad de Málaga, Universidad Carlos III de Madrid and Universidad de Jaén). People working at Blue Ocean Robotics or MetraLabs GmbH might also know about the work. The authors declare no further Conflict of interest.

**Ethical Approval** The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Provincial Ethical Committee of the Andalusian Public Healthcare system (Comité de Ética de la Investigación Provincial de Málaga, session hold on 24/09/2020).

**Informed Consent** Informed consent was obtained from all subjects involved in the study.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. World Population Prospects (2022) <https://population.un.org/wpp/>
2. Proyección de población de Andalucía por ámbitos subregionales 2009–2035 (2012)
3. Ribeiro O, Araújo L, Figueiredo D, Paúl C, Teixeira L (2022) The caregiver support ratio in Europe: estimating the future of potentially (un)available caregivers. *Healthcare* 10(1):11. <https://doi.org/10.3390/healthcare10010011>

4. Feil-Seifer D, Mataric MJ (2005) Defining socially assistive robotics. In: 2005 IEEE 9th international conference on rehabilitation robotics, pp 465–468
5. Gómez-Donoso F, Escalona F, Rivas FM, Cañas JM, Cazorla M (2019) Enhancing the ambient assisted living capabilities with a mobile robot. *Comput Intell Neurosci*. <https://doi.org/10.1109/HUMANOIDS.2017.8246957>
6. Anghel I, Cioara T, Moldovan D, Antal M, Pop CD, Salomie I, Pop CB, Chifu VR (2020) Smart environments and social robots for age-friendly integrated care services. *Int J Environ Res Public Health*. <https://doi.org/10.3390/ijerph17113801>
7. Robotics 2020 multi-annual roadmap for robotics in Europe (2015) Technical report, SPARC: The Partnership for Robotics in Europe, euRobotics Aisbl, Brussels, Belgium
8. Vandemeulebroucke T, Casterle BD, Gastmans C (2020) Ethics of socially assistive robots in aged-care settings: a socio-historical contextualisation. *J Med Ethics* 46:128–136
9. Keizer RACMO, van Velsen L, Moncharmont M, Riche B, Ammour N, Signore SD, Zia G, Hermens H, N'Dja A (2019) Using socially assistive robots for monitoring and preventing frailty among older adults: a study on usability and user experience challenges. *Health Technol* 9:595–605
10. Seibt J, Damholdt MF, Vestergaard C (2020) Integrative social robotics, value-driven design, and transdisciplinarity. *Interact Stud* 21(1):111–144
11. Iglesias A, Viciano R, Pérez-Lorenzo JM, Lan Hing Ting K, Tudela A, Marfil R, Dueñas A, Bandera JP (2020) Towards long term acceptance of socially assistive robots in retirement houses: use case definition. In: 2020 IEEE international conference on autonomous robot systems and competitions (ICARSC), pp 134–139
12. Lan Hing Ting K, Voilmy D, De Roll Q, Iglesias A, Marfil R (2021) Fieldwork and field trials in hospitals: co-designing a robotic solution to support data collection in geriatric assessment. *Appl Sci* 11(7):3046
13. Breazeal C (2001) Affective interaction between humans and robots. In: Kelemen J, Sosík P (eds) *Advances in artificial life*. Springer, Berlin, pp 582–591
14. Payr S (2015) Towards human–robot interaction ethics. In: Trapp R (ed) *A construction manual for robots' ethical systems*. Cognitive Technologies. Springer, Cham
15. Pulido JC, González JC, Suárez-Mejías C, Bandera A, Bustos P, Fernández F (2017) Evaluating the child-robot interaction of the natherapist platform in pediatric rehabilitation. *Int J Soc Robot* 9:343–358
16. Wainer J, Robins B, Amirabdollahian F, Dautenhahn K (2014) Using the humanoid robot Kaspar to autonomously play triadic games and facilitate collaborative play among children with autism. *Auton Ment Dev* 6:183–199
17. Ali S, Mehmood F, Ayaz Y, Sajid M, Sadia H, Nawaz R (2022) An experimental trial: multi-robot therapy for categorization of autism level using hidden Markov model. *J Educ Comput Res* 60:722–741. <https://doi.org/10.1177/07356331211040405>
18. Chang WL, Šabanovic S (2015) Interaction expands function: social shaping of the therapeutic robot Paro in a nursing home. In: *Proceedings of the tenth annual ACM/IEEE international conference on human–robot interaction, HRI '15*, Portland, OR, USA, pp 343–350
19. Kitt ER, Crossman MK, Matijczak A, Burns GB, Kazdin AE (2021) Evaluating the role of a socially assistive robot in children's mental health care. *J Child Fam Stud* 30(7):1722–1735. <https://doi.org/10.1007/s10826-021-01977-5>
20. Kriegel J, Grabner V, Tuttle-Weidinger L, Ehrenmüller I (2019) Socially assistive robots (SAR) in in-patient care for the elderly. *Stud Health Technol Inform* 260:178–185
21. Papadopoulos I, Koulouglioti C, Lazzarino R, Ali S (2020) Enablers and barriers to the implementation of socially assistive humanoid robots in health and social care: a systematic review. *BMJ Open*. <https://doi.org/10.1136/bmjopen-2019-033096>
22. Iglesias A, García J, García-Olaya A, Fuentetaja R, Fernández F, Romero-Garcés A, Marfil R, Bandera A, Ting KLH, Voilmy D, Dueñas A, Suárez-Mejías C (2021) Extending the evaluation of social assistive robots with accessibility indicators: the AUSUS evaluation framework. *IEEE Trans Hum Mach Syst* 51(6):601–612. <https://doi.org/10.1109/THMS.2021.3112976>
23. Arthanat S, Begum M, Gu T, LaRoche DP, Xu D, Zhang N (2020) Caregiver perspectives on a smart home-based socially assistive robot for individuals with Alzheimer's disease and related dementia. *Disabil Rehabil Assist Technol* 15(7):789–798. <https://doi.org/10.1080/17483107.2020.1753831>
24. Gallego J, Lohse M, Evers V (1987) D6.3: acceptability of a home companion robot. Technical Report Final Report v 1.3, Accompany Project
25. Kachouie R, Sedighdeli Khosla R, Chu M-T (2014) Socially assistive robots in elderly care: a mixed-method systematic literature review. *Int J Hum Comput Interact* 30(5):369–393. <https://doi.org/10.1080/10447318.2013.873278>
26. Hall A, Wilson CB, Stanmore E, Todd C (2017) Implementing monitoring technologies in care homes for people with dementia: a qualitative exploration using normalization process theory. *Int J Nurs Stud* 72:60–70. <https://doi.org/10.1016/j.ijnurstu.2017.04.008>
27. McGinn C, Bourke E, Murtagh A, Donovan C, Cullinan MF (2019) Meeting stevie: perceptions of a socially assistive robot by residents and staff in a long-term care facility. In: 2019 14th ACM/IEEE international conference on human–robot interaction (HRI), pp 602–603. <https://doi.org/10.1109/HRI.2019.8673161>
28. Clabaugh C, Mahajan K, Jain S, Pakkar R, Becerra D, Shi Z, Deng E, Lee R, Ragusa G, Mataric M (2019) Long-term personalization of an in-home socially assistive robot for children with autism spectrum disorders. *Front Robot AI*. <https://doi.org/10.3389/frobt.2019.00110>
29. Céspedes N, Irfan B, Senft E, Cifuentes CA, Gutiérrez LF, Rincon-Roncancio M, Belpaeme T, Múnica M (2021) A socially assistive robot for long-term cardiac rehabilitation in the real world. *Front Neurobot*. <https://doi.org/10.3389/fnbot.2021.633248>
30. Getson C, Nejat G (2022) The adoption of socially assistive robots for long-term care: during COVID-19 and in a post-pandemic society. *Front Neurobot* 35(5):301–309. <https://doi.org/10.1177/08404704221106406>
31. Abras C, Maloney-Krichmar D, Preece J (2004) User-centered design. In: Bainbridge W (ed) *Encyclopedia of human–computer interaction*, vol 37. Sage Publications, Thousand Oaks, pp 445–456
32. Van der Velden M, Mörtberg C et al (2015) Participatory design and design for values. *Handbook of ethics, values, and technological design: sources, theory, values and application domains*. Springer, Cham, pp 41–66
33. Simonsen J, Robertson T (eds) (2012) *Routledge international handbook of participatory design*. Routledge, New York
34. Bratteteig T, Wagner I (2014) *Disentangling participation: power and decision-making in participatory design*. Springer, Berlin
35. Nygaard K (1979) The iron and metal project trade union participation. In: Sandberg A (ed) *Computers dividing man and work*, vol 12. Utbildningsproduktion, Malmö, pp 94–107
36. Weiss A, Bernhaupt R, Lankes M, Tscheligi M (2009) The USUS evaluation framework for human–robot interaction. In: *AISB2009: proceedings of the symposium on new frontiers in human–robot interaction*, vol 4, pp 11–26
37. Joshi SG, Bratteteig T (2016) Designing for prolonged mastery on involving old people in participatory design. *Scand J Inf Syst* 28(1):3–36

38. Voilmy D, Suarez C, Romero-Garcés A, Reuther C, Pulido JC, Marfil R, Manso LJ, Lan Hing Ting K, Iglesias A, González JC, García J, García Olaya A, Fuentetaja R, Fernández F, Dueñas A, Calderita LV, Bustos P, Barile T, Bandera JP, Bandera A (2017) CLARC: a cognitive robot for helping geriatric doctors in real scenarios. In: *ROBOT. Advances in intelligent systems and computing*, vol 693. Springer, pp 403–414
39. Bandera A, Bandera JP, Bustos P, Calderita L, Dueñas A, Fernández F, Fuentetaja R, Olaya A, García-Polo F, González Dorado JC, Iglesias A, Manso L, Marfil R, Pulido JC, Reuthe C, Romero-Garcés A, Suárez C (2016) CLARC: a robotic architecture for comprehensive geriatric assessment. In: *Proceedings of the XVII workshop of physical agents*
40. Stanford Artificial Intelligence Laboratory et al. Robotic Operating System. <https://www.ros.org>
41. Bustos P, Manso LJ, Bandera AJ, Bandera JP, García-Varea I, Martínez-Gómez J (2019) The cortex cognitive robotics architecture: use cases. *Cogn Syst Res* 55:107–123
42. Lobo A, Saz P, Marcos G, Dña JL, De La Cámara C, Ventura T, Asín FA, Pascual LAM, Montañés JMR, Aznar SN (1999) Revalidation and standardization of the cognition mini-exam (first Spanish version of the Mini-Mental Status Examination) in the general geriatric population. *Med Clin* 112(20):767–74
43. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H (2005) The Montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* 53:695–699. <https://doi.org/10.1111/j.1532-5415.2005.53221>
44. Joshi SG, Bratteteig T (2015) Assembling fragments into continuous design: on participatory design with old people. In: Oinas-Kukkonen H, Iivari N, Kuutti K, Öörni A, Rajanen M (eds) *Nordic contributions in IS research*. Springer, Cham, pp 13–29
45. Qbilat M, Iglesias A, Belpaeme T (2021) A proposal of accessibility guidelines for human–robot interaction. *Electronics* 10(5):561
46. Fernandes Í, Rocha S, Portela C, Braz Junior G, Almeida J, Silva A, Viana D, Rabelo J, Paiva A, Rivero L (2022) Defining an a/b testing process for usability and user experience evaluation through the analysis of the results of a literature review. In: *HCI international 2022-late breaking papers. Design, user experience and interaction: 24th international conference on human–computer interaction, HCII 2022, virtual event, 26 June–1 July, 2022*, Proceedings. Springer, pp 204–213
47. Joshi A, Kale S, Chandel S, Pal DK (2015) Likert scale: explored and explained. *Br J Appl Sci Technol* 7(4):396
48. Saplacan Lindblom D, Herstad J, Zada P (2020) An analysis of independent living elderly's views on robots a descriptive study from the Norwegian context. In: *Proceedings of the thirteenth international conference on advances in computer–human interactions (ACHI 2020)*, pp 199–208
49. Barnard Y, Bradley MD, Hodgson F, Lloyd AD (2013) Learning to use new technologies by older adults: perceived difficulties, experimentation behaviour and usability. *Comput Hum Behav* 29(4):1715–1724. <https://doi.org/10.1016/j.chb.2013.02.006>
50. Yogeewaran K, Zlotowski J, Livingstone M, Bartneck C, Sumioka H, Ishiguro H (2016) The interactive effects of robot anthropomorphism and robot ability on perceived threat and support for robotics research. *J Hum Robot Interact* 5(2):29–47. <https://doi.org/10.5898/JHRI.5.2.Yogeewaran>
51. Hinds PJ, Roberts TL, Jones H (2004) Whose job is it anyway? A study of human–robot interaction in a collaborative task. *Hum Comput Interact* 19(1–2):151–181. <https://doi.org/10.1080/07370024.2004.9667343>
52. Martinetti A, Chemweno PK, Nizamis K, Fosch-Villaronga E (2021) Redefining safety in light of human–robot interaction: a critical review of current standards and regulations. *Front Chem Eng* 3:666237. <https://doi.org/10.3389/fceng.2021.666237>
53. Abdi J et al (2018) Scoping review on the use of socially assistive robot technology in elderly care. *BMJ Open* 8:e018815
54. Anghel I, Cioara T, Moldovan D, Antal M, Pop CD, Salomie I, Pop CB, Chifu VR (2020) Smart environments and social robots for age-friendly integrated care services. *Int J Environ Res Public Health* 17(11):3801. <https://doi.org/10.3390/ijerph17113801>
55. Rojas Sierra X, Nørskov S (2023) Mobile telepresence robots and meaningful work in health care. In: *39th European Group for Organizational Studies (EGOS) Colloquium; Conference date: 06-07-2023 through 08-07-2023*
56. Bensch S, Sun J, Bandera JP, Romero-Garcés A, Hellström T (2023) Personalised multi-modal communication for HRI. In: *Proceedings of the RO-MAN conference, weighting the benefits of autonomous robot personalization (WARN) workshop*

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

**Antonio Jerez** finished his degree in Healthcare Engineering in 2021 and since then holds a position as Teaching and Research Personnel. He is a researcher in the TIC-125 (Ingeniería de Sistemas Integrados) research group. His work, and also the main theme of the thesis he is currently working on, focuses on socially assistive robots, design research and UX evaluation. He is involved in research projects that aim to implement socially-aware robots, integrated in both privated and shared smart environments, always following both participatory and user centered methodologies in order to design accepted, useful and accessible tools.

**Ana Iglesias** obtained her Ph.D. in 2004 and has been an assistant professor at the Carlos III University of Madrid since 2019. She is a researcher in the Department of Computer Science and Engineering. Her research interests focus on accessibility, human-machine interaction, advanced database technologies, natural language processing and information retrieval, among others. She has received several national and international research awards for her work in the field of accessible electronic devices for all.

**José Manuel Pérez-Lorenzo** obtained his Ph.D. in 2010, and he is a lecturer at the University of Jaén, where he is a member of the TIC-257 (Advances in Intelligent Systems and Applications) research group. His work is related to fields such as artificial perception algorithms, networks and internet technologies and cloud services. He has been involved in research projects that try to apply different technological solutions to the field of health and personal care.

**Alberto Tudela** Telecommunication Engineer with specialization in electronics by the University of Malaga in 2016. In September of the same year he joined the Integrated Systems Engineering group of the Department of Electronic Technology of the University of Malaga in a project of integration of manipulator arms by 3D printing within the framework of the CLARC system architecture. In June 2017, he obtained the Master's degree in Mechatronics Engineering from the University of Malaga with a TFM focused on pedestrian detection for autonomous vehicles. Subsequently, he has continued in the same line of research developing perceptual agents that allow monitoring the state of the world to be integrated into the cognitive architecture of a social assistive robot. Since then, he has participated in 5 R+D+I

projects, made 3 contributions to congresses, contributed with 2 chapters in specialized books and published 2 articles.

**Alejandro Cruces** holds a higher engineering degree in telecommunications and a master's degree about teaching. He has worked for more than ten years in the research group TIC-125 (Integrated Systems Engineering) at the University of Málaga, where he has also held a full-time teaching position for two years in the area of Electronic Technology. He is currently in the process of obtaining a PhD in the field of social robotics, focused on improving the adaptation of these robots to their context thanks to their integration in intelligent environments, promoting the synergy between both technologies.

**Juan Pedro Bandera** obtained his Ph.D. in 2010 and since 2017 holds a position as Assistant Professor at the University of Málaga. He is a researcher in the TIC-125 (Ingeniería de Sistemas Integrados) research group. His work focuses on robotics, artificial vision, smart environments and socially assistive robots. He is involved in research projects that aim to implement socially-aware robots, integrated in smart environments and able to adapt to different users and contexts. These projects also aim to move the robots from the lab to daily life environments, following user centered design processes. These robots aim to become accepted, useful and accessible tools.