

Results of Field Trials with a Mobile Service Robot for Older Adults in 16 Private Households

MARKUS BAJONES, DAVID FISCHINGER, and ASTRID WEISS, TU Wien, Austria

PALOMA DE LA PUENTE, Universidad Politécnica de Madrid, Spain

DANIEL WOLF and MARKUS VINCZE, TU Wien, Austria

TOBIAS KÖRTNER and MARKUS WENINGER, Academy for Aging Research, Austria

KONSTANTINOS PAPOUTSAKIS, DAMIEN MICHEL, AMMAR QAMMAZ, PASCHALIS

PANTELERIS, MICHALIS FOUKARAKIS, ILIA ADAMI, DANAE IOANNIDI, ASTERIOS

LEONIDIS, MARGHERITA ANTONA, and ANTONIS ARGYROS, FORTH, Greece

PETER MAYER and PAUL PANEK, TU Wien, Austria

HÅKAN EFTRING, Lund University, Sweden

SUSANNE FRENNERT, Malmö University, Sweden

In this article, we present results obtained from field trials with the Hobbit robotic platform, an assistive, social service robot aiming at enabling prolonged independent living of older adults in their own homes. Our main contribution lies within the detailed results on perceived safety, usability, and acceptance from field trials with autonomous robots in real homes of older users. In these field trials, we studied how 16 older adults (75 plus) lived with autonomously interacting service robots over multiple weeks.

Robots have been employed for periods of months previously in home environments for older people, and some have been tested with manipulation abilities, but this is the first time a study has tested a robot in private homes that provided the combination of manipulation abilities, autonomous navigation, and non-scheduled interaction for an extended period of time. This article aims to explore how older adults interact with such a robot in their private homes. Our results show that all users interacted with Hobbit daily, rated most functions as well working, and reported that they believe that Hobbit will be part of future elderly care. We show that Hobbit's adaptive behavior approach towards the user increasingly eased the interaction between the users and the robot. Our trials reveal the necessity to move into actual users' homes, as only there, we encounter real-world challenges and demonstrate issues such as misinterpretation of actions during non-scripted human-robot interaction.

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 288146 (HOBbit).

Authors' addresses: M. Bajones, D. Fischinger, A. Weiss, D. Wolf, M. Vincze, P. Mayer, and P. Panek, TU Wien, Vienna, Austria; emails: bajones@acin.tuwien.ac.at, david.fischinger@gmail.com, astrid.weiss@tuwien.ac.at, dinsn87@gmail.com, markus.vincze@tuwien.ac.at, {mayer, panek}@fortec.tuwien.ac.at; P. de la Puente, Universidad Politécnica de Madrid, Madrid, Spain; email: paloma.delapuate@upm.es; T. Körtner and M. Weninger, Academy for Aging Research, Vienna, Austria; emails: Tobias.Koertner@gmx.at, herr.weninger@gmail.com; K. Papoutsakis, D. Michel, A. Qammaz, P. Panteleris, M. Foukarakis, I. Adami, D. Ioannidi, A. Leonidis, M. Antona, and A. Argyros, FORTH, Heraklion, Greece; emails: {papoutsakis, michel, ammarkov, padeler, foukas, iadami, ioanidi, leonidis, antona, argyros}@ics.forth.gr; H. Efring, Lund University, Lund, Sweden; email: hakan.eftring@certec.lth.se; S. Frennert, Malmö University, Malmö, Sweden; email: susanne.frennert@mau.se.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

2573-9522/2019/12-ART10

<https://doi.org/10.1145/3368554>

CCS Concepts: • **Human-centered computing** → **User studies; Field studies**; • **Computer systems organization** → **Robotic autonomy**; • **Computing methodologies** → Vision for robotics;

Additional Key Words and Phrases: Social robot, assistive robotics, long-term human-robot interaction, mutual care, field trials, robotic companions

ACM Reference format:

Markus Bajones, David Fischinger, Astrid Weiss, Paloma de la Puente, Daniel Wolf, Markus Vincze, Tobias Körtner, Markus Weninger, Konstantinos Papoutsakis, Damien Michel, Ammar Qammaz, Paschalis Panteleris, Michalis Foukarakis, Ilia Adami, Danae Ioannidi, Asterios Leonidis, Margherita Antona, Antonis Argyros, Peter Mayer, Paul Panek, Håkan Efrting, and Susanne Frennert. 2019. Results of Field Trials with a Mobile Service Robot for Older Adults in 16 Private Households. *Trans. Hum.-Robot Interact.* 9, 2, Article 10 (December 2019), 27 pages.

<https://doi.org/10.1145/3368554>

1 INTRODUCTION

The problem of an aging society sparks the need for technological advances that could enable social service robots to advance in elderly care tasks. Different companies and research teams provide robots for the goal of improved health care such as, e.g., Giraff,¹ Care-O-Bot,² and Kompai.³ Studies on needs and expectations of older adults towards socially assistive robots [4] indicate that people expect robots to help with household chores (cleaning the kitchen, bath, and toilet), lifting heavy objects, and reaching for, picking up, and delivering objects. While some of these tasks can be performed by state-of-the-art robotic platforms, only a few have so far entered private homes of older adults. One of the biggest challenges is offering sufficient useful and social functionalities autonomously and safely to achieve the ultimate goal of prolonging independent living at home.

In this article, we present field trials with the Hobbit PT2 robot, in the remainder of this article referred to as *Hobbit*. A former version of the robot, Hobbit PT1, was presented in detail in Reference [15]. Hobbit is an assistive robot that offers useful personal and social functionalities to enable independent living at home for seniors. Previous studies have been conducted in users' homes with robots that stayed active for a longer period or provided manipulation tasks (e.g., picking up objects). However, these studies often used a robot with either limited autonomy or limited manipulation capabilities or were only available during pre-scheduled sessions. The novelty we report for the Hobbit trials is the fact that for the first time a social service robot with extended manipulation abilities autonomously operated in the homes of end-users offering multiple functionalities, during which it was the users' choice to decide how and when to interact with the robot.

The main contribution of this article is the results from field trials with 16 participants (three weeks for each user) during which they were not limited by a fixed schedule or set of tasks they had to perform with the robot. Participants were allowed to freely decide how and when to interact with Hobbit.

The reason for the chosen approach was that we wanted to explore actual usage patterns of the robot as a whole instead of studying the usability and acceptance of single functionalities. Up to now, we are lacking data on how users interact with service robots over a longer period in

¹<http://www.giraff.org>.

²<http://www.care-o-bot.de>.

³<http://www.robosoft.com/robotic-solutions/healthcare/kompai/kompai-rd.html>.

non-controlled environments. We completed all field trials in private homes in Austria, Greece, and Sweden.

The article proceeds as follows: Section 2 discusses relevant related work on studies of similar robots outside of laboratory environments. In Section 3, we give a short overview of the Hobbit robot. Section 4 presents the field trials, insights we gained from the autonomous interaction between our robot and the older adults in different private households, and discussion. Summary and conclusions follow in Section 5.

2 RELATED WORK

2.1 Care Robots

As early as 1996, Dario et al. [7] designed and developed mobile robots to assist the elderly at home (MOVAID), patients in a hospital (URMAD), or people in a wheelchair (IMMEDIATE). The URMAD goals were to assist humans with searching objects in the environment, picking them up, and carrying them to the user. MOVAID, however, was supposed to warm up food in a microwave oven and clean kitchen surfaces as well as remove dirty sheets from a bed within a user's home. Prototypes of the MOVAID robot have been tested and evaluated in a residential site for disabled people, to inform a survey on the acceptability and perceived usability of the system. The design of the IMMEDIATE robot is of a wheelchair equipped with an arm and a manipulator to pick up objects for the user.

Pearl,⁴ developed by Pollack et al. [27] in the Nursebot project, is a mobile robot developed to fulfill two main tasks for older adults living in a retirement center: (a) to give reminders about drinking, eating, and taking medicine, and (b) to guide a person to their appointments.

Charles, a robot based on the Peoplebot⁵ platform, was designed by Kuo et al. [22] to serve users who need vital signs to be regularly monitored. Their main focus was on the inclusion of a multitude of medical equipment to collect a more comprehensive range of vital signs and to explore the ability to perceive and respond to users using multiple modalities. When the robot was supposed to carry an object, however, it was limited by the 2-DoF gripper. Thus, it relied on the user to place the object between the gripper's fingers before closing them.

Multiple robots were developed in the Healthbots project [19] during recent years. The hardware itself is based on a mobile platform by Yujin Robot Company (Korea) and extended to accommodate sensors to measure blood pressure, oxygen saturation, and glucose levels as well as cameras and microphones.⁶ In New Zealand, they conducted three different field trials with one, two, or four robots at the same time. During the trials, the robots provided measurements of blood properties with the included sensors, simple entertainment, and brain training functionality. Further, it featured the capability to detect falls with the help of external sensors (connected via a ZigBee network).

Care-O-Bot [17] was developed for a safe Human-Robot Interaction, fetch-and-carry tasks, and patrolling in elderly care facilities. The robot can grasp objects using its 7-DoF arm (equipped with a 7-DoF dexterous hand) from the floor, the tray on the robot's front, and the area behind the robot; thus, enabling it to assist in cleaning up the environment and serve water, both extensions of the well-known fetch-and-carry tasks. Disadvantages of the Care-O-Bot are, however, the size (making it difficult to navigate in narrow environments, such as private homes) and the high cost.

⁴<https://www.cmu.edu/cmtoday/issues/dec-2004-issue/feature-stories/human-health/index.html>.

⁵<http://www.mobilerobots.com/Libraries/Downloads/PeopleBot-PPLB-RevA.sflb.ashx>.

⁶<http://robotics.auckland.ac.nz/current-work/healthcare-assistive-technologies/healthbots/>.

Robots with a focus on providing companionship and close physical contact include Paro [35],⁷ AIBO⁸ and Huggable [38].⁹ Such robots often have little or no navigation and manipulation skills, thus being distinctively different from the former mentioned robots that provide at least some kind of functionality to move and manipulate objects in the environment.

2.2 Field Trials in HRI

Multiple projects attempted to develop and deploy assistive, service robots into the homes of users in recent years. Wada et al. [41–43] focused their work on psychological, physiological, and social effects that Paro has on elderly in supported care facilities during long-term studies.

In 2006, Forlizzi and Disalvo [16] used Roomba¹⁰ in 14 households. Even though their study used a rather simple robot, offering only little functionality, the results gave insights on how a service robot's introduction into a household changes the view upon such technology. They further unveiled how the introduction of a more capable robot needs to consider the social norms of the target environment.

Sung et al. [39] explored the effect of possible customization and consequently, passive adaptation to the users in a study within 30 households for six months each. They also used Roomba robots and found that the possibility to adapt the robot's appearance in such a way that it blends into the environment can enhance the experience with a robot.

The *Companionable* project [33] prepared two domestic-like environments, outfitted with Ambient Assisted Living (AAL) technology, and invited six participants (accompanied by five secondary users) to spend two days in one of these homes. Their robot (Scitos G3 platform¹¹) navigated autonomously within a well-prepared environment but was limited to non-restricted guidance-free use for six hours per day. The study itself focused on participants with mild cognitive impairments and showed that the functionality to provide reminders and the ability to locate a person was especially well received by the five primary users.

In several user studies a baseline approach was chosen, defining the main functionality the robot should provide, as well as how it should be delivered in a real user home. The *ALIAS* robot [29], for example, was set up to mainly explore multi-modal input possibilities for Human-Robot interaction, including Automatic Speech Recognition (ASR) and a Brain-Computer Interface (BCI). During field trials, lasting two days with five participants, an emergency call, a game, and an e-ticket event booking were tested, and the navigational features including human-aware navigation adaptation were demonstrated.

In what can be considered as one of the first long-term user trials with a mobile service robot in users' homes, the *DOMEO* project [12] placed a robot in the homes of four elderly users for 287 days. Even though they used a mobile robot, it rarely made use of its ability to move within the home, as indicated by their results.

The *Serroga* project [18], direct successor of *Companionable*, moved from the constraints of their AAL-equipped living labs into users' homes. They successfully tested their robot with nine users for a total duration of 16 days in which the elderly users could use the robot for multiple hours without any guidance or intervention.

The recently finished *MARIO* project [13] focused on people with dementia living in care facilities. The robot provided a variety of applications on its touch interface and recorded the patient's

⁷<http://www.parorobots.com/>.

⁸<https://aibo.sony.jp/en/>.

⁹<http://robotic.media.mit.edu/portfolio/huggable/>.

¹⁰<https://www.irobot.com/For-the-Home/Vacuuming/Roomba.aspx>.

¹¹<https://www.metralabs.com/en/>.

motion behavior for later analysis. A focus of the project was to assist caregivers in comprehensive geriatric assessment during trials in care facilities in three countries.

Broadbent et al. [6] investigated the effect of six stationary robots in multiple aged care facilities. The robots (Guide¹² and Cafero¹³) were deployed for three months in total during which the robots were switched on between 6 AM and 8 PM. The robots provided entertainment, health monitoring, and communication functionality to users in the intervention group, while the control group (in different care homes) did not have any contact with the robots. Due to the limited number of interactions (that needed to be sought after and started by the users) the amount of collected data did not seem sufficient to show a significant effect between the two groups.

In a more recent study Broadbent [5] deployed IRobi¹⁴ in the homes of 30 participants for four months. The robots' main tasks were the measurement of a set of vital signs and to present Clinical COPD Questionnaire (CCQ),¹⁵ provide reminders, and an *I am feeling unwell* function. Their findings suggest that such a robot may be useful for COPD¹⁶ patients who struggle with keeping their exercise and medication schedule. The authors note, however, that *A number of technical issues would need to be improved before the robot could be implemented on a larger scale.*

In different studies, Kanda and Shiomi et al. placed multiple robots in a science museum [36] and in a shopping mall [20]. They used different robots¹⁷ for guiding visitors through an exhibition, interaction (playing rock, paper, and scissors, hugging, or handshaking), and bidding farewell. Out of about 10K visitors who interacted with the robots, a third returned questionnaires. Their evaluation suggests that the provided guidance increased the interest in science and technology. In the shopping mall a stationary Robovie, using pressure plates to detect visitors and partly controlled by a remote operator for speech recognition and behavior selection/decision making, was used to provide information to visitors. Questionnaires from 235 participating visitors indicated increased usefulness of given information, interest in shops, and visiting and shopping frequency when interacting with the robot compared to an information display.

The discussed work shows a common trade-off between the duration of the study and the complexity of the robot. When the complexity and the number of available functionalities of the robot increased, trade-offs had to be made in the study design, such as (1) reducing the interaction's duration to a few hours (sometimes split over multiple days), (2) carrying out the studies in custom set-up living laboratories, (3) providing a rather small set of functionalities on the robot, (4) reducing the autonomy of the robot, or (5) a combination.

In Table 1, we show how the Hobbit field trials compare against relevant related state-of-the-art studies. As can be seen, prior work excelled in some of the shown categories; Hobbit's novelty lies in achieving a combination of successfully testing extensive manipulation capabilities and the long duration of non-scheduled interactions and autonomous navigation within the users' homes.

3 THE HOBBIT ROBOT

The motivation for Hobbit's development was to create a low-cost, social robot to enable older adults to independently live longer in their own homes. The functionality (described in more detail in Section 3.1) provided by Hobbit has been based on the given requirements and laboratory studies with the first generation of the Hobbit platform [15] conducted with 49 users in Austria, Greece, and Sweden. One reason for older adults to move into care facilities is the risk of falling

¹²<http://www.ed.co.kr/eng/>.

¹³<http://yujinrobot.com/eng/>.

¹⁴<http://en.yujinrobot.com/archives/portfolio-items/irobi-q>.

¹⁵<http://ccq.nl/>.

¹⁶COPD: Chronic Obstructive Pulmonary Disease.

¹⁷https://www.vstone.co.jp/english/products/robovie_x/.

Table 1. Comparison of Related User and Field Trials (Including Hobbit)

	Robotic platform	scheduled vs. non-scheduled interactions	Provided navigation capabilities	Provided extensive manipulation capabilities	Duration	Environment	Provided functionalities
Wada [41, 42, 43]	Paro	non-scheduled	no	no	six weeks, one year	care facilities	companionship
Forlizzi and Disalvo [16]	Roomba	non-scheduled	yes	no	three to six weeks	households	floor cleaning
Sung [39]	Roomba	non-scheduled	yes	no	six months	households	floor cleaning
Schroeter [33]	Scitos G3	scheduled	yes	no	6 hours per day for 2 days	lab	
Rehrl [29]	Scitos G5	scheduled	yes	no	2 days	-	Emergency call, games, online ticket purchase
Fazekas [12]	Kompai	non-scheduled	yes	no	3 months	households	Emergency call, games, calendar, weather forecast, monitoring of blood pressure and weight
Gross [18]	Scitos G3	scheduled	yes	no	8 hours 2-3 hours	lab households	vital sign measurements, reminders, physical activity monitor, calendar
Felzmann [13]	Kompai	scheduled	yes	no	6 hours per day for 2 days	lab	Smart-home integration, cognitive training, calendar
Broadbent [6]	Guide, Cafero	non-scheduled	no	no	14 hours per day for 3 months	care facilities	Fall detection, blood pressure and oxygen level measurements, reminders, games
Smarr [37]	PR2	-	yes	yes	video based study	-	-
Zsiga [46]	Kompai	non-scheduled	yes	no	90 days	households	calendar communication games
Shiomi [36]	Robovie	non-scheduled	yes	no	2 months	science museum	reading RFID tags, physically guiding people to exhibitions
Kanda [20]	Robovie	non-scheduled	no	no	25 days	shopping mall	advertisement, verbally guiding visitors
Broadbent [5]	iRobi	non-scheduled	no	no	4 months	households	vital sign measurements, reminders
Bajones	Hobbit	non-scheduled	yes	yes	3 weeks	households	see Section 3.2

and eventually inflicting injuries. To reduce this risk, the “must-haves” for the Hobbit robot were *emergency detection* (detecting a previously fallen person or recognizing the fall of a user if it happens in the field-of-view of the robot’s sensors), *emergency handling* (automatic calls to relatives or an ambulance service after an emergency was detected), as well as *fall prevention* (searching and bringing known objects to the user and picking up objects from the floor pointed to by the user). Hobbit also provided a *safety check* feature that guided the user through possible risks in specific



Fig. 1. Rendering and hardware description of the second prototype of Hobbit (PT2).

rooms (e.g., wet floor in the bathroom, slippery carpets on wooden floors) and explained how to reduce those risks. Given the requirements and the laboratory studies, the wish from potential end-users and care personnel for Hobbit to provide a *basic fitness program*, designed to enhance the user's overall fitness, emerged and was integrated in the second prototype.

The *Mutual Care* concept [23], envisioned as reciprocal help between user and robot, served as fundamental interaction paradigm for the design of the adaptive robot behavior.

3.1 Robot Platform and Sensor Set-up

The Hobbit robot, as seen in Figure 1, is a custom-built platform with a rectangular footprint with a width of 0.48m, a length of 0.55m, and a height of 1.25m. On its right side, the robot is equipped with a six-DoF arm with a two-finger gripper, mounted in such a way that objects lying on the floor can be picked up and placed on a tray on top of the robot's body. Furthermore, the arm can grasp a small turntable stored on the right side of the body, which is used by the robot to learn unknown objects.

On the front of the robot is a tablet computer that provides the Multi-Modal User Interface (MMUI) as shown in Figure 2. In general, the MMUI is a framework containing the following main building blocks: Graphical User Interface (GUI) with touch, Automatic Speech Recognition (ASR), Text to Speech (TTS), and Gesture Recognition Interface (GRI). The MMUI provides the emergency

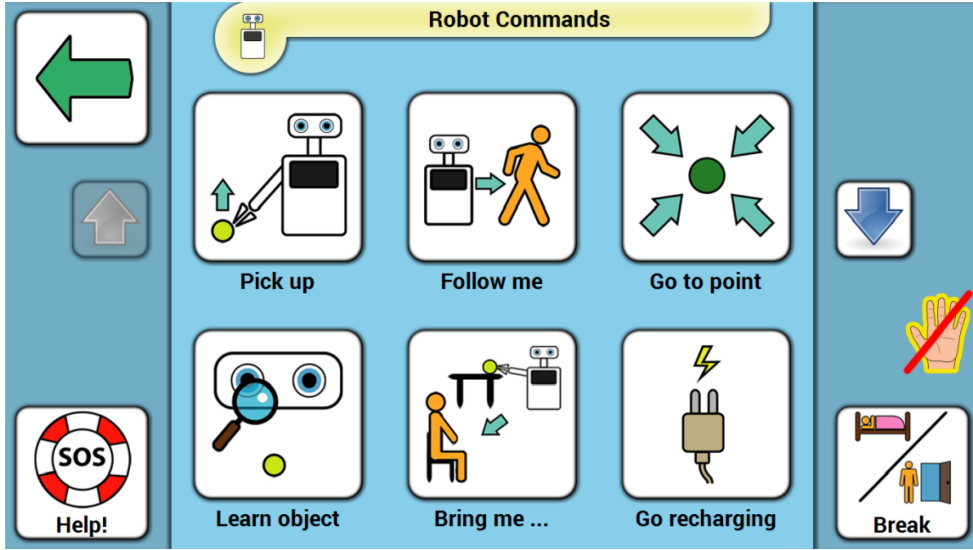


Fig. 2. *Robot commands* interface of the Multi-Modal User Interface of Hobbit PT2. The six most often wished for functions are shown in the center of the screen. On the right side an indicator shows that gesture recognition is currently unavailable. On the bottom left the software version of the emergency button (available in every menu of the user interface) is shown; the break button (used to inform the robot that the user is leaving the apartment or going to bed) is shown on the bottom right.¹⁸

call feature, web services (e.g., weather, news, shared calendar, and social media), control of robotic functions, and entertainment features. Figure 3 shows two use-cases of the MMUI. More details about the technical implementation of Hobbit are presented in References [2, 40].

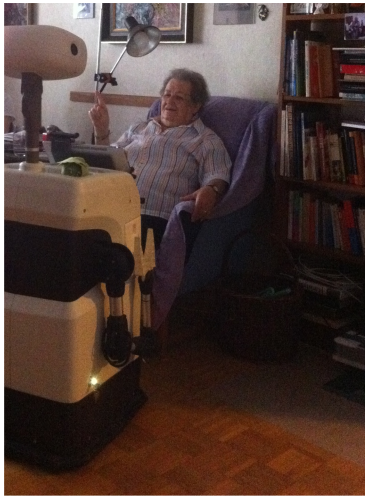
3.2 Robot Functionality

In the Hobbit project, we extensively investigated the users' needs, desires, and expectations towards the robot [21]. As typical in user-centered design approaches, not all user expectations can be fulfilled in the system implementation. Therefore, an informed decision was made by the interdisciplinary project consortium about which user requirements derived from the results obtained from the laboratory trials with the first robotic platform [14] can be realized in the next version.

To implement the Human-Robot-Interaction aspects of Hobbit, we needed to translate the user requirements into concrete scenarios regarding how the user will interact with the robot. We started based on ranked user requirements, which focused on safety aspects for the end-users, and decided which requirements were most important and also technically feasible in the context of the robot development.

Each interaction scenario Hobbit can perform was designed as a flow-chart with well-defined states in which the robot expects input from the user. However, input from outside of these states was not ignored, but the behavior coordination decided if a running task should be stopped beforehand and to start a new one or not. This decision making is based on the priorities of the active and to-be-started scenario (where *Emergency* has the highest priority). The trigger to start a scenario

¹⁸The symbols for *gesture or speech input not available* as well as all other parts of the user interface were designed in cooperation with our end-user partner (Austrian Academy for Aging Research) to ensure that they are easily understandable by our target group.



(a)



(b)

Fig. 3. Hobbit provides multi-modal input methods such as simple gestures, touch, and speech input. (a) shows the *No* gesture that can be used to answer Yes-or-No questions. In (b) a similar question is answered with *No* via the touch interface.

could come from different sources, such as speech, touchscreen, gesture recognition,¹⁹ scheduled event, or by pressing a physical button (e.g., emergency help button on the robot or remote button in the bathroom). All of the following behaviors were available within the field trials.

Call Hobbit was designed to initiate any interaction with the robot. Participants could press a wireless call button, associated with a specific location within their homes, to call the robot to this location. Another option was to use a speech command that prompted the robot to locate the user in the apartment. This was done by navigating from room to room, executing a full rotation in every room using the head camera to detect a person. When detected, the robot would move closer to the user, announce its arrival, and wait for further input.

Patrol was developed to make sure a fallen person would be detected after no more than three hours since the last interaction between Hobbit and the person occurred. The robot moved through the entire home, stopped in each room, and tried to detect the user in the room, focusing especially on the option of the user lying on the floor. If the person could not be found, Hobbit would call out to the user three times in five-minute intervals at full audio volume to capture the attention of the user if she or he is in a room not accessible to the robot (e.g., bathroom). When the user would not respond for five more minutes, Hobbit started the *Emergency* function.

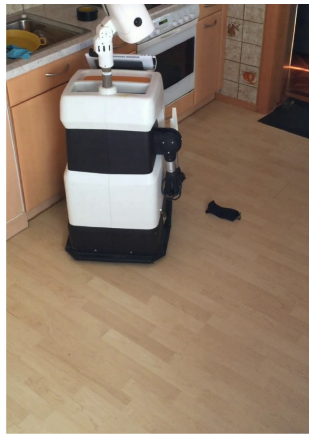
Emergency is the most important use case. As such, it can be started autonomously by the robot or through multiple modalities by the user:

- Speech commands, such as *Help me* or *Call help*.
- Input on the touchscreen mounted on the front of Hobbit.
- Pressing a physical button mounted on the front bottom of the robot.
- Pressing a wireless emergency button that is placed within the bathroom.

¹⁹The trigger from the gesture recognition was ignored under certain circumstances. For example, when the robot was looking at the rotating turntable during the *Users teaches Hobbit a new object* scenario.



(a) In the first step the user points at the object on the floor while standing in front of Hobbit



(b) The robot moves to a location from where it can better inspect the area on the floor



(c) After the object was detected and a grasping point is calculated, Hobbit grasps the object and places it on its tray.

Fig. 4. Hobbit performs the task of picking up a soft spectacle case during the trials for a user in Vienna.

But it can also be initiated in the following situations:

- The user falling in front of the robot.
- Hobbit patrolling through the apartment (three hours after the last interaction with the user) and is not able to detect her or him.

When the robot is not near the user, it first needs to move sufficiently close for speech recognition to work reliably. After proximity has been established the procedure tries to calm the person and call a relative or the emergency contact using the integrated Voice over IP solution (VoIP). The called person is then connected to the user so they can discuss the emergency and to reassure the user that help is on its way.

Safety check raises the awareness of potential risks that exist in a typical apartment. Examples include a wet floor in the bathroom, a rug that can slide, a stove unintentionally left on. The task for Hobbit is to move through the home alongside the user and provide safety tips on how to address such risks.

User goes away from home or to sleep is the method of letting the robot know that the user is not to be disturbed for a given amount of time (e.g., until next morning or for one hour). Hobbit presented the user with a short checklist of what to check before leaving the apartment or going to sleep. Items on this list included *Make sure all windows are closed*; *You should check if you turned off the stove*; *Do you want to be reminded of today's appointments?* This further disabled the *Patrol* functionality until the user returned home or until the set duration was over.

Pick up an object from the floor is directly related to the main goal of preventing falls by clearing the floor of unwanted objects to minimize the risk of a person stepping and slipping on them. When started, Hobbit simply asks a person to point to a desired object on the floor by using a fully stretched arm in the process (see Figure 4). It calculates the area where this gesture points to, moves to a location from where it can be inspected, and uses the head camera to detect any objects that stick out from the floor plane. Using the HAF algorithms [15] Hobbit calculates a possible grasp point and approach vector before moving to a position from which the object can be picked up. After it arrives at

the appropriate location, the arm moves to actually pick up the object from the floor and place it on the tray.

User teaches Hobbit a new object for later object-search and fetch-and-carry tasks. Hobbit provides the functionality to learn a 3D object model with the help of the user. The object is placed on the integrated turntable that is held in the gripper on the outstretched arm of the robot. While turning the object, color and depth information is constantly recorded using the head camera and combined into a full 3D model [28]. The user needs to name the object afterward so it can be referenced by the user in the *Bring object* task.

Bring object Hobbit presents a list of previously learned objects on the touchscreen for the user to choose from. Hobbit moves through the home and uses RGB-D-based object recognition [1] to detect the object and its pose. If possible, Hobbit would grasp the object and place it on its tray. If not, Hobbit guides the user to the location of the object.

Reminders were set up on a shared calender for primary (PU) and secondary users (SU). PUs as well as SUs, e.g., relatives, could add and edit appointments to be presented to the user on the robot's tablet at an appropriate time. Further daily hydration reminders were added to make sure our users would drink enough water during the day.

Transport object Hobbit follows the user to a desired location in the home, after an object was placed on its tray. To follow a person, we used RGB-D person tracking to retrieve the user's pose and calculate a path towards this location.

Fitness functionality as a means to increase the physical fitness of the PU was integrated after the laboratory studies with the first prototype. As a precondition, the user sits on a simple chair with sufficient space surrounding her/him to be able to move freely with the arms. Hobbit moves into a position about 1.5m away, pointing in the direction of the PU. On the integrated tablet, instructional videos were shown to the user on the left side of the screen and a skeletal representation of the tracked user's upper body on the right side. The user's movements are tracked and analyzed for correctness. If the exercise was accurately performed, then positive feedback was given. If not, then instructions on how to correct and improve were presented to the user.

Robot recharging fulfills the technical necessity of the robot to keep a certain battery level to guarantee autonomous operation and can be started by the user or by the robot itself. The charging procedure starts with Hobbit moving to the location of the docking station and detecting its unique shape for fine adjustment, while driving onto the charging pads. Metal contacts on the bottom of the robot touch these charging pads on the docking station for power transfer. For safety reasons the charging pads were not enabled until the robot identified itself to the station.

3.3 Mutual Care as Underlying Interaction Paradigm

The *Mutual Care* concept was implemented through two different social roles, one that enforces this concept and one that does not. Hobbit started in the non-mutual care (*device*) mode during the field trials and changed after 11 days to the Mutual Care (*companion*) mode. The differences between these two modes or *social roles* of the robot were mainly in its dialogues, pro-activity, and the proximity in which the robot would remain when the user stopped interacting with the robot. In more detail, the main characteristics of the *companion* mode were (1) Return of favor: Hobbit asked if it could return the favor after situations in which the user had helped Hobbit to carry out a task, (2) Communication style: Hobbit used the user's name in the dialogue and was more human-like, such as responding to a reward from the user by saying *You are welcome* instead of *Reward has been received*, (3) Pro-activity: Hobbit was more proactive and initiated interactions with the user, and (4) Presence: Hobbit stayed in the room where the last interaction took place for

at least 30 minutes instead of directly returning to the charging station. To avoid potential biases, users were not told about the behavioral change of the robot beforehand.

4 FIELD TRIALS

To evaluate Hobbit outside of controlled lab situations, field trials were performed in Austria, Greece, and Sweden. The trials ran in parallel with five robots and lasted approximately 21 days for each of the 18 households, resulting in 371 days of testing overall. Hobbit was placed in the private home of older adults, living by themselves, where the participants could interact with the robot on a 24/7 basis. The overall evaluation goal was to assess the Human-Robot-Interaction with Hobbit in terms of usability, user acceptance, and affordability (for space reasons, we only report on usability and user acceptance in this manuscript; see References [2, 44] for details). Therefore, an evaluation plan was needed, which we understood as the overall methodological concept for the household trials; it described the overall research questions, applied methods, and the timeline for conduction. Overall, the evaluation goal was to explore the following main question: Do older adults experience the Hobbit robot as a suitable means to maintain independent living in their private household? During the process of setting up the evaluation methodology for the household trials, we encountered several methodological challenges that we had to consider in our study design. We will briefly outline them before we go into detail of the trial procedure:

Exploring HRI with Older Adults. It was shown in Human-Computer Interaction research that older people tend to praise the developers rather than giving an objective view, thereby being very positive about prototypes they are presented and tending to blame themselves rather than the interaction modalities if not being able to cope with the system [11]. Age-related factors can also make self-reporting inaccurate (for example, in questionnaires), with recent research showing that there are differences due to age in the way in which people respond in self-reports, which we intended as an indication for a multi-factor/method mix for the long-term household evaluation [25]. A good example of these effects was observed in our first laboratory user study. In the final interviews of the study, some users considered themselves as too healthy and active to need a Hobbit at home. We partly consider this as an answer effect, as it would be stigmatizing for an older adult to admit that they need a Hobbit to independently live at home. Moreover, the prototype level will have impacted this reply, as it might be difficult for an older adult to imagine the full capacity of Hobbit at the next prototype stage. Subsequently, we decided to use a multi-factor/method mix and a multi-informant approach for the field trials combining self-reporting data with logging data to gather more objective data about the usage of Hobbit and to involve also secondary users (i.e., friends and relatives).

Field Studies. So-called evaluations “in the wild” are studies [30] where researchers are decamping from their labs and moving into actual users’ environments (e.g., homes), which is often claimed to be of utmost importance for HRI research [32]. A central part of evaluating outside of the lab involves observing and recording what people do and how this changes over suitable periods. Whereas the burning question in HRI studies used to be “How many participants do I need?” the hotly debated question now is “How long should my study run for?” [31]. Some researchers/papers say a few weeks, others argue for months or even suggest years are needed to show sustainable and long-term effects. However, the most crucial question is: “How can we measure what we need to know without having a researcher present?” and “How can we conclude causality from statistical correlations without being present?”

Exploring Various Factors over Time. It is well-known that the first impressions of robots are most often positive (so-called encounters at the zero acquaintance level) and to study these impressions is valuable for human-robot interaction scenarios [8] where the application context requires brief non-repeated interaction (e.g., a museum guide robot). However, this is not the case for an elderly-care robot at home. Preferences and attitudes of users are likely to change over time and novelty effects will wear off. In our studies over several weeks in private homes, the intention was to collect data on diverse quality factors, which is labor-, time-, and equipment-intensive, but required to obtain a better understanding of how a robotic product may become a social product.

Several long-term studies in HRI choose the approach to study only specific functionalities at specific phases of the study, e.g., first week one function, second week another function. (see, e.g., Reference [24]). This approach should guarantee that the participants do not experience a cognitive overload at the beginning of a long-term study. However, we intentionally decided against such an approach, as it is our overall goal to assess the holistic experience of Hobbit as a product. As soon as Hobbit becomes a real product, it will also offer all its functionalities at once and there will only be an initialization phase in which the user becomes acquainted with the robot. Therefore, we need the field trials as an ecological valid reference to identify: Which functionalities are used by the participants over a longer period and how does the interaction change after the novelty effect is over? A reasonable benchmark can, to our conviction, only be achieved if the robot offers the same interaction spectrum over three weeks.

4.1 Evaluating Usability and User Acceptance

Usability is defined as the overall ease of use; meaning how easily the user can interact with the robot. For the Hobbit field trial, the evaluation of the following usability indicators was considered as relevant: ease of learning, flexibility, and utility. The following research questions were investigated:

Ease of learning. Is Hobbit intuitive to interact with for the older adults?

Flexibility. Does Hobbit provide the relevant input modalities that enable the user to command it effectively?

Utility. Does Hobbit offer the right functionalities, therefore make older adults feel supported to maintain independent living at home?

User acceptance is in general defined as “the demonstrable willingness within a user group to employ technology for the task it is designed to support” [10]. To investigate this multi-layered aspect of Hobbit, a variety of different indicators needed to be assessed, namely: attitude towards robots, perceived safety, self-efficacy, emotional attachment, and perceived reciprocity. The following research questions were investigated:

Attitude towards robots. Does the general attitude towards robots change over time due to the interaction with Hobbit?

Perceived safety. Does the perceived safety change over time due to having Hobbit in the household?

Self-efficacy. Is the self-efficacy maintained on the same level due to having Hobbit in the household?

Emotional attachment. Do the users develop an emotional bond with Hobbit over time?

(Perceived) reciprocity. Do the users perceive the interaction with Hobbit as a give-and-take relationship?

Representative target users for the trials were older adults who were still relatively independent, but already in need of some support and at high risk of severe outcomes if emergencies are not recognized in due time. Given that the sample for our field study can be considered small from a statistical point of view (18 participants was the ultimate possible number concerning project resources, i.e., robot platforms built and time), a careful selection of the study sample was crucial. Tightly specifying the participant group provides more reliability given the small group size. Therefore, primary users (PU) were defined according to the following inclusion and exclusion criteria:

Inclusion criteria:

- Men and women aged 75 or older.
- Single-living at home (due to considerations that acceptance of a robot system among senior couples might be lower than for single-living persons).
- At least moderate impairments in one or several of the following areas: mobility/motor skills, sight, and hearing.
- Possible multimorbidity (different impairments).
- Possibly also receiving (moderate) home care; help in the household.
- Sufficient mental capacity to understand the project and ability to give consent.
- Possibility of having an Internet connection in the home (at least mobile Internet).
- Preferably, having a secondary user (SU) such as a relative, friend, or neighbor willing to accompany the PU and let Hobbit call them for help if needed.

Exclusion criteria:

- Bedridden person/complete immobility.
- Any chronic disease that demands regular and longer treatment in hospitals.
- Blindness or deafness or unable to interact according to a user screening procedure.
- Home is unsuitable for Hobbit to be autonomously tested (e.g., stairs, two-storied homes, high doorsteps or other barriers, carpeted flooring (carpet not removable), Internet connection not possible to install (e.g., no signal for mobile Internet and no cable-based solution available).
- 24-hour home-care/regular medical home care that serves to avoid or at least shorten hospitalization of a person.
- Cognitive impairments that render the use and understanding of Hobbit impossible (e.g., progressed Alzheimer's disease).
- Pacemaker.
- Pets moving freely in the home.

Secondary users, such as friends or relatives, were also invited to take part in the trials as support for the PU.

4.2 Trial Procedure

All trials took place in private homes of senior adults in Austria, Greece, and Sweden between April and August 2015.

4.2.1 Screening of the Users and Their Homes. Before the actual trials, the homes of potential users were examined to make sure that they matched the inclusion criteria and to discuss possible necessary changes to the home environment for the trials. Most notably, carpets sometimes had to be removed, as well as mirrors. Thresholds between rooms could sometimes be overcome by using ramps. Other obstacles that had to be removed for the trials were usually cables and wires or pieces

User's home check and screening procedure	Pre-phase	Day 1: Mutual care mode set to device	Midterm assessment (day 11)	After Midterm assessment: Mutual care mode set to companion	End-of-trial assessment (day 21)	Post-phase (one week after end of trial)
	Questionnaires		Questionnaires		Questionnaires	Questionnaires
			Interview with PU		Interview with PU	Interview with PU
					Interview with SU	
					Log data analysis	

Fig. 5. Overview of the trial procedure.

of furniture in narrow spaces. A connection to the Internet was made available where necessary. If a user's home fit the criteria, then potential users were screened for the inclusion and exclusion criteria: difficulties with sight, hearing, and mobility were assessed in a screening questionnaire. Potential users were also asked if they had fallen in the last two years at home and if they were worried about falling. The higher the participants scored on these scales, the lower their confidence in managing activities of their daily routine. The ability to use the robot independent of sight and hearing difficulties was additionally tested with a more practical approach during the screening phase: A tablet with the Hobbit MMUI was shown to potential users to test if they were able to read the display from the real distance. Similarly, speech communication was simulated to assess if the potential users were able to communicate via the speech communication tool. Additionally, experience with computers and mobile phones was assessed beforehand. Based on the results and reported difficulties, users were either included or excluded.

4.2.2 Three Week Trial. The actual trial lasted for three weeks for every participating user and consisted of the following phases (for an overview, see Figure 5). During the introduction and *Pre-phase* assessment the robot was brought into the home. The technical set-up took place and users signed informed consent. The PUs completed the *Pre-phase* assessment questionnaires and were instructed in the use of Hobbit. They also received a manual on how to use the robot. Technical set-up and a step-by-step introduction of the functions of the robot took 2–4 days. Afterward, the user was able to interact with the robot on his/her own. At the *Midterm* assessment (about day 11 of the trial) users were visited again to complete questionnaires and answer a qualitative interview, then the robot's social role was switched from "device-like" to Mutual Care (MuC) "companion" mode. At the *End-of-trial* assessment (last day of the trial, i.e., about day 21) users were interviewed and completed questionnaires once more. Besides, SUs (if available) were also interviewed at the end of the trial. In the *Post-phase* assessment, approximately one week after the trial had ended, we visited PUs again for *Post-phase* questionnaires and a short interview.

4.2.3 Trial Support. Each user was provided with a contact telephone number they could call anytime during the trial in case of questions and problems. It was consciously taken care of to provide PUs with one fixed contact for the whole trial instead of several different members of the consortium. Technical partners provided the user partners with a list of responsibilities depending on the technical issues in question. This made communication simpler in case of technical issues.

Depending on what problem was reported by users, the contact researcher could refer to the list and quickly contact the responsible technical partner for quick trouble-shooting. In the case of more complex issues, visits by technical partners to the user's home were arranged to fix hardware malfunctions.

4.3 Study Design and Evaluation Materials

The fact that Hobbit should autonomously interact with the older adults and also manipulate objects in their homes brought along challenges in terms of a controlled methodological approach. Primarily, we had to ensure that users were able to interact with the robot on their own, without an influencing observer constantly present. To moderate these difficulties, multiple methods were applied to gather insightful data on usability and acceptance. We chose a combination of qualitative measures and quantitative measures (attitudinal as well as behavioral). Data were collected from multiple sources ("multi-informant approach"): data of the PU, interview data from the SU, and data logged by the robot itself.

User behavior was assessed several times during the user trial and continuously tracked through the robot's logging data, respectively. All participants experienced both roles of the robot: "device-like" and "companion" (Mutual Care). For the first 11 days of each trial, the robot was set to "device-mode"; afterward it was changed to "companion-mode." To avoid a bias brought about by expectations or information that could prime the users, participants were not told about this change and its effects beforehand. Due to the small sample size, we decided on this within-design in which all participants experienced the robot in the same fashion. Quantitative and qualitative data collection was scheduled following the change of the MuC mode, i.e., the *Midterm* assessment allowed for collecting baseline data of the period in which the PUs experienced Hobbit for 11 days with all features except MuC. After that, MuC "companion-mode" was switched on and users were surveyed at the *End-of-trial* assessment again. Changes in user acceptance between *Midterm* and *End-of-trial* are therefore mainly due to the MuC feature. Additional *Pre-phase* and *Post-phase* assessments complemented the data source and allowed for changes to be clarified over time.

4.3.1 Questionnaires/Scales. For the quantitative assessment of the interaction with Hobbit, we aimed at using existing, validated scales as much as possible, as the sample size of the trial is rather small. However, for the acceptance indicators emotional attachment and perceived reciprocity, we needed to generate our items. Moreover, feedback from PUs and SUs showed that even already existing scales had to be adjusted for the users in terms of wording or answering format.

Questionnaires used in the trials were: (1) Falls efficacy scale, (2) Self-efficacy scale, (3) Negative Attitudes Towards Robots Scale (NARS), and (4) Self-developed items on emotional attachment and perceived reciprocity.

The Falls-efficacy Scale [45] measures fear of falling by looking at self-reported difficulties with physical and social activities. For our trials, a shortened version was used, which consisted only of the items that deal with the fear of falling in the home.

The Self-efficacy scale [34] is an established questionnaire tool to measure to what extent a person has optimistic self-beliefs to cope and solve problems (in everyday life) on his or her own. This psychometric instrument is available in validated versions in various languages, including German, Greek, and Swedish. The idea of using this questionnaire was to monitor if using the robot could have a positive influence on the PU's self-efficacy, which then could be interpreted as a motivating effect as well as an increase in quality of life and increase of feeling independent at home.

The NARS [26] serves the purpose of measuring negative attitudes towards robots. We expected that the attitude of users would become more positive (or at least not become more negative) in the course of the trials. The NARS exists in different non-standardized translations. To render the

single items easier to understand for the Hobbit trials, the wording was changed where necessary and formulated more objectively.

To study the acceptance indicators of emotional attachment and perceived reciprocity between PU and robot, we generated some specific items, asking users whether they felt the robot was like a companion or a pet, whether they had the feeling that the robot had its own will, and so on.

For the exact items of all questionnaires, see the Appendix.

4.3.2 Interview Guidelines. The in-depth interviews should serve the purpose to learn about individual experiences and opinions, as well as an opportunity for the participants to reflect on their interaction with Hobbit over time. Topics covered in the first interview with the PU were the perception of the robot's behavior and how the user would describe the robot. Topics covered in the second interview were usability aspects (utility, flexibility, and ease of learning) and user acceptance aspects (perceived safety, self-efficacy, emotional attachment, and perceived reciprocity). Similarly, in the last interview topics covered were the utility, perceived safety, emotional attachment, and perceived reciprocity. SUs were interviewed at the end of week three. Their interviews covered the following aspects: the perceived utility of the robot, PU's self-efficacy, and perceived safety.

During the times of the visits, the involved researchers documented how the users interacted with Hobbit. Each interview lasted between 60 and 90 minutes. The interviews were recorded, listened to repeatedly, and summaries of each interview were translated into English and sent to a project partner for further analysis. NVivo (a software for qualitative analysis) was used to search for patterns, themes, and contradictions within one household and between all the households. The interpretations were cross-examined by all the involved researchers, and the results presented in this journal article have been cooperatively generated. The qualitative analysis was based on much more material than the reader will have access to by reading this manuscript. The selected quotes primarily serve as illustrations and have been selected by how they illustrate the most important findings. All questions can be found in the interview guidelines in the Appendix.

4.3.3 Logging Data. To obtain a clearer picture of how often the PUs interacted with Hobbit and how these interactions took place, we recorded timestamped log data on the internal state of the robot and its behavior coordination system. Further input and its modality from the PU, as well as pre-processed input from sub-systems such as the fallen-user detection, gesture recognition, or calendar-based events, were saved to the internal logging system. A database for relevant logging information was built (for more details on data logging see Reference [3]). In the trial analysis, the focus was on the following data on usability and user acceptance indicators:

Flexibility. Logging the number of times different modalities are successfully used.

Utility. Logging the number of robot commands issued by the user.

4.3.4 Pilot Trials. Two pilot trials with older users were conducted with target-users in Vienna, Austria. Those pilot trials yielded feedback for the planned evaluation methodology and also showed further challenges and issues to be taken into consideration when bringing a robot into a real home. This included how to transport the robot, what to look for in the set-up phase, what to take care of in the user's home, and so on. The first took place during July 2014. The individual scenarios were tested with Hobbit in the home environment and technical errors, as well as script failures, were protocolled. Another pilot test took place during September 2014 and lasted for two weeks. Functionalities were tested again, as well as autonomous behavior of Hobbit in a real-life environment and communication of a naive user with Hobbit through the MMUI interface. All evaluation data were tested and then edited based on the received user feedback. The functional limitations of the prototype identified during the two-week pilot test were used to trigger further

development and system integration. Frequent lab trials with feedback from researchers were continued until the start of the actual trials to make the system more robust and improve identified faults.

4.4 Trial Results

We report the data from 16 participating primary users. Broken down by site, the trials were distributed as follows: Vienna, Austria: seven PUs and six SUs; Heraklion, Greece: four PUs and four SUs; and Lund, Sweden: seven PUs and six SUs. The age of PUs ranged from 75 to 89 years with the average age of the sample being 79.75 years. The majority ($n = 14$, 87.5%) of participants were female. In regards to the living situation, 68.8% of the sample (11 users) were living in apartments, and the others in houses. Computer literacy (assessed in the screening questionnaire by the question “How often do you use a computer?”) of the sample was divided between either “never/less than once a week” (37.5%) and “every day” (56.3%). Only one PU stated to be using a computer “two to three times a week.” 14 of the 16 PUs (87.5%) had used a cell phone, six of those (37.5%) were smart-phones. In total, 13 PUs (81.3%) had some form of multiple impairment (e.g., severe vision and also moderate mobility problems). All users fulfilled the sample requirement of having at least one impairment graded as “moderate” or higher. Based on the screening questionnaire, seven users had a severe vision impairment (43.8%), six had a severe hearing impairment (37.5%), and seven had a severe mobility impairment (43.8%). However, all were able to communicate via all modalities offered by the robot as assessed via the practical screening procedure described above.

4.4.1 Insights on Usability. In the following, we present an overview of the findings of usability perception through the various data collection methods split up according to the pre-defined indicators. The guiding research question for the data analysis/interpretation is mentioned in the paragraph heading.

Ease of Learning: Is Hobbit intuitive to interact with older adults?

This indicator was only informed by qualitative interview data. The majority of the PUs thought that Hobbit was easy to use. Many PUs stated that it had been easier to learn how to use Hobbit than they initially expected. As an example, one of the PUs said: *I feel the robot is easy to understand and easy to use. I have read the manual, but I did not have to use the manual because the robot is very intuitive to use.* Many PUs also mentioned that they had been in frequent interaction with the research team and that they had received the needed support to be confident enough to try out the robot. However, when Hobbit did not react as expected to a command, then some participants were unsure if they did something wrong or if the response was due to a malfunction of the robot.

Flexibility: Does Hobbit provide the relevant input modalities that enable the user to command it effectively?

In the interviews, almost all PUs stated that they mainly used the touchscreen and call buttons to communicate with Hobbit. Some would have preferred to have had more verbal communication with Hobbit, but the speech recognition failed most of the time for most of the PUs. Most PUs also found it difficult to remember the gestures and, in most instances, Hobbit did not respond to their gestures. All the PUs expressed a belief that the touchscreen commands work to a higher degree than voice and gestures. Most participants would have liked to have been able to communicate more with Hobbit via speech and some expressed frustration regarding limitations in the verbal interaction with Hobbit.

Additionally, we analyzed the logging data. We assumed that one input modality would be preferred more than others and that specific commands would be more suitable for one input modality than the other. We saw some interesting tendencies in the data concerning preferred interaction

modalities (for details on the numbers, we refer the reader to Reference [3]). At first glance, it seemed that gestures were the most preferred input modality, even more preferred than the touchscreen. This counter-intuitive result was often caused by false positives from the gesture recognition system. This was above all true for the over-sensitive “Come closer” gesture and also, as mentioned above, by falsely interpreted emergency triggers. Similarly, we observed high numbers for the “Come closer” gesture in the first trials. However, content-wise, it is also convincing that “Come closer” and “Yes” are the most frequently used gestures. Similarly, it was convincing that “Yes” was the most often used speech command. The most often used touchscreen inputs were the “Go to” command and the “Recharge” command, which was also a very plausible result. Considering the many false positives in gesture recognition at the beginning of the trials, it is rather the case that the touchscreen was used most, followed by speech and gesture.

Utility: Does the robot offer the right functionalities making older adults feel supported to maintain independent living at home?

To see how actively users interacted with Hobbit, we first had a look at the logging of the clicks carried out by the users on the MMUI, as this is the most robustly logged input. These data reveal that most households show a click count above 1K clicks. In only 6 out of the 16 cases, more clicks were achieved in the MuC condition; however, this tendency in the data has to be handled with care, as the chances of recovering valid logging data were decreasing at later stages of the trials (for details on the numbers, we refer the reader to Reference [3]). The overall click rates are in any case convincing. We also analyzed how often an emergency was actively triggered by the user or automatically triggered by the robot. It has to be mentioned at this point that user triggers were counted as pressing the emergency button on the MMUI, saying help, or using the hardware button on the front of the robot. Robot triggers were: a person lying on the floor, a person currently falling, and the gesture recognition for emergency (as a gesture can also falsely be recognized without any user input it cannot be counted as user input at that point). Comparing the two categories showed that, in general, more emergency calls were issued by the robot ($n = 555$) than pro-actively by the user ($n = 94$). This is easy to explain by false positives from the fall detection and also from the fitness function that often led to falsely recognized emergency gestures. However, it needs to be mentioned that no actual emergency happened during any of the trials. No explainable tendencies for the MuC and Non-MuC conditions could be observed.

In almost all trials *Call Hobbit* was the command issued most often. However, a closer examination also shows that almost all commands were canceled about half of the time. One reason for this high cancellation rate could be that starting one command before another running one is finished is counted as canceling the currently executed task. Secondly, in cases where the robot autonomously started an action, this canceled a user action (a circumstance that caused dissatisfaction on the user side, e.g., when the robot was moving to the docking station while the user was still playing a game).

It was more often the case that users sent the robot to the charging station than the robot going there autonomously. This almost removed the need for Hobbit to use its autonomous recharge capability. In the single situation when it was necessary to recharge on its own, the battery charge had dropped to 9%.

The interview data revealed interesting complementary insights. There was a general concern regarding whether Hobbit offered the right functionalities to maintain living independently at home for older people. Most PUs perceived Hobbit more as a toy than an aid to prolong living at home for a longer time. However, users also said: *He could become convenient ... when all functions would work. The concept is very good. I think it is ideal. I can imagine that, as a result of having a robot, it would be possible to live at home for a longer period. He would need to access all rooms in a*

Table 2. Self-reporting of 18 Users on How Often Each of the 17 Hobbit Functions Was Used and How They Valued the Functions as Working Well, Neither Well Nor Badly, or Badly

Function	Used often	Used sometimes	Worked well	Worked neither well nor badly	Worked badly
Call Hobbit	17	1	16	0	2
Go recharging	17	1	16	0	2
Go to	15	2	15	1	1
Break	13	1	14	0	0
Come closer	12	5	16	1	0
Entertainment	12	4	12	2	2
Calendar/Reminding	8	4	6	3	3
Reward Hobbit	7	6	9	2	2
Information	7	5	11	1	0
Pick up object	4	7	1	1	9
Make Hobbit stop	3	5	6	1	1
Teach object	2	10	4	1	7
Surprise me	2	9	4	3	4
Telephone	2	7	8	0	1
Bring object	2	6	2	2	4
Follow me	2	6	0	1	7
Emergency	0	5	4	0	1

Functions used often by most users sorted first.

home. Hobbit is an inspirational tool.... The robot could increase the quality of life. I live alone and the robot is like a treasure.... At our age, we lack the stimulus. Nothing is very interesting because we have seen it all before. The robot gives stimulus and stimulates activity. It is something new to explore. According to the results of the interviews, 13 of the 17 functions, or 76% (see Table 2), were used often or sometimes by at least 50% of the 18 users. Two of 18 users said they used every function at least once during the trials. For the other functions, one or more users did not use them. Only six users valued the *Calendar/Reminding* function as working well. This could be explained either by the *Calendar/Reminder* function having low reliability (Hobbit not reminding when it should) or by a frequent use of the *Break* function as a way of canceling reminders in the process of their delivery. The *Surprise me* function triggered Hobbit to randomly suggest playing some music or a game. Some of the users liked it while some were confused when Hobbit just showed the music or game menu and did not start playing music or the game. That could be the explanation for why it was not used that often. Finally, the reliability of the *Follow me* command was not sufficiently high to be able to follow the user, but it worked sometimes for one of the users' grandchildren.

4.4.2 Insights on User Acceptance. In the following, we present an overview of the findings of user acceptance through the various data collection methods split up according to the pre-defined indicators. The guiding research question for the data analysis/interpretation is mentioned in the paragraph heading.

Attitude towards robots: Does the general attitude towards robots change over time due to the interaction with Hobbit?

For the quantitative assessment of attitude change during the trial, we used the NARS questionnaire, which consists of three subscales (1): Negative attitude towards interactions with robots

Table 3. Self-reported NARS Questionnaire Scores²⁰

	Pre-phase			Mid-term			End of trial			Post-phase		
	min	max	median	min	max	median	min	max	median	min	max	median
NARS subscale 1	9	18	13	6	19	12.5	9	20	13	6	20	14
NARS subscale 2	11	18	14	7	19	15.5	7	23	14.5	9	22	15
NARS subscale 3	5	13	9	6	15	10.5	4	14	10.5	5	15	11

(highest possible score: 30), (2): Negative attitude towards social influence of robots (highest possible score: 25), (3): Negative attitude towards emotions in interaction with robots (highest possible score: 15). The respective items are added together to reach the score for each subscale. Higher scores reflect a more negative attitude [26].

Table 3 gives an overview of the sub-scale ratings and the different measurement points. Due to the small sample size and not normally distributed data, we calculated non-parametric tests to observe whether the changes in the scores were significant. Friedman tests showed a significant difference only for subscale 3, *Negative attitude towards emotions in interaction with robots* ($\chi^2(3) = 9.217$, $p = 0.027$). The results revealed that the score for the negative attitude towards emotions in interaction with robots was lower (i.e., less negative) in the *Pre-phase* than at all other points of assessment. The strongest difference was found between the *Pre-phase* and the *Post-phase*, as well as *Midterm* assessment. Wilcoxon-matched-pairs post hoc comparisons identified two statistically significant differences. It shows that the attitude towards emotions in interaction with robots was significantly more negative during *Midterm* of the trial ($U = -2.288$, $p = 0.022$, $r = -0.404$) and in the *Post-phase* ($U = -2.763$, $p = 0.006$, $r = -0.488$) than before the trials. The switch from MuC-device to MuC-companion mode did not influence the ratings. To summarize, according to the NARS the general attitude towards robots changed over time for subscale 3. Only 12.5% (two PUs) had a more positive attitude after the trials than before, with 62.5% (10 PUs) of the attitudes becoming more negative. This shows that the trial caused insecurity and anxiety with the older users, which is an effect that has to be taken into careful consideration when introducing new technology or prototypes into the homes of a potentially vulnerable group.

Perceived Safety: Does the perceived safety change over time due to living with Hobbit in the household?

We developed our short version of the Falls Efficacy Scale, consisting of only those items that deal with the fear of falling in the home. For this reason, the validated scoring of the scale cannot be used. For the (unvalidated) Hobbit-version of the scale, we used a score ranging from 7 (no concern about falling) to 28 (severe concern about falling). *Pre-phase*: 87.5% of the sample ($n = 14$) had a score below 13, indicating no concern about falling. The remaining two PUs had a score below 20, indicating a slight concern about falling. *Midterm*: On day 11, 100% of the users reached a score below 13, indicating no concern about falling. *End-of-trial*: At the *End-of-trial* assessment, this distribution had changed. Now, 13 of the users (81.25%) had a score of 12 or less, indicating no concern about falling. Yet three users (18.75%) scored below 20 but above 12, indicating at least a slight concern about falling. *Post-phase*: One week after the end of the trial, all 16 users had scores below 13 again. The distribution of these results is presented in Table 4.

Non-parametric testing (Friedman ranking test) showed that the differences in scores between the four points of assessment were not statistically significant ($\chi^2(3) = 5.571$, $p = 0.134$). To summarize, according to the adapted FES, perceived safety was not significantly influenced by the robot. 43.8% of the users ($n = 7$) had a higher score in the falls efficacy scale in the *Post-phase* than in

²⁰We report median instead of mean values, as the data were not normally distributed.

Table 4. Self-reported Perceived Safety Scores²⁰

	Pre-phase			Mid-term			End of trial			Post-phase		
score	min	max	median	min	max	median	min	max	median	min	max	median
	7	19	9	7	12	8	7	20	9	7	12	10

Table 5. Scores on the Self-efficacy Scale²⁰

	min	max	median
Pre-phase	26	40	31
Mid-term	21	40	30.5
End of trial	20	40	31
Post-phase	24	40	30

the *Pre-phase*; with 37.5% ($n = 6$) the score remained the same. Only three PUs (18.8%) were less concerned about falling in the *Post-phase* in comparison to the *Pre-phase*.

The interview data revealed that all participants—except one—perceived Hobbit as a safe device to be in a home environment. The participant said: *In the first week, he drove around although I've sent him to take a break. I don't think that Hobbit is a safe device to live with because he didn't work all the time. And he couldn't help me. If I would lie on the floor, I wouldn't be able to press the red SOS-button.*

All the participants said that Hobbit did not increase their perceived feeling of safety at home. The opinion that Hobbit did not increase the feeling of safety maintained during the three weeks of trials as well as after Hobbit had moved out. Most mentioned that patrolling is the desired functionality that might increase the feeling of safety, if reliable: *I think that the patrolling function can be very useful for older people who require being looked after.*

Self-Efficacy. The quantitative assessment of Self-Efficacy was done using the self-efficacy scale. The individual score of self-efficacy was generated by adding the values of all items for each user, with a minimum of 10 and a maximum of 40.

Despite the impression that the self-efficacy was lower after the robot trial than before, non-parametric testing revealed that the differences between the points in time were not significant ($\chi^2(3) = 3.972$, $p = 0.265$). As Table 5 shows, the individual self-efficacy did not change during the trial. It did not increase due to having a robot at home. Only 25% of PUs had a higher self-efficacy score at day 21. With three PUs (18.8%) the score remained unchanged, whereas nine participants (56.3%) had a lower self-efficacy score at day 21 than in the *Pre-phase*. A reason for this might be that the self-efficacy of the users decreased slightly due to the robot and problems handling the robot.

The interview data additionally showed that none of the PUs felt that Hobbit had helped them in situations, which they would have had difficulties mastering on their own. However, some mentioned that their level of exercise increased when they had Hobbit due to the fitness program, while others mentioned that Hobbit had kept them busy. Many felt obligated to try out Hobbit and its functionality every day, since they were partaking in the trial. However, by the end of the trial, most of the participants still expressed a belief in robots as helpful in the future.

Emotional Attachment and Perceived Reciprocity. The questionnaire on emotional attachment and perceived reciprocity consisted of 14 self-generated items for the *Pre-phase*, *Midterm* and *End-of-trial* assessment, and a 15th item for the final assessment (*I will miss the robot.*). The questions were grouped by four factors: (1) Emotional Attachment, (2) Reciprocity, (3) Reliability and

Importance, and (4) Machine-like vs. Human-like perception. Data were analyzed in two steps: non-parametric Friedman tests were calculated to identify differences across the four time-points, and a post hoc analysis was done by Wilcoxon-matched-pair comparisons between phases to assess MuC induced changes. In the following, we summarize the key findings for emotional attachment and perceived reciprocity:

Emotional attachment. Initially, users expected the robot to be like a friend or pet. However, as soon as people gained experience with it, they rated it as rather not like a friend or pet. Statistical comparisons between phases are significant. The same is true for Item 9 *empathy with the robot*. Data show that people expected to become emotionally attached, but that this expectation was not fulfilled. There is no difference between the MuC-device and MuC-companion mode. It follows that MuC thus did not contribute to establishing emotional attachment.

Perceived reciprocity. The PUs agreed to Item 2 *I often supported the robot...* consistently across all time-points, and agreed to Item 6 *I support it like a good friend* a little bit less during *Post-phase* assessment—the *good-friend* phrase might be the reason for the drop in the ratings to disagreement during the *Post-phase* assessment. The PUs agreed to Item 2 *I often supported the robot...* consistently across all time-points. The expectation of users to be supported by the robot and to support each other shifted to disagreement: *I was often supported by the robot* and *The robot and I often supported each other* was disagreed upon from the *Midterm* assessment on. To summarize, the users were motivated to support the robot for the whole trial duration. They also expected to be supported by the robot, but the robot could not fulfill this expectation (in analogy to the *Emotional Attachment* factor). The switch from *device-mode* to *MuC-companion-mode* did not influence the ratings.

The interview data revealed additional interesting facts concerning emotional attachment and perceived reciprocity. There was significant ambivalence in terms of feelings toward Hobbit. Participants positively talked about Hobbit. Hobbit was perceived as cute and easy to use, although some of the PUs complained that Hobbit is too big and generated too much heat and noise. One of the PUs expressed this ambivalence eloquently: *He is too big; he gets very hot. He's not a human. He's interesting and amusing, but I don't want to live with him forever*. While Hobbit was perceived as unreliable and underdeveloped, most of the PUs were happy that they had had the chance to use it. Participants envisioned that Hobbit's presence in their home could be like having another living inhabitant (e.g., a pet), although PUs stressed that Hobbit is a machine and not a human. As one of the PUs expressed it: *He moves and he talks. I can feel the presence of him in my house. It is great fun when he is working well*. Overall, all the participants liked the verbal response of the robot. *When he moves and his head moves, I get happy and compassionate. I think that I want to see him like something alive. He's charming*. During the trial, PUs had become familiar with Hobbit's presence and when Hobbit was taken away his presence was missed for a day. However, most PUs were rather happy that life returned to normal and that they could re-establish their home to its original state (e.g., putting back carpets and furniture).

Moreover, most PUs noticed that dialogues changed in the companion mode. Hobbit using the participant's name in the dialogue was perceived as very positive and was appreciated. However, the *return of a favor* did not make sense to most of the PUs. It was perceived as an annoyance. None of the PUs noticed increased interaction and presence of Hobbit in the companion mode compared to the device mode. During our field trials, the need for the predictable behavior of the robot became increasingly evident, as users sometimes failed to understand the robot's actions during an interaction scenario.

4.5 Additional Insights

The trials yield a great deal of insight on how Hobbit can be improved despite the trials having been plagued with technical issues. All participants contributed valuable insights and ideas for improvements, such as:

- A battery level icon that is always visible on the screen.
- A screen saver is necessary, since it would save both energy and screen.
- The display is too bright to be placed in the bedroom during the night.
- The *Recharge batteries* command should be placed under the *Go To* command.
- The tray should not be divided into two so that one can use Hobbit for transporting, e.g., a couple of coffee cups, a plate of food.
- The sound level of the fan is too high.
- Hobbit discharges a significant amount of heat, which warms up the participants' homes.
- Much shorter response times to the different commands.
- Hobbit should have a round body shape to better operate without bumping into walls, furniture, and people while turning around.
- Generally, Hobbit is perceived as too large, resulting in poor mobility in usual homes.
- The ability of the robot to reverse is desirable. Users mentioned that they would like to operate the robot with the help of voice or remote control "right, left, forward, reverse, and stop" during operation. Most participants think that it is too limited for the robot to just be able to go from a predetermined point to another predetermined point.
- Possibility to shut it down completely and for the user to be able to turn it on, possibly with passwords.
- The *Safety checklist* and *Reminders* should have a time delay before the answer OK is accepted. This is to prevent one from routinely pressing OK without even thinking whether you have turned off the stove, and so on.
- Regarding the calendar, a user should be able to erase an incorrect input.
- Moreover, one should be able to "step-through" the month and maybe even a week for a long period.

4.6 Discussion

Usability. Features such as *delivering reminders* and *robust navigation in different scenarios* were among the highest rated features and are in line with previous work [9, 18, 33, 37]. Only a small number of features were rated as working bad, mainly caused by their lack of sufficient robustness (Table 2). This suggests that most functionalities Hobbit provided are useful for the elderly living alone, were well received, and should be considered by developers of an elderly-care robot. Our participants were able to reliably interact with Hobbit over a longer period, even though most of them were not too familiar with technology such as robots in general or devices like smart-phones in particular.

User Acceptance. Previous work indicates that attention needs to be paid to how robots are introduced into the home of users, as this will have a lasting impact on a user's established routines. For instance, Forlizzi [16] suggests that *Homes and service robots must adapt to each other*. A recurring theme in long-term studies is the robustness and physical limitations of the deployed robots. Fazekas [12] stresses that essential features such as safe, autonomous navigation, and speech recognition need to improve until they reach an acceptable level of robustness. Broadbent [6] implies that the low usage of their robot may be due to its immobility, thus leaving a certain amount of potential users without a chance to interact with the robot. This suggests that

the choice to have a stationary robot needs to be well thought through, and should be only done based on the outcome of a user-centered design approach [29].

While the quantitative data did not reveal any increase in user acceptance, the qualitative results do suggest that Hobbit was well received. Participants noted that they *feel the presence of him* and *want to see him as something that is alive*; they also stated, *I help the robot, it helps me*, something we see as an indication that they want to see more in Hobbit than simply a tool.

We strongly believe that just adding more features (e.g., transporting people, help them stand up after a fall) or changes to the appearance (e.g., higher levels of anthropomorphism) to a robot will not be enough to increase the user acceptance of the utility. A strong focus needs to be placed to improve the reliability of all aspects of the robotic system so long-term effects can finally be observed without the influence of the sub-optimal system.

5 CONCLUSION

Exploring how older adults interact with a mobile social service robot in their private homes provides important insights not only for the acceptance of the specific robot but for the design and further development of this type of technology in general. Our investigation on how users experience the Hobbit robot in their private space over a longer period has revealed not only success stories, but also obstacles for acceptance and technological challenges. This knowledge may enable developers to better adapt robot innovations for the care domain that benefit older adults and increase the chances of robot-supported independent aging. The results of our field trial showed that all users interacted with Hobbit daily, rated most functions as well working, and reported that they believed Hobbit will be part of future elderly care. However, the results also highlighted several technical/ practical issues, some which can be easily fixed (e.g., having a screen saver, bigger tray) and some that will need substantial improvements in, e.g., navigation, grasping, and behavior coordination. The challenge for HRI researchers and developers is to create robots that can autonomously safely and robustly operate in end-users' homes and provide functionalities that are experienced as relevant and are robustly performed by the robot. End users in our study provided vital quantitative and qualitative data that demonstrated the high motivation and willingness of older adults to use such technology and to integrate it into their everyday lives, as well as that end-users had the necessary skills and self-efficacy to do so.

REFERENCES

- [1] Aitor Aldoma, Federico Tombari, Luigi Di Stefano, and Markus Vincze. 2012. A global hypotheses verification method for 3d object recognition. In *Proceedings of the European Conference on Computer Vision*. Springer, 511–524.
- [2] Markus Bajones, David Fischinger, Astrid Weiss, Daniel Wolf, Markus Vincze, Paloma de la Puente, Tobias Körtner, Markus Weninger, Konstantinos Papoutsakis, Damien Michel, Ammar Qammaz, Paschalis Panteleris, Michalis Foukarakis, Ilia Adami, Danai Ioannidi, Asterios Leonidis, Margherita Antona, Antonis Argyros, Peter Mayer, Paul Panek, Håkan Efrting, and Susanne Frennert. 2018. Hobbit: Providing fall detection and prevention for the elderly in the real world. *J. Robot.* 2018 (2018). DOI: <https://doi.org/10.1155/2018/1754657>
- [3] Markus Bajones, Astrid Weiss, and Markus Vincze. 2016. I know what you did last week. Log data analysis of long term household trials. In *Proceedings of the Workshop on "The Challenge (Not) to Go Wild! Challenges and Best Practices to Study HRI in Natural Interaction Settings" at HRI'16*.
- [4] Jenay M. Beer, Cory-ann Smarr, Tiffany L. Chen, Akanksha Prakash, Tracy L. Mitzner, Charles C. Kemp, and Wendy A. Rogers. 2012. The domesticated robot: Design guidelines for assisting older adults to age in place. In *Proceedings of the 7th ACM/IEEE International Conference Human-Robot Interaction (HRI'12)*. 335–342.
- [5] Elizabeth Broadbent, Jeff Garrett, Nicola Jepsen, Vickie Li Ogilvie, Ho Seok Ahn, Hayley Robinson, Kathryn Peri, Ngaire Kerse, Paul Rouse, Avinesh Pillai et al. 2018. Using robots at home to support patients with chronic obstructive pulmonary disease: Pilot randomized controlled trial. *J. Med. Internet Res.* 20, 2 (2018).
- [6] Elizabeth Broadbent, Ngaire Kerse, Kathryn Peri, Hayley Robinson, Chandimal Jayawardena, Tony Kuo, Chandan Datta, Rebecca Stafford, Haley Butler, Pratyusha Jawalkar et al. 2016. Benefits and problems of health-care robots in aged care settings: A comparison trial. *Australasian J. Age.* 35, 1 (2016), 23–29.

- [7] Paolo Dario, Eugenio Guglielmelli, Vincenzo Genovese, and Maurizio Toro. 1996. Robot assistants: Applications and evolution. *Robot. Auton. Syst.* 18, 1–2 (1996), 225–234.
- [8] Kerstin Dautenhahn. 2007. Methodology & themes of human-robot interaction: A growing research field. *Int. J. Adv. Robot. Syst.* 4, 1 SPEC. ISS. (2007), 103–108. DOI: <https://doi.org/10.5772/5702>
- [9] Paloma de la Puente, Markus Bajones, Christian Reuther, Daniel Wolf, David Fischinger, and Markus Vincze. 2019. Robot navigation in domestic environments: Experiences using RGB-D sensors in real homes. *J. Intell. Robot. Syst.* 94, 2 (2019), 455–470.
- [10] Andrew Dillon. 2001. User acceptance of information technology. *Encyclopedia of Human Factors and Ergonomics*. Retrieved from <http://hdl.handle.net/10150/105880> <http://arizona.openrepository.com/arizona/handle/10150/105880>.
- [11] R. Eisma, A. Dickinson, J. Goodman, A. Syme, L. Tiwari, and A. F. Newell. 2004. Early user involvement in the development of information technology-related products for older people. *Univ. Access Inform. Soc.* 3, 2 (2004), 131–140. DOI: <https://doi.org/10.1007/s10209-004-0092-z>
- [12] Gabor Fazekas, Andras Toth, Pierre Rumeau, Katalin Zsiga, Tamas Pilissy, and Vincent Dupurque. 2012. Cognitive-care robot for elderly assistance: Preliminary results of tests with users in their homes. In *Proceedings of the AAL Forum*. 145–148.
- [13] Heike Felzmann, Kathy Murphy, Dymna Casey, and Oya Beyan. 2015. Robot-assisted care for elderly with dementia: Is there a potential for genuine end-user empowerment. In *Proceedings of the 10th ACM/IEEE International Conference on Human–Robot Interaction*.
- [14] David Fischinger, Peter Einramhof, Konstantinos Papoutsakis, Walter Wohlking, Peter Mayer, Paul Panek, Stefan Hofmann, Tobias Körtner, Astrid Weiss, Antonis Argyros, and Markus Vincze. 2016. Hobbit, a care robot supporting independent living at home: First prototype and lessons learned. *Robot. Auton. Syst.* 75 (2016), 60–78. DOI: <https://doi.org/10.1016/j.robot.2014.09.029>
- [15] David Fischinger, Astrid Weiss, and Markus Vincze. 2015. Learning grasps with topographic features. *Int. J. Robot. Res.* 34, 9 (2015), 1167–1194. DOI: <https://doi.org/10.1177/0278364915577105>
- [16] Jodi Forlizzi and Carl Disalvo. 2006. Service robots in the domestic environment: A study of the Roomba vacuum in the home. *Design* 2006 (2006), 258–265. DOI: <https://doi.org/10.1145/1121241.1121286>
- [17] Birgit Graf, Ulrich Reiser, Martin Hägele, Kathrin Mauz, and Peter Klein. 2009. Robotic home assistant Care-O-Bot® 3—Product vision and innovation platform. In *Proceedings of the IEEE Workshop on Advanced Robotics and Its Social Impacts (ARSO'09)*. 139–144. DOI: <https://doi.org/10.1109/ARSO.2009.5587059>
- [18] Horst-Michael Gross, Steffen Mueller, Christof Schroeter, Michael Volkhardt, Andrea Scheidig, Klaus Debes, Katja Richter, and Nicola Doering. 2015. Robot companion for domestic health assistance: Implementation, test and case study under everyday conditions in private apartments. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'15)*. 5992–5999. DOI: <https://doi.org/10.1109/IROS.2015.7354230>
- [19] Chandimal Jayawardena, I-Han Kuo, Elizabeth Broadbent, and Bruce A. MacDonald. 2016. Socially assistive robot healthbot: Design, implementation, and field trials. *IEEE Syst. J.* 10, 3 (2016), 1056–1067.
- [20] Takayuki Kanda, Masahiro Shiomi, Zenta Miyashita, Hiroshi Ishiguro, and Norihiro Hagita. 2009. An affective guide robot in a shopping mall. In *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction (HRI'09)*. ACM, New York, NY, 173–180. DOI: <https://doi.org/10.1145/1514095.1514127>
- [21] Tobias Körtner, Alexandra Schmid, Daliah Batko-Klein, Christoph Gisinger, Andreas Huber, Lara Lammer, and Markus Vincze. 2012. How social robots make older users really feel well—A method to assess users' concepts of a social robotic assistant. In *Proceedings of the International Conference on Social Robotics*. Springer, 138–147. DOI: https://doi.org/10.1007/978-3-642-34103-8_14
- [22] I-Han Kuo, Elizabeth Broadbent, and Bruce MacDonald. 2008. Designing a robotic assistant for healthcare applications. In *Proceedings of the 7th Conference of Health Informatics*.
- [23] Lara Lammer, Andreas Huber, Astrid Weiss, and Markus Vincze. 2014. Mutual care: How older adults react when they should help their care robot. In *Proceedings of the 3rd International Symposium on New Frontiers in Human-Robot Interaction*. Retrieved from <http://hobbit.acin.tuwien.ac.at/publications/AISB2014-HRIpaper.pdf>.
- [24] Iolanda Leite, Carlos Martinho, and Ana Paiva. 2013. Social robots for long-term interaction: A survey. *Int. J. Soc. Robot.* 5, 2 (2013), 291–308. DOI: <https://doi.org/10.1007/s12369-013-0178-y>
- [25] J. C. Marquié, L. Jourdan-Boddaert, and N. Huet. 2002. Do older adults underestimate their actual computer knowledge? *Behav. Inform. Technol.* 21, 4 (2002), 273–280. DOI: <https://doi.org/10.1080/0144929021000020998>
- [26] Tatsuya Nomura, Tomohiro Suzuki, Takayuki Kanda, and Kensuke Kato. 2006. Measurement of negative attitudes toward robots. *Interact. Stud.* 7, 3 (2006), 437–454. DOI: <https://doi.org/10.1075/is.7.3.14nom>
- [27] Martha E. Pollack, Laura Brown, Dirk Colbry, Cheryl Orosz, Bart Peintner, Sailesh Ramakrishnan, Sandra Engberg, Judith T. Matthews, Jacqueline Dunbar-Jacob, Colleen E. McCarthy et al. 2002. Pearl: A mobile robotic assistant for the elderly. In *Proceedings of the AAAI Workshop on Automation as Eldercare*, Vol. 2002. 85–91.

- [28] Johann Prankl, Aitor Aldoma Buchaca, Alexander Svejda, and Markus Vincze. 2015. RGB-D object modelling for object recognition and tracking. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'15)*. 8.
- [29] Tobias Rehr, Raphaël Troncy, Andreas Bley, Susanne Ihlen, Katharina Scheibl, Sebastian Glende, Stefan Goetze, Jens Kessler, Christoph Hintermueller, and Frank Wallhoff. 2012. The ambient adaptable living assistant is meeting its users. In *Proceedings of the AAL Forum*. 629–636. DOI : <https://doi.org/10.1.1.364.3969>
- [30] Yvonne Rogers. 2011. Interaction design gone wild. *Interactions* 18, 4 (July 2011), 58. DOI : <https://doi.org/10.1145/1978822.1978834>
- [31] Astrid Marieke Rosenthal-von der Pütten, Astrid Weiss, and Selma Šabanović. 2016. The challenge (not) to go wild!: Challenges and best practices to study HRI in natural interaction settings. In *Proceedings of the 11th ACM/IEEE International Conference on Human Robot Interaction (HRI'16)*. IEEE Press, Piscataway, NJ, 583–584. Retrieved from <http://dl.acm.org/citation.cfm?id=2906831.2906991>.
- [32] Selma Sabanovic, Marek P. Michalowski, and Reid Simmons. 2006. Robots in the wild: Observing human-robot social interaction outside the lab. In *Proceedings of the International Workshop on Advanced Motion Control (AMC'06)*, Vol. 2006. 576–581. DOI : <https://doi.org/10.1109/AMC.2006.1631758>
- [33] Ch. Schroeter, S. Mueller, M. Volkhardt, E. Einhorn, C. Huijnen, H. Van Den Heuvel, A. Van Berlo, A. Bley, and H. M. Gross. 2013. Realization and user evaluation of a companion robot for people with mild cognitive impairments. In *Proceedings of the IEEE International Conference on Robotics and Automation*. 1153–1159. DOI : <https://doi.org/10.1109/ICRA.2013.6630717>
- [34] Ralf Schwarzer and Matthias Jerusalem. 1995. Generalized self-efficacy scale. In *Measures in Health Psychology: A User's Portfolio. Causal and Control Beliefs*. Number 2008. Elsevier Publishers, 35–37. DOI : <https://doi.org/10.1037/t00393-000>
- [35] T. Shibata and K. Tanie. 2001. Physical and affective interaction between human and mental commit robot. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'01)*, Vol. 3. 2572–2577. DOI : <https://doi.org/10.1109/ROBOT.2001.933010>
- [36] Masahiro Shiomi, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. 2006. Interactive humanoid robots for a science museum. In *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction (HRI'06)*. ACM, New York, NY, 305–312. DOI : <https://doi.org/10.1145/1121241.1121293>
- [37] Cory-Ann Smarr, Akanksha Prakash, Jenay M. Beer, Tracy L. Mitzner, Charles C. Kemp, and Wendy A. Rogers. 2012. Older adults' preferences for and acceptance of robot assistance for everyday living tasks. In *Proceedings of the Human Factors and Ergonomics Society Meeting*, Vol. 56. SAGE Publications, Los Angeles, CA, 153–157.
- [38] Walter Dan Stiehl, Cynthia Breazeal, Kuk-Hyun Han, Jeff Lieberman, Levi Lalla, Allan Maymin, Jonathan Salinas, Daniel Fuentes, Robert Toscano, Cheng Hau Tong, et al. 2006. The huggable: A therapeutic robotic companion for relational, affective touch. In *Proceedings of the ACM SIGGRAPH Conference on Emerging Technologies*. ACM, 15.
- [39] Jayoung Sung, Rebecca E. Grinter, and Henrik I. Christensen. 2009. "Pimp My Roomba": Designing for personalization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*1 (2009), 193–196. DOI : <https://doi.org/10.1145/1518701.1518732>
- [40] Markus Vincze, Markus Bajones, Markus Suchi, Daniel Wolf, Astrid Weiss, David Fischinger, and Paloma da la Puente. 2016. Learning and detecting objects with a mobile robot to assist older adults in their homes. In *Proceedings of the European Conference on Computer Vision*. Springer, 316–330.
- [41] K. Wada and T. Shibata. 2007. Living with seal robots—Its sociopsychological and physiological influences on the elderly at a care house. *IEEE Trans. Robot.* 23, 5 (Oct. 2007), 972–980. DOI : <https://doi.org/10.1109/TRO.2007.906261>
- [42] K. Wada, T. Shibata, T. Saito, K. Sakamoto, and K. Tanie. 2005. Psychological and social effects of one year robot assisted activity on elderly people at a health service facility for the aged. In *Proceedings of the IEEE International Conference on Robotics and Automation*. 2785–2790. DOI : <https://doi.org/10.1109/ROBOT.2005.1570535>
- [43] K. Wada, T. Shibata, T. Saito, and K. Tanie. 2004. Effects of robot-assisted activity for elderly people and nurses at a day service center. *Proc. IEEE* 92, 11 (Nov. 2004), 1780–1788. DOI : <https://doi.org/10.1109/JPROC.2004.835378>
- [44] Astrid Weiss, Lara Lammer, Andreas Huber, Markus Vincze, Nina Hess, Tobias Körtner, and Alexandra Schmid. 2014. Developing an assistive robot for older adults: Methodological considerations for field trials. In *Proceedings of the Workshop on "Socially Assistive Robots for the Elderly: Are We Trapped in Stereotypes?" at HRI'14*.
- [45] Lucy Yardley, Nina Beyer, Klaus Hauer, Gertrudis Kempen, Chantal Piot-Ziegler, and Chris Todd. 2005. Development and initial validation of the Falls Efficacy Scale-International (FES-I). *Age and Ageing* 34, 6 (2005), 614–619. DOI : <https://doi.org/10.1093/ageing/afi196>
- [46] Katalin Zsiga, András Tóth, Tamás Pilissy, Orsolya Péter, Zoltán Dénes, and Gábor Fazekas. 2018. Evaluation of a companion robot based on field tests with single older adults in their homes. *Assist. Technol.* 30, 5 (2018), 259–266. DOI : <https://doi.org/10.1080/10400435.2017.1322158>

Received November 2017; revised August 2019; accepted September 2019