Robotic Versus Human Coaches for Active Aging: An Automated Social Presence Perspective

Martina Čaić[1] · João Avelino[2] · Dominik Mahr[1,3] · Gaby Odekerken-Schröder[1] · Alexandre Bernardino[2]

Abstract

This empirical study compares elderly people's social perception of human versus robotic coaches in the context of an active and healthy aging program. In evaluating hedonic and utilitarian value perceptions of exergames (i.e., video games integrating physical activity), we consider elderly people's judgments of the warmth and competence (i.e., social cognition) of their assigned coach (human vs. robot). The field experiments involve 58 elderly participants in the real-life context. Leveraging a mixed-method approach that combines quantitative and qualitative data, we show that (1) socially assistive robots activate feelings of (automated) social presence (2) human coaches score higher on perceived warmth and competence relative to robotic coaches, and (3) social cognition affects elderly people's experience (i.e., emotional and cognitive reactions and behavioral intentions) with respect to exergames. These findings can inform future developments and design of social robots and systems for their smoother inclusion into elderly people's social networks. In particular, we recommend that socially assistive robots take complementary roles (e.g., motivational coach) and assist human caregivers in improving elderly people's physical and psychosocial well-being.

Keywords Automated social presence · Social cognition · Socially assistive robots · Exergames · Elderly care · User experience

1 Introduction

Many societies face the challenges of aging populations, at risk of reduced physical activity [54], whereas an active lifestyle has proven health-related benefits for

✉ Martina Čaić
  m.caic@maastrichtuniversity.nl

  João Avelino
  javelino@isr.tecnico.ulisboa.pt

  Dominik Mahr
  d.mahr@maastrichtuniversity.nl

  Gaby Odekerken-Schröder
  g.odekerken@maastrichtuniversity.nl

  Alexandre Bernardino
  alex@isr.tecnico.ulisboa.pt

1 School of Business and Economics, Maastricht University, Maastricht, The Netherlands

2 Institute for Systems and Robotics Instituto Superior Técnico, University of Lisbon, Lisbon, Portugal

3 Service Science Factory, Maastricht University, Maastricht, The Netherlands

aging adults [3]. When a sedentary lifestyle becomes routine though [5], falls among elderly people can develop into an alarming problem, often leading to hospitalization and reduced physical autonomy [77]. Evidence links decreased motivation for physical activity to advanced age [67], which implies the need for healthcare systems to develop effective solutions to ensure the physical well-being of elderly people. Exergames offer an innovative way for seniors to avoid a sedentary lifestyle and combat the degenerative effects of aging; their easy-to-follow steps and gamified nature motivate seniors to remain physically active through playful interactions. The application of exergames in healthcare settings also has yielded positive outcomes for both physical and cognitive well-being [9, 76]. Recent studies of the use of exergames in rehabilitation [16, 71] suggest comparable or slightly better results (e.g., balance, gait) than achieved with conventional fitness programs.

According to a study of the usability of exergame platforms [70], elderly people's cognitive deficiencies can hinder their engagement with exergames though. We propose that a social agent can serve a supportive function, guiding and

motivating elderly people through the gaming process. Professionals in care institutions do their best to keep elderly patients physically and mentally active, but the unprecedented relative growth of the elderly population is creating a gap between the supply of and demand for care services [42]. As an aid, gaming platforms can include virtual agents that serve in guidance and motivational roles, with an automated social presence. Still, the presence of a physical entity enhances interactions with an autonomous agent, compared with a virtual or real entity presented on screen [49, 58, 69]. These findings have motivated researchers to turn their attention to socially assistive robots (SAR), like Vizzy [51], MBOT [74], and GrowMu [57], which understand social cues through facial and voice recognition technology, are capable of engaging in quasi-social interactions, and can cause users to feel as if they are in the company of a social entity [14, 72].

The extent to which current SARs are capable of helping people enjoy and perceive exergames as useful is a crucial research topic. The adoption of a user-centered design perspective, with continuous benchmarks and comparisons between SARs and care professionals, can support technological improvements that better suit professionals and elderly patients. In this study, we take the perspective of elderly users to address the following research question: How does an automated (i.e., SAR) versus human presence affect elderly people's experience with exergames to improve their physical activity? For this purpose, we use the Vizzy robot (Fig. 1) as a robotic coach and the Portable Exergames Platform for Elderly (PEPE) [70] as an exergame system. Vizzy is a semiautonomous robot, and PEPE is a fixed gaming platform. These systems were developed and designed specifically for use in elderly care institutions. With

an experimental design, applied in five elderly care locations, we investigate users' experience with playing one of PEPE's exergames, in the company of a human or robotic coach (Vizzy), by gathering their responses to a questionnaire augmented with a set of probing questions. The resulting contributions and practical implications include:

- An empirical test of the new concept of automated social presence and how it influences exergame experiences for seniors.
- Quantitative and qualitative comparisons of the social perceptions elderly people form about human and robotic coaches.
- Evaluations of elderly people's experiences (i.e., emotional and cognitive reaction and behavioral intentions) with the exergame, depending on the company.

In Sect. 2, we discuss prior end-user studies with SARs in motivational and conversational roles, which constitutes the background for our conceptual model in Sect. 3. We present the hypotheses in Sect. 4. Next, we describe Vizzy, PEPE, the research setup, experimental design, and data collection procedures in Sect. 5. After we detail the results in Sect. 6, we discuss contributions and implications of our study in Sect. 7.

2 Related Work

2.1 Social Robots for Older Adults

Applications of SARs in elderly care constitute a trending research topic. Several studies with older adults test their acceptability and possible range of applications. In Iwamura et al.'s work [41], seniors used two robots as shopping partners, and the findings suggest that the robots' social skills (e.g., conversation) improve people's intentions to use them. McColl and Nejat [50] present an exploratory study at an elderly care facility to investigate user engagement and compliance during mealtime interactions with a robot. Participants were engaged and compliant with the robot's instructions and also perceived the robot as enjoyable.

Other studies have investigated the evolution of seniors' perceptions of SARs over a longer period of time. In a study of which factors determine long-term user acceptance of social robots [34], the evidence reveals that hedonic factors gain the most attention, but utilitarian factors are a fundamental prerequisite of long-term interactions (i.e., the robot must have a clear purpose). A more recent contribution investigates older adults with dementia at residential aged care facilities over four years [46]. After each trial, the robot's services improved, reflecting staff and resident feedback, leading to statistically significant increases in engagement and robot acceptability.

Fig. 1 Vizzy interacting with an elderly lady

2.2 Robotic Coaches

Robotic coaches are platforms that engage, monitor, and support physical exercise activities, through verbal and non-verbal communication, demonstration of exercises, and real-time corrective or motivational feedback. An early study [25] explored interaction enjoyment and perceived utility, with a comparison of two different approaches. The first relied on a relational robot that praised, addressed users by name, and expressed humor and empathy. The second used a non-relational robot that just provided scores and help as needed. Users regarded the relational robot as a better companion and exercise coach than the non-relational one. Robotic coaches thus need to provide motivational support to users during the exercise. In turn, users perceive relational robots as more intelligent and more helpful. Feasibility tests of a robot coach architecture with older adults also indicate promising acceptability results [24]. This architecture includes a rule-based decision process but also can learn users' preferences, estimate their affective state, and recall past interactions. Users' perceptions of their interaction with the robot thus improved, relative to their initial expectations. Schneider and Kümmert [66] also investigate the differences in motivation created by a robot that solely instructs users (instructor role) and one that exercises alongside the user (companion role); people appear more motivated by the robot companion role. Finally, in a robot coaching scenario [33], an elderly user study cites engagement and enjoyment while exercising but also reveals some difficulties, such as hearing instructions, focusing on the physical aspects of the robot while ignoring verbal instructions, and confusion associated with performing sequences of gestures.

2.3 Comparing Human and Robot Actors

Extant studies also compare the performances of human and robotic actors in a variety of contexts. To summarize these findings, we start with studies in which participants have no direct interaction with the human/robotic actor, then continue on to studies in laboratory settings, and finally end our analysis in this section with findings from real-world field experiments.

The first group of studies relies on online surveys that present participants with written [75] or video recorded [39] scenarios, followed by a set of questions designed to evaluate their perceptions of human versus robotic actors. The foci of the studies varies, from moral dilemmas to the communication ability of the human or robot agent. For example, one online study [75] reports that participants assign higher moral obligations to human than robotic actors, such that they believe it is morally wrong for a hypothetical human actor to choose to act and save four, while sacrificing one, coal miner, but they expect the robot to make this choice when

faced with the same moral dilemma. Another study [39] compares human and robotic doctors' ability to inform patients about their health conditions, following a standardized script. Participants evaluate the robot better at communicating bad news, though these researchers predict that human social presence might invert the results in a real-life scenario. Although the design of these experiments allow for replication and comparability between humans and robots, they have limited external validity, because they refer to specific, artificial scenarios, and participants have no contact with actors or no immersion in the scenario. Such a limitation is especially important in critical scenarios [39], such as when participants do not experience a real health diagnosis.

Another group of studies relies on laboratory-based methodologies that enable participant–agent interactions. These studies generally report no significant differences in the behaviors of participants in the presence of robotic versus human agents. For example, in a shared attention memorization task with a human or robot agent [47], the participant and agent point at images and look at them together. Later, participants reviewed another set of pictures and had to identify which they had seen before. Recognition performance was better when the participant initiated the pointing behavior than when the collaborator did so, but it did not matter if that collaborator was a human or a robot. In a turn-based variant of the Tower of Hanoi experiment, conducted with a human or robot collaborator [44], participants performed significantly better with a collaborator than when playing individually, but again, there were no significant differences between the human and robot collaborator conditions. Because the collaborator always made the best move, the lack of performance differences could be because the game was too simple. Although these studies suggest that participants benefit equally from collaborating with other humans or robots, one contribution [44] assigns an advantage to humans, in terms of perceived social presence and perceived credibility. These designs are replicable and controlled, and they rely on participant–collaborator interactions. However, the laboratory setting might limit internal and external validity.

Finally, some studies use realistic field experiments to compare human and robotic teachers [28, 45]. For example, when teachers gave a tutorial on prime numbers, no significant differences emerged in the student evaluations of human or robotic teachers [45]. The robot in this study was fully autonomous but could not adapt its gestures to events in the classroom or engage in mutual gazes with students. In a similar study of programming principles [28], students between 14 and 17 years of age performed better with the robotic teacher, whereas those 10–13 years of age performed better with a human teacher. Younger students (6–9 years) showed no significant performance differences. Although they do not suffer the limitations of the previously mentioned studies, strong novelty effects are present in these

studies, especially [28]. Beyond the novelty of interacting with a robot, the participants were visiting students, unfamiliar with the university environment.

This literature review reveals that prior research has focused on different domains of social interaction, including moral expectations, diagnosis communication, shared attention benefits, gaming performance effects with a collaborator, and teaching. They also provide valuable insights into the issues and methodologies that are required to achieve successful experiments involving human and robot actors. First, in particular, they reveal that people evaluate humans and robots differently, even if the robot is perceived to be intelligent [75]. Second, lab experiments allow for more control over the stimuli, but they might not be valid in real-life scenarios [39]. Third, some elements of these methodologies are standard and necessary for interaction studies, to increase their internal validity. The human actor must be trained to follow the same script and move similarly to the robot; the task should match the robots' skills to reduce possible biases. However, literature is still sparse in this domain, and to the best of our knowledge, no study compares older adults' perceptions of exergames when they are coached by either a human or a robot. The following section substantiates the importance and relevance of such a study.

3 Conceptual Background

Social interactions are not reserved to humans. Technologies mimic humans, in their appearance and behaviors (i.e., social robotics, [11]), and quasi-social interactions increasingly appear in diverse service settings, such as robotic waiters, bank tellers, and receptionists. These artificially intelligent social actors can engage in quasi-social interactions [6]. Robots are not social by nature but instead are programmed to act as conversational partners and socially interactive peers [32]. Yet, when they engage in these quasi-social interactions, humans tend to perceive robotic systems as social beings [36, 60]. This phenomenon is closely linked to the concept of social presence, broadly defined as the sense of being with others (e.g., in virtual reality, [38]) or having the access to others' minds (i.e., cognitive, affective, and intentional states, [7]). Social presence is a property of humans, not technologies [6], so current theories may benefit from using social presence as a lens to compare interactions with either human or non-human actors. The concept of automated social presence, as introduced in service literature, refers to "the extent to which machines (e.g., robots) make consumers feel that they are in the company of another social entity" ([72], p. 44). If robots can interact with people in human-like ways, such as by listening, conversing, or reading emotional cues, then people might automatically respond socially to them [60]. In particular, this study argues that people may activate their

social cognition and judge the non-human social actor on the basis of warmth and competence dimensions, which are the universal dimensions of social cognition.

3.1 Social Cognition

According to Fiske et al. [31], when interacting with members of the same species, humans need to determine whether the other is friend or foe (i.e., warmth dimension), and whether it is capable of acting on good or ill intentions (i.e., competence dimension). Although robots are not members of the same species, their anthropomorphism [53] and increasing social dexterity make them resemble humans in many aspects. This study in particular focuses on SARs, which offer assistance through social interactions in a humanlike manner [26]. In a healthcare context, SARs are autonomous, understand social cues through facial and voice recognition, and can provide both child (e.g., autism therapy, [65]) and elderly (e.g., medication reminders, [13]) care. In an elderly care context, SARs offer services such as health monitoring and safety, encouragement to engage in rehabilitation or general health-promoting exercises, social mediation, interactions, and companionship [27]. By performing more socially engaging tasks, these robots take roles as companions, collaborators, partners, pets, or friends [21, 32]. We build on the idea that humans judge a social interactor (even a non-human actor) on warmth (i.e., friendliness, kindness, caring) and competence (i.e., efficacy, skill, confidence) dimensions, which then determines their emotional, cognitive, and behavioral reactions. Thus we address recent calls for research on the effects of automated social presence on the perceived enjoyment and usefulness of service interactions and the acceptance of new technologies [36, 72].

4 Hypotheses Development

We test our empirical model through field experiments with elderly people. The research aim is to investigate the differences between human and robotic companions when it comes to motivating elderly persons to engage in physical activity. We introduce hypotheses related to robotic coaches only (H1), comparisons of human and robotic coaches (H2 and H3), and user experiences with the exergame (H4 and H5). Figure 2 depicts the empirical model.

Using the argument that humans tend to treat non-human actors/systems as social beings [36, 60] and the definition of automated social presence [72], we investigate whether SARs, which display human-like social behavior and communicative skills [10], evoke feelings of social presence in humans. Furthermore, this study extends previous research [30, 43] by involving socially dexterous robots that are capa-
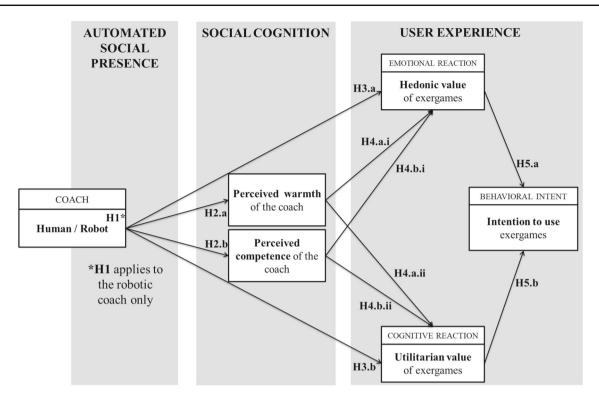
Fig. 2 Empirical model. *Notes*: H1 is tested with one-sample t-tests; H2 and H3 are tested with independent sample t-tests; H4 and H5 are tested with seemingly unrelated regressions (SUR)

ble of engaging in a dialog, gazing, and responding to social cues. Thus, we hypothesize:

H1 Robotic coaches are perceived to have automated social presence (i.e., people feel they are in the company of a social entity).

If people feel like they are in the company of a social entity, they likely evaluate robots using the mechanisms of social cognition (i.e., warmth and competence), similar to their evaluations of human companions [72]. We extend previous research [30] with a comparison of human and robotic coaches, with the prediction that

H2 There is a difference in the (a) perceived warmth and (b) perceived competence of human versus robotic coaches.

Moreover, in line with calls for further research [35, 37], we explore the effects of automated social presence versus the presence of human coaches on users' emotional and cognitive reactions, which may influence their acceptance of the technologies [73]. Being in the company of a human or a robot may determine whether elderly participants find playing the exergame enjoyable and effective. We therefore postulate:

H3 There is a difference in the (a) hedonic and (b) utilitarian value perceptions of an exergame, depending on whether the coach is a human or a robot.

Finally, assuming users evaluate both human and robotic entities using universal social cognition dimensions, we further consider how differences in people's social cognition–based evaluations of human versus robot coaches affect their emotional and cognitive responses and, ultimately, their intentions to use the exergame. Therefore:

H4 (a) Perceived warmth and (b) perceived competence of the coach affects users' (i) hedonic and (ii) utilitarian value perceptions of the exergame experience.

H5 (a) Hedonic and (b) utilitarian value perceptions affect users' intentions to use the exergame.

5 Methods and Materials

5.1 Robot Description

The robotic platform in these experiments is the Vizzy robot [51], developed by the Institute for Systems and Robotics (ISR-Lisboa/ISR). Vizzy is a 1.3-m-tall, wheeled robot with a humanoid upper torso and a friendly, marsupial-like design (Fig. 3). It has a total of 30 mechanical degrees of freedom (DOF). The two DOFs of the differential drive base allow it to navigate planar surfaces easily and plan its trajectories using well-established algorithms [59, 64]. The head

Fig. 3 The Vizzy robot

(a)

(b)

Fig. 4 Dialogue control GUI. **a** Default view with several buttons grouping verbal intentions. **b** One of button pressed, showing several options

can perform pan and tilt movements, and its eyes can do tilt, vergence, and version movements, for a total of five DOFs. Vizzy's arms and torso also have 23 DOFs, such that it can perform human-like motions. Two laser scanners at the bottom of the mobile base capture planar point-clouds of the environment, providing valuable information for the robot's safe navigation. Each of Vizzy's eyes contains a camera, used for object/people detection and tracking. A depth camera is mounted on Vizzy's chest, further enhancing the robot's sensing capabilities for human movement analysis. Other components include a loudspeaker and a microphone, useful for human–robot interactions, and 12 tactile sensors on each hand [55], mounted on the finger phalanges.

Vizzy's actuators and controllers allow it to navigate, grasp objects, perform gestures, and gaze in a biologically inspired way [1, 63]. The robot also has software that enables it to detect and follow people in the environment, using an implementation of the Aggregate Channel Features Detector [22] with appearance-based tracking using color features [29]. A set of RVIZ plug-ins allow the remote control of the robot (head and base) in Wizard-of-Oz (WoZ) experiments, in a gaming-like way, while visualizing the obstacles gathered by the laser and the people detected by the cameras. A web interface (see Fig. 4) allows a "wizard" to select utterances for the robot to speak while listening through the robot's microphone. This approach supports faster replies and interactions, compared with explicitly typing utterances, though at the cost of speech flexibility.
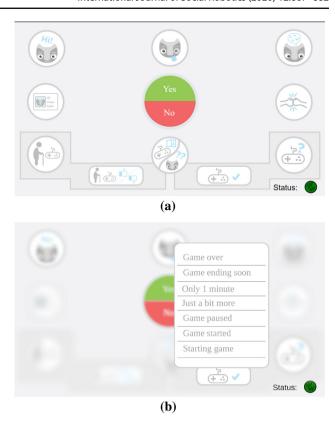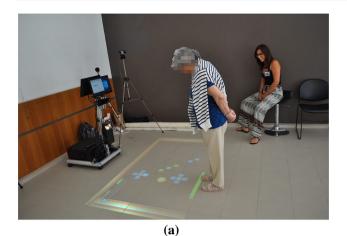
5.2 Gaming Platform Description

To perform the exergames with seniors, we used the Portable Exergames Platform for Elderly (PEPE) [70], an augmented reality gaming platform that projects exergames on the floor, while a Kinect sensor captures the person's movements to control the game elements. Users can play through PEPE without requiring any wearable sensor or controller, which minimizes the burden and complexity for older adults. Two touchscreens on the top of the platform provide additional information to staff members monitoring the exercise. The games currently need to be initiated and terminated by a person, but in the future, the social robot will be able to control the gameflow. The chosen game for our experiments is Exer-Pong [16], which requires players to control a green paddle (Fig. 5a) with their body movements. The objective is to hit the yellow ball with the green paddle, making sure it does not leave the game area. Colored boxes populate the gaming area. When touched by the yellow ball, the colored boxes are destroyed and yield points to the player. After destroying all the boxes, the player completes the level. Every time the player fails to hit the yellow ball, and it leaves the playing area, a box reappears. During the game, audiovisual stimuli give performance information to the player, in the form of a red visual feedback and success and failure sounds. The
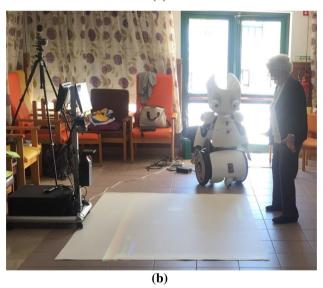
(a)

(b)

Fig. 5 Experimental conditions. **a** Scenario 1, human coaches elderly user in playing the exergame using PEPE. **b** Scenario 2, Vizzy performs the coaching role

players can either control the paddle by walking sideways or with horizontal arm movements. Thus ExerPong can be played by people with different degrees of mobility.

5.3 Participants

In total, 58 elderly persons (42 female, 16 male) participated in the study; they ranged in age between 48 and 94 years ($\mu = 78.79$, $\sigma = 9.85$). We excluded people suffering from severe physical (e.g., full physical immobility) or mental (e.g., dementia) health problems, as well those incapable of giving their consent. The target population comprised elderly persons living autonomously (e.g., in their own home) or in a nursing home who accept services from a care institution. Thus, the main inclusion criterion was that they were clients of the elderly care centers with which we partnered, which offer a wide range of tailored activities (e.g., card

games, fitness, handcrafts). We randomly divided the overall sample into two experimental groups: 22 participants in the human coach group (16 female, 6 male), aged between 65 and 94 years ($\mu = 75.46$, $\sigma = 12.79$), and 36 participants in the robotic coach group (26 female, 10 male), aged between 48 and 91 years ($\mu = 80.83$, $\sigma = 6.98$). The human coach experiments were carried out at two locations (Location A: 10, Location B: 12 elderly participants), and the robotic coach experiments were carried out at three locations (Location C: 11, Location D: 10, Location E: 15 elderly participants).

5.4 Experimental Setup

The data collection, in collaboration with five elderly care institutions, took place during July–September 2017. Each participating elderly care institution provided a spacious activity area, an adjoining room for surveys/interviews, and an in-house psychologist. The activity area was essential, because we needed to fit the gaming platform and secure enough space for the robot to navigate throughout the area. We requested an additional room to ensure each informant's privacy during the data collection. Finally, the in-house psychologist helped screen elderly clients, to evaluate their fit for the experiment (exclusion criteria: severe physical or mental health problems). The visits were scheduled during the regular activities time held by each care center.

5.5 Experimental Procedure

As a part of the study, the exergame platform was installed in the main activities room of each care institution. During the regular activities time, when people usually participate in various activities offered by the institution (e.g., drawing, knitting classes, yoga, card or board games), the seniors were approached either by a human or the robot, depending on their assigned experimental group, and invited to try a new activity that reportedly was being considered as an addition to the existing activity offers. Because all the activities took place in a main activities room, peers could observe the exergame activity, though some remained engaged in their own activities. To keep the experimental scenarios as constant as possible, we developed a standardized set of steps for each actor (i.e., human/robot) to follow: (1) Approach an elderly person who satisfies the inclusion criteria; (2) introduce the new activity (i.e., exergames) and invite the elderly person to join the game; (3) if the elderly person accepts, escort her or him to the gaming area; (4) provide instructions for how to play the game; (5) motivate the elderly person by words of encouragement and feedback on game progression; and (6) ask the elderly person whether to continue or terminate the game. The locations of the coach during Steps 4–6 were standardized, as depicted in Fig. 5. Although the "wizards" directing the robots were present in

the same room, they were described as technicians, available resolve any problems with the exergame system. Furthermore, they were standing in a non-visible area, out of the field of view of the participant during the game. The communication was standardized; both human and robot actors used the same sentence to introduce the exergames, give instructions, and motivate players. Past studies have shown that gendered voices can influence perceived trust, likeability [48], and competence [18] if the agent is a computer with no gender-specific characteristics other than the voice. People tend to give higher ratings to agents whose voice gender matches their own, and they also apply gender stereotypes. Therefore, Vizzy had a female voice to match the gender of the caregivers, who were all women. On average, the time participants spent playing the exergame was 08:07 min, with a standard deviation of 02:32 min. The minimum playing time was 1:58 min, and the maximum was 12:25 min. After the game, a member of the research team escorted the elderly participants to the separate room to conduct the survey and interview. Prior to being surveyed, all elderly informants signed an informed consent, outlining the main objectives of the research project and guaranteeing the anonymity of their responses. We consciously decided not to ask for informed consent before their interaction with Vizzy or the exergame platform, because we wanted to mimic a regular activity at the center, rather than create a research environment.
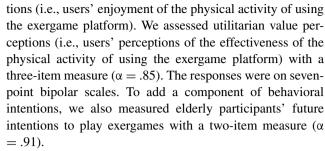
5.6 Data Collection

To test whether a robot (i.e., automated social presence) is comparable to humans in terms of motivating elderly people to engage in physical activity, we chose a mixed-method approach, combining quantitative and qualitative data. The main instrument was a questionnaire, augmented with probing questions to elicit more in-depth, qualitative insights.

5.6.1 Quantitative Data Collection

The 58 elderly participants were asked to allocate 10–15 min of their time and join the first author to discuss their overall experience with the activity. For that purpose, a questionnaire was constructed to guide the discussion. With the decrease of their cognitive functioning, elderly people often struggle with question and rating scales comprehension [68], so we decided to conduct guided surveys. The researcher read the question to the elderly person and explained the rating scale (e.g., totally disagree to totally agree), then asked the elderly person to indicate her or his response.

To measure the experience with playing exergames (i.e., emotional and cognitive reactions), we adopted two bipolar scales [17, 52]. Specifically, we employed a five-item measure ($\alpha = .85$) to assess their hedonic value percep-

tions (i.e., users' enjoyment of the physical activity of using the exergame platform). We assessed utilitarian value perceptions (i.e., users' perceptions of the effectiveness of the physical activity of using the exergame platform) with a three-item measure ($\alpha = .85$). The responses were on seven-point bipolar scales. To add a component of behavioral intentions, we also measured elderly participants' future intentions to play exergames with a two-item measure ($\alpha = .91$).

Furthermore, we gathered the elderly participants' evaluations of whether the human/robotic coach was good- or ill-intended (i.e., warmth) and competent to perform the task. In line with prior conceptualizations of social cognition [20], we adopted a three-item measure for perceived warmth ($\alpha = .76$) and three-item measure for perceived competence ($\alpha = .88$). Participants responded to all items on five-point Likert scales ("totally disagree" = 1 to "totally agree" = 5).

Finally, to evaluate the experience of sensing a social entity when interacting with the robot, we developed a four-item measure ($\alpha = .7$) of automated social presence [2, 37]. Again, all items were measured on five-point Likert scales. Tables 1 and 2 contain the constructs and items used in the questionnaire.

5.6.2 Qualitative Data Collection

We chose to employ a convergent parallel mixed methods design [19], in which we collected both quantitative and qualitative data at the same time, with the goal of determining whether the findings associated with both types of data support each other. During the survey process, the researcher probed particular items as appropriate, using either "tell me more" probes [4] or a laddering technique [61]. For example, after elderly participants indicated their enjoyment level (from "I hated it" = 1 to "I enjoyed it" = 7), the researcher would ask what in particular they found enjoyable or not. After they indicated their intentions to use the exergame platform again (scale from 1 to 5), the researcher would ask why and thereby potentially uncover the core values that guide their behavioral intentions.

All collected narratives were digitally audio-recorded and then transcribed, translated, and reviewed. The analysis proceeded in the following order: First, the authors all read the transcripts independently to form their own understanding of each participant's narratives [62]. Second, four co-authors met for a joint analysis session to share emerging codes and develop a more focused coding scheme. Third, the first two authors coded the remaining narratives according to the coding scheme. Due to the exploratory nature of the study, codes and thematic descriptions emerged from the data rather than from preconceived categories or theories [40].

Table 1 Items related to the exergame

Hedonic value ($\alpha = .85$)		
1: I hated it (1)	...	I enjoyed it (7)
2: I felt bored (1)	...	I felt interested (7)
3: I disliked it (1)	...	I liked it (7)
4: I found it unpleasurable (1)	...	I found it pleasurable (7)
5: It was not fun at all (1)	...	It was a lot of fun (7)
Utilitarian value ($\alpha = .85$)		
1: I find it harmful for my physical health (1)	...	I find it useful for my physical health (7)
2: I find exergames ineffective (1)	...	I find exergames effective (7)
3: I feel like it cannot improve my physical health (1)	...	I feel like it can improve my physical health (7)
Intention to use (1–5 Likert scale, $\alpha = .91$)		
1: I think I will play exergames during the next few days		
2: I will use the exergame platform in the future		

Table 2 Items related to the human/robotic coach

Perceived warmth (1–5 Likert scale, $\alpha = .76$)
1: I feel the human/robotic coach understands me
2: I think the human/robotic coach is well-intentioned
3: I think the human/robotic coach is friendly
Perceived competence (1–5 Likert scale, $\alpha = .88$)
1: I think the human/robotic coach is competent
2: I think the human/robotic coach is reliable
3: I think the human/robotic coach is an expert/knowledgeable
Automated social presence (1–5 Likert scale, $\alpha = .70$)
1: I can image the robot to be a living creature
2: When interacting with the robot I felt I'm talking with a real person
3: Sometimes the robot seems to have real feelings
4: I felt like the robot was actually looking at me throughout the interaction

6 Results

6.1 Automated Social Presence

To investigate H1, we ran a one-sample *t* test to determine whether the mean value of automated social presence is statistically different from a neutral value, that is, the midpoint (= 3) on the 5-point Likert scale. In support of H1, the mean score of 3.5 (SD = .97) is significantly different from 3, t(35) = 2.97, $p = .005$ (Fig. 6). That is, elderly people experience some sort of automated social presence when interacting with the robot. When we complement these quantitative findings with the qualitative data, we identify the emergence of a thematic paradox that we label "We know it's a machine, but it feels like a human." From the narratives, we recognize that the elderly participants experience some conflicting thoughts, as captured in the following quotes:

> It's almost the same thing [as talking with a real person]. (Female, 80)
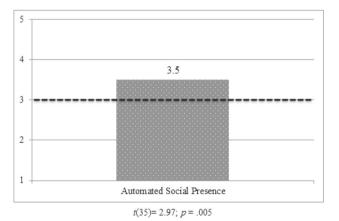

$t(35) = 2.97; p = .005$

Fig. 6 Automated social presence

> No, I mean… we can see that it's a robot, right? But we can see that it's something intelligent… but it's not a person… but it's as if it is a person. (Female, 89)

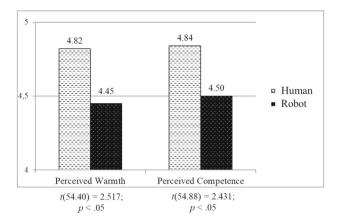> It's a robot but… I imagined it as a person there talking to me. (Female, 94)

Fig. 7 Perceptions of warmth and competence

It felt like talking to an adult person. It's like the grown-ups. (Male, 71)

The comparison of quantitative and qualitative data thus reveals similarities and support for H1.

6.2 Social Cognition

To test H2, we ran an independent-samples t-test and determine if the means of two groups (human vs. robot) differ statistically. We find a significant difference in the mean scores for perceived warmth when the coach is a human (M = 4.82, SD = .37) or a robot (M = 4.45, SD = .73), t(54.40) = 2.517, p <.05. Similar results emerge for the perceived competence of the coach, as a human (M = 4.84, SD = .47) or robot (M = 4.50, SD = .67), t(54.88) = 2.431, p <.05 (Fig. 7).

These results support H2, the difference for both perceived warmth and perceived competence is statistically significant. Elderly people perceive the robot to be friendly, well-intentioned, and understanding of their needs, but to a lesser extent than the human coach. Perceiving the robot as high on the warmth dimension is in itself an important finding, which we also can affirm with the qualitative data. In coding the narratives, we identified another, broader theme, which we label "It's a machine, but it's kind and gentle." As emphasized by the participants:

The robot spoke very politely to us [laughs]. He is really friendly. (Female, 86)

The robot seems gentle to me. (Female, 64)

If he has revealed all of his intentions… they are good. (Male, 72)

I was very loyal to the robot and it became my friend. (Male, 90)

Similarly, the elderly participants perceive the robot as competent, reliable, and a knowledgeable expert for the par-
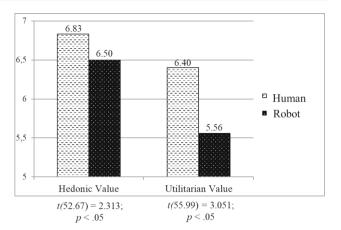

Fig. 8 Hedonic and utilitarian value perceptions of exergames

ticular task at hand. They assign it an average of 4.5 on a 5-point scale, slightly lower than the human coach. The qualitative data yield similar results, in further empirical support of the high perceived competence of the robotic coach, as in the following quotes:

For what he is programmed, he is competent. He seemed smart to me. He did not say that much. If he had spoken about football… [laughs] (Male, 72)

Wasn't it intelligent? It is more intelligent than I am. (Female, 84)

It is well-informed. If it wasn't well informed it wouldn't come here. (Female, 80)

It's smart to the max! (Male, 81)

6.3 Hedonic and Utilitarian Value

To investigate H3, we ran another independent-samples t-test, to determine whether the different types of coaches (human vs. robot) affect elderly participants' value perceptions regarding their enjoyment and the effectiveness of exercising using the exergame platform. As hypothesized, we find a significant difference in the mean scores, for both hedonic and utilitarian value perceptions, across the experimental groups. In particular, the mean for hedonic value when the coach is a human is 6.83 (SD = .33), whereas when it is a robot, the mean is 6.50 (SD = .73), t(52.67) = 2.313, p <.05. For utilitarian value, the means again are higher for a human coach (M = 6.40, SD = .78) compared with a robotic one (M = 5.56, SD = 1.31), t(55.99) = 3.051, p <.05 (Fig. 8).

These results support H3 but do not explain the source of these differences. When we integrate the qualitative data, we realize that the elderly people had fun and enjoyed the experience with the exergames while in the company of a robot:

Table 3 Descriptive statistics and correlations

Correlations	1	2	3	4	5	6
1. Coach (Human vs. Robot)	1					
2. Perceived warmth	− .28*	1				
3. Perceived competence	− .29*	.61**	1			
4. Hedonic value	− .25	.56**	.63**	1		
5. Utilitarian value	− .34**	.42**	.59**	.54**	1	
6. Intention to use	− .13	.43**	.55**	.56**	.53**	1
Descriptive statistics						
Minimum	0	3	3	4	2	1
Maximum	1	5	5	7	7	5
Mean	.62	4.59	4.64	6.62	5.87	4.52
SD	.49	.64	.62	.63	1.20	.85

I loved it. I found it very interesting… and believe me, I don't participate if I don't enjoy the activity…. I had the maximum score. It was very funny. And the robot was a good lad. (Male, 72)

I liked it a lot, lot, lot, lot, lot! The robot was always saying:'Excellent, excellent'. That motivated me. (Female, 84)

Yet they enjoyed the experience slightly less than they did with a human coach, as indicated by their narratives:

I was a bit shy and anxious, right? Without knowing what to say, or what to do… the movements I should do. (Female, 89)

I found it funny [enjoyable]. But I almost can't hear, so I couldn't understand most things that the robot said. (Female, 91)

Similarly, when asked to share their narratives of whether they found the exergames effective (i.e., utilitarian value perceptions) for improving their physical condition, some participants shared concerns:

Robot's explanations were not enough! I was there doing something that I didn't know if it was right or wrong. And if I really was… as the robot was saying "Excellent", but those are the words it has inside and we cannot trust that very much. (Female, 81)

Let's see… if the robot only does that exercise, that isn't much more than moving left and right, I think it's insufficient. So it should have more, more advantages… (Female, 81)

6.4 Regressions

Table 3 contains an overview of the descriptive statistics and correlations. To test H4 and H5 (Fig. 2) we estimate a sys-tem of five linear equations (see Table 4 for an overview of the dependent and independent variables in each equation), using the seemingly unrelated regressions (SUR) procedure in STATA. The SUR method is based on generalized least squares and is more efficient than ordinary least squares regression when the independent variables differ across a system of equations, because it accounts for correlated errors [78]. The significant Breusch–Pagan–Lagrange multiplier test for error independence ($\chi^2(10) = 28.306$, $p < .01$) indicates correlated errors in the five equations, and the R-square for each individual equation is statistically significant at $p < .01$ (see Table 5. Therefore, using SUR is appropriate [15, 23]. The first two equations have perceived warmth and perceived competence as dependent variables and the type of coach as the independent variable. These two equations control for the influence of the coach on social cognition. The third equation includes hedonic value as the dependent variable and perceived warmth and competence as independent variables; the fourth equation features utilitarian value as the dependent variable and perceived warmth and competence as independent variables. These two equations test H4. Finally, the fifth equation includes intention to use as the dependent variable and hedonic and utilitarian value as the independent variables, to test H5.

The unstandardized coefficients and their standard errors, obtained with the SUR estimator (sureg command) in STATA 14.2, are in Table 5. The perceived warmth of the coach has a positive, significant coefficient for hedonic value (.29; $p < .05$), in support of H4a.i; its impact on utilitarian value does not produce a significant coefficient though (.23; $p > .05$), so we cannot confirm H4a.ii. The perceived competence of the coach indicates a positive, significant coefficient for both hedonic (.46; $p < .01$) and utilitarian (1.08; $p < .01$) value, in support of both elements of H4b. Finally, hedonic value has a positive, significant coefficient for intention to use (.52; $p < .01$), as does utilitarian value (.22; $p < .05$). Therefore, we find support for H5a and H5b.

Table 4 Overview of Seemingly Unrelated Regressions (SUR) Equations

Equation	Hypothesis	Independent variable	Dependent variable	RMSE	R^2	χ^2	p
(1)	NA	Coach (Human vs. Robot)	Perceived warmth	.61	.08	5.64	.02
(2)	NA	Coach (Human vs. Robot)	Perceived competence	.59	.08	6.21	.01
(3)	H4.i	Perceived warmth Perceived competence	Hedonic value	.47	.44	48.11	.00
(4)	H4.ii	Perceived warmth Perceived competence	Utilitarian value	.96	.35	37.09	.00
(5)	H5	Hedonic value Utilitarian value	Intention to use	.66	.39	36.16	.00

Sample size = 58; NA = not applicable

Table 5 Overview of Seemingly Unrelated Regressions (SUR) Results

	(1) Perceived warmth		(2) Perceived competence	
	Coefficient	Standard error	Coefficient	Standard error
Intercept	4.83**	.13	4.89**	.12
Coach (Human vs. Robot)	− .39*	.16	− .39*	.16
	(3) Hedonic value		(4) Utilitarian value	
	Coefficient	Standard error	Coefficient	Standard error
Intercept	3.15**	.51	− .20	1.04
Perceived warmth	.29*	.12	.23	.25
Perceived competence	.46**	.12	1.08**	.26
	(5) Intention to use			
	Coefficient	Standard error		
Intercept	− .19	.91		
Hedonic value	.52**	.16		
Utilitarian value	.22*	.08		

Sample size: 58; Breusch–Pagan–Lagrange: $\chi^2(10) = 28.306$, $p < .01$

Number in parentheses above each dependent variable refer to the equations in Table 4

$*p < .05; **p < .01$

7 Discussion

This research evaluates the application of robots (i.e., SARs) in social contexts, formerly exclusively reserved for human agents. Motivated by the importance of physical activity for healthy aging, and the shortages of elderly care staff [54], our aim is to examine if SARs can take a coaching role and motivate elderly people to undertake physical activity, as well as potentially expand their social networks. In particular, we test whether SARs, which are not inherently social entities but are programmed to exhibit social behaviors, can give rise to social perceptions in elderly people and affect them in ways similar to those induced by human agents [36]. With the idea that SARs equipped with social capabilities are comparable to human agents, we also seek to investigate their efficiency in motivating elderly people to engage in a new kind of physical activity (i.e., exergames), relative to the impact of human coaches.

7.1 Automated Social Presence and Social Cognition

Our findings suggest that elderly people sense that they are in the company of social entities when they interact with SARs [72]. Our mixed-method empirical tests affirm that they struggle to classify the robot as a "machine," because of its expressiveness, friendly voice, politeness, and considerateness. In their narratives, we detect evidence that they see the robot as a "metal box," but due to the robot's social dexterity, they humanize the machine in their imagination. This evidence implies that SARs potentially could extend elderly people's social networks and support human caregivers in not just mechanical tasks but also tasks of a social nature (e.g., motivational coaches, playing games, conversational partners).

In these quasi-social interactions, in which elderly people perceive SARs as social entities, they also tend to evaluate SARs using the social cognition mechanisms they apply to humans. Our experiments shed light on social perceptions of automated agents [60]. Elderly people activate their

warmth and competence judgments when interacting with SARs. However, when we compare human and robot actors, performing the same (coaching) task, we find that they evaluate humans as superior to the robot. We explain this slight difference with qualitative data, which reveal that the human coaches are better at responding to elderly persons' individual needs. Even though they find Vizzy cute and funny, some elderly participants complained that they could barely hear what it was saying. Unlike Vizzy, for which the voice pace and pitch remained the same, and the volume could not be adjusted (i.e., it was already set at its maximum), human coaches could adjust their voices to the needs of hearing-impaired participants. This difference likely affected both warmth and competence judgments.

7.2 Hedonic and Utilitarian Values

By conducting these experiments in a real-life context, we could test whether elderly people enjoy and appreciate this new type of gamified workout, depending on their coaching partner (human vs. robot). We demonstrate that seniors' hedonic and utilitarian value perceptions of exergames are driven by the social cognition they develop of their coaches. Elderly participants both enjoy and find the exergame more effective if they are instructed and motivated by a human, rather than a robot. The reason for this difference might be the slightly higher levels of anxiety that we find expressed in the narratives, due to interacting with the robot. For many of the elderly participants, this experience was the first time they had seen or engaged with an automated agent. Some noted their nervousness when the robot first approached them ("Initially, I wasn't very relaxed, but I ended up being more relaxed as we continued," Male, 86). Some seniors also lacked the information they needed to enjoy playing exergames ("I did not quite understand the purpose of the game, because I could not hear well what the robot was saying," Female, 87). These quotes provide insights into the participants' psychological states and physical (e.g., hearing) impairments, which hinder their full enjoyment of the exergames. These important inputs for technology developers and service designers can help them design robotic platforms to ensure the best experience for senior users.

Our analysis also demonstrates that perceived competence has a positive impact on both hedonic and utilitarian value perceptions. The effect of competence on hedonic value is even stronger than that of perceived warmth. We explain the effect by taking the exercise context into account. That is, the activity the participants were evaluating involved physical exercise, so they were particularly interested in the reliability and ability of the coach to react flexibly. Thus, it makes sense that competence judgments exhibit primacy over warmth judgments. Furthermore, for those with the robotic coach, the participants were being exposed to a new and unfamiliar

agent, which increased their need to trust in its abilities before being able to enjoy the activity. We argue that human coaches might seem to perform better because they are able, if necessary, to demonstrate the movement. That is, some seniors had difficulty understanding the requirements of the exergame. In such cases, the robot could only repeat the instructions, in the hope that the participant would understand what was expected; human coaches could augment their voice instructions with a movement demonstration, which likely improved the evaluations of their competence. Despite the robots' affective capabilities (e.g., kindness, care, emotional sensitivity), which enhance warmth judgments, their lack of competence might hinder the adoption of robotic coaches.

7.3 Contributions and Implications

With these findings, we primarily contribute to the theory of social presence and explain cognitive mechanisms activated by interacting with artificial minds [8]. We advance current knowledge on the ways SARs, through their social resemblance to humans, allow people interacting with them to feel as if they are in a presence of social entities [72]. In turn, by demonstrating that elderly people sense a social presence when interacting with SARs and develop social perceptions about them, we offer insights for further developments and designs of social robotics and systems. Including robots in people's social networks, as exercise coaches, could improve both physical and psychosocial well-being. However, in contexts that require high social engagement, humans still perform better than SARs. Social robots need continued improvements, in terms of their empathy, flexibility, and spontaneity, to adjust their behavior to the heterogeneity of human needs.

Still, SARs can be used to enhance elderly people's experience and assist human caregivers in certain tasks. Assisted living facility managers should make well-considered decisions about which tasks to automate and which to leave in human hands. Reflecting on our coaching tasks, SARs can be used to motivate seniors to be more active, but human caregivers should make sure the elderly users perform the exercises correctly and assist them if necessary.

7.4 Limitations and Further Research

This study provides an interesting and necessary perspective on ways in which SARs can support human caregivers in motivating elderly people to engage in physical activity, but it also contains some limitations. First, none of the elderly care institutions we collaborated with houses as many physically active clients as would be ideal for our experiments. Therefore, we conducted the study across several locations. Second, the participating care institutions, though located in the same region in the same country, feature heterogeneous

elderly populations and varying conditions for the field experiments. Additional research might find other ways to expand the sample but also to minimize the differences across elderly populations and care institutions.

Third, the short-term nature of the study (i.e., single session) introduces a novelty effect, for both the robot and exergames. To rule out the effect of novelty, a longitudinal study might measure the same constructs in different phases of the experiment. Researchers then could track not only behavioral intentions but also actual behavior (i.e., whether elderly people keep playing the exergames in the future). Fourth, the robot's speech abilities were controlled with a WoZ method, which limits the spontaneity of dialog. Continued research could provide for more flexibility, moving beyond preset utterances to allow for more natural social interactions (e.g., typing responses that exactly address the questions/statements raised by the elderly participants), at least until we reach the point that the human granting agency to the robot can be eliminated and the robot can interact completely autonomously. Although the volume of the robot's speech was already at its maximum, it seemed insufficient for some elderly participants. Robot developers thus should find ways to enhance robots' volume and produce sounds on a range of frequencies that older adults can hear more easily [56].

Fifth, in addition to speech, we controlled the robot's navigation and gaze, so our experiment does not involve fully autonomous robot behavior. Efforts to endow the robot with autonomous navigation, perception, speech, and decision making could reduce the degree of human intervention needed to ensure appropriate functioning. However, this level of autonomous behavior is rare and difficult to achieve, considering the currently limited capabilities of social robots [12]. Sixth, another limitation is the robot's inability to mimic humans' nonverbal communication, such as facial expressions and gestures. It would be interesting to investigate how robots endowed with nonverbal capabilities, in addition to verbal ones, affect users' perceptions of their warmth and competence. We expect that facial expressiveness (e.g., smile, eye contact) influences warmth judgments, and gestures (e.g., use of the limbs to demonstrate the exercise) might affect competence judgments.

References

1. Avelino J, Figueiredo R, Moreno P, Bernardino A (2016) On the perceptual advantages of visual suppression mechanisms for dynamic robot systems. Procedia Comput Sci 88:505–511
2. Bailenson JN, Blascovich J, Beall AC, Loomis JM (2001) Equilibrium theory revisited: mutual gaze and personal space in virtual environments. Presence: Teleoperators Virtual Environ 10(6):583–598
3. Bauman A, Merom D, Bull FC, Buchner DM, Fiatarone Singh MA (2016) Updating the evidence for physical activity: Summative reviews of the epidemiological evidence, prevalence, and interventions to promote "active aging". Gerontology 56(2):S268–S280
4. Bernard HR, Bernard HR (2012) Social research methods: qualitative and quantitative approaches. Sage, Thousand Oaks
5. Bernstein MS, Morabia A, Sloutskis D (1999) Definition and prevalence of sedentarism in an urban population. Am J Public Health 89(6):862–867
6. Biocca F, Harms C (2002) Defining and measuring social presence: contribution to the networked minds theory and measure. Proc Presence 2002:7–36
7. Biocca F, Nowak K (2001) Plugging your body into the telecommunication system: mediated embodiment, media interfaces, and social virtual environments. In: Lin C, Atkin D (eds) Communication technology and society. Hampton Press, Waverly Hill, pp 407–447
8. Biocca F, Harms C, Burgoon JK (2003) Toward a more robust theory and measure of social presence: review and suggested criteria. Presence: Teleoperators Virtual Environ 12(5):456–480
9. Boyle EA, Hainey T, Connolly TM, Gray G, Earp J, Ott M, Lim T, Ninaus M, Ribeiro C, Pereira J (2016) An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. Comput Educ 94(Supplement C):178–192
10. Breazeal C (2003) Toward sociable robots. Robot Auton Syst 42(3–4):167–175
11. Breazeal C (2004) Social interactions in hri: the robot view. IEEE Trans Syst Man Cybern C (Appl Rev) 34(2):181–186
12. Broadbent E (2017) Interactions with robots: the truths we reveal about ourselves. Annu Rev Psychol 68:627–652
13. Broekens J, Heerink M, Rosendal H et al (2009) Assistive social robots in elderly care: a review. Gerontechnology 8(2):94–103
14. Čaić M, Odekerken-Schröder G, Mahr D (2018) Service robots: value co-creation and co-destruction in elderly care networks. J Serv Manag 29(2):178–205
15. Cameron AC, Trivedi PK (2009) Microeconometrics with stata. StataCorp LP, College Station
16. Cardona JEM, Cameirao MS, PaulinoT, iBadia SB, Rubio E (2016) Modulation of physiological responses and activity levels during exergame experiences. In: 2016 8th international conference on games and virtual worlds for serious applications (VS-GAMES), pp 1–8
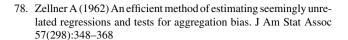
17. Childers TL, Carr CL, Peck J, Carson S (2002) Hedonic and utilitarian motivations for online retail shopping behavior. J Retail 77(4):511–535

18. Clifford N, Youngme M, Nancy G (1997) Are machines gender neutral? Gender-stereotypic responses to computers with voices. J Appl Soc Psychol 27(10):864–876

19. Creswell JW, Creswell JD (2017) Research design: qualitative, quantitative, and mixed methods approaches. Sage, Thousand Oaks

20. Cuddy AJ, Fiske ST, Glick P (2008) Warmth and competence as universal dimensions of social perception: the stereotype content model and the bias map. Adv Exp Soc Psychol 40:61–149

21. Dautenhahn K, Woods S, Kaouri C, Walters ML, Koay KL, Werry (2005) What is a robot companion-friend, assistant or butler? In: 2005 IEEE/RSJ international conference on intelligent robots and systems, 2005 (IROS 2005). IEEE, pp 1192–1197

22. de Graaf MM, Allouch SB, Klamer T (2015) Sharing a life with Harvey: exploring the acceptance of and relationship-building with a social robot. Comput Hum Behav 43:1–14

23. Dollár P, Appel R, Belongie S, Perona P (2014) Fast feature pyramids for object detection. IEEE Trans Pattern Anal Mach Intell 36(8):1532–1545

24. Dotzel T, Shankar V, Berry LL (2013) Service innovativeness and firm value. J Mark Res 50(2):259–276

25. Fan J, Bian D, Zheng Z, Beuscher L, Newhouse PA, Mion LC, Sarkar N (2017) Arobotic coach architecture for elder care (rocare) based on multi-user engagement models. IEEE Trans Neural Syst Rehabil Eng 25(8):1153–1163

26. Fasola J, Mataric MJ (2012) Using socially assistive human 2013; robot interaction to motivate physical exercise for older adults. Proc IEEE 100(8):2512–2526

27. Feil-Seifer D, Mataric MJ (2005) Defining socially assistive robotics. In: 9th international conference on rehabilitation robotics, 2005. ICORR 2005. IEEE, pp 465–468

28. Feil-Seifer D, Skinner K, Matarić MJ (2007) Benchmarks for evaluating socially assistive robotics. Interact Stud 8(3):423–439

29. Fernández-Llamas C, Conde MÁ, Rodríguez-Sedano FJ, Rodríguez-Lera FJ, Matellán-Olivera V (2017) Analysing the computational competences acquired by K-12 students when lectured by robotic and human teachers. Int J of Soc Robot. https://doi.org/10.1007/s12369-017-0440-9

30. Figueira D, Bazzani L, Minh HQ, Cristani M, Bernardino A, Murino V (2013) Semi-supervised multi-feature learning for person re-identification. In: 2013 10th IEEE international conference on advanced video and signal based surveillance, pp 111–116

31. Fiore SM, Wiltshire TJ, Lobato EJ, Jentsch FG, Huang WH, Axelrod B (2013) Toward understanding social cues and signals in human–robot interaction: effects of robot gaze and proxemic behavior. Front Psychol 4:1–15

32. Fiske ST, CuddyAJ Glick P (2007) Universal dimensions of social cognition: warmth and competence. Trends Cogn Sci 11(2):77–83

33. Fong T, Nourbakhsh I, Dautenhahn K (2003) A survey of socially interactive robots. Robot Auton Syst 42(3):143–166

34. Görer B, Salah AA, Akın HL (2017) An autonomous robotic exercise tutor for elderly people. Auton Robots 41(3):657–678

35. Heerink M, Kröse B, Wielinga B, Evers V (2008) Enjoyment intention to use and actual use of a conversational robot by elderly people. In: Proceedings of the 3rd ACM/IEEE international conference on human robot interaction. ACM, New York, NY, USA, HRI'08, pp 113–120

36. Heerink M, Kröse B, Evers V, Wielinga B (2010) Assessing acceptance of assistive social agent technology by older adults: the almere model. Int J Soc Robot 2(4):361–375

37. Heerink M, Kröse B, Evers V, Wielinga B (2010) Relating conversational expressiveness to social presence and acceptance of an assistive social robot. Virtual Real 14(1):77–84

38. Heeter C (1992) Being there: the subjective experience of presence. Presence: Teleoperators Virtual Environ 1(2):262–271

39. Hoorn JF, Winter SD (2017) Here comes the bad news: doctor robot taking over. Int J Soc Robot 10(4):519–535

40. Hsieh HF, Shannon SE (2005) Three approaches to qualitative content analysis. Qual Health Res 15(9):1277–1288

41. Iwamura Y, Shiomi M, Kanda T, Ishiguro H, Hagita N (2011) Do elderly people prefer a conversational humanoid as a shopping assistant partner in supermarkets? In: 2011 6thACM/IEEE international conference on human–robot interaction (HRI), pp 449–457

42. Janssen D, Jongen W, Schrder-Bck P (2016) Exploring the impact of austerity-driven policy reforms on the quality of the long-term care provision for older people in Belgium and the Netherlands. J Aging Stud 38:92–104

43. Jerčić P, Wen W, Hagelbäck J, Sundstedt V (2017) The effect of emotions and social behavior on performance in a collaborative serious game between humans and autonomous robots. Int J Soc Robot 10:115–129

44. Jerčić P, Wen W, Hagelbäck J, Sundstedt V (2018) The effect of emotions and social behavior on performance in a collaborative serious game between humans and autonomous robots. Int J Soc Robot 10(1):115–129

45. Kennedy J, Baxter P, Senft E, Belpaeme T (2016) Heart vs. hard drive: children learn more from a human tutor than a social robot. In: 2016 11th ACM/IEEE international conference on human–robot interaction (HRI), pp 451–452

46. Khosla R, Nguyen K, Chu MT (2017) Human robot engagement and acceptability in residential aged care. Int J Hum-Comput Interact 33(6):510–522

47. Kim M, Kwon T, Kim K (2018) Can human–robot interaction promote the same depth of social information processing as human–human interaction? Int J Soc Robot 10(1):33–42

48. Lee EJ, Nass C, Brave S (2000) Can computer-generated speech have gender? An experimental test of gender stereotype. In: CHI'00 extended abstracts on human factors in computing systems. ACM, New York, NY, USA, CHI EA'00, pp 289–290. https://doi.org/10.1145/633292.633461

49. Li J (2015) The benefit of being physically present: a survey of experimental works comparing copresent robots, telepresent robots and virtual agents. Int J Hum-Comput Stud 77:23–37

50. McColl D, Nejat G (2013) Meal-time with a socially assistive robot and older adults at a long-term care facility. J Hum-Robot Interact 2(1):152–171

51. Moreno P, Nunes R, Figueiredo R, Ferreira R, Bernardino A, Santos-Victor J, Beira R, Vargas L, Aragão D, Aragão M (2016) Vizzy: a humanoid on wheels for assistive robotics. In: Robot 2015: second Iberian robotics conference. Springer, New York, pp 17–28

52. Mullen SP, Olson EA, Phillips SM, Szabo AN, Wójcicki TR, Mailey EL, Gothe NP, Fanning JT, Kramer AF, McAuley E (2011) Measuring enjoyment of physical activity in older adults: invariance of the physical activity enjoyment scale (paces) across groups and time. Int J Behav Nutr Phys Activity 8(1):103

53. Nowak KL, Biocca F (2003) The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. Presence: Teleoperators Virtual Environ 12(5):481–494

54. Organization WH (2015) World report on ageing and health. World Health Organization, Geneva

55. Paulino T, Ribeiro P, Neto M, Cardoso S, Schmitz A, Santos-Victor J, Bernardino A, Jamone L (2017) Low-cost 3-axis soft tactile sensors for the human-friendly robot vizzy. In: 2017 IEEE international conference on robotics and automation (ICRA), pp 966–971

56. Pearson JD, Morrell CH, GordonSalant S, Brant LJ, Metter EJ, Klein LL, Fozard JL (1995) Gender differences in a longitudinal study of age associated hearing loss. J Acoust Soc Am 97(2):1196–1205

57. Portugal D, Santos L, Alvito P, Dias J, Samaras G, Christodoulou E (2015) Socialrobot: an interactive mobile robot for elderly home care. In: 2015 IEEE/SICE international symposium on system integration (SII), pp 811–816

58. Powers A, Kiesler S, Fussell S, Torrey C (2007) Comparing a computer agent with a humanoid robot. In: Proceedings of the ACM/IEEE international conference on human–robot interaction. ACM, New York, NY, USA, HRI'07, pp 145–152

59. Quinlan S, Khatib O (1993) Elastic bands: connecting path planning and control. In: Proceedings of IEEE international conference on robotics and automation, 1993. IEEE, pp 802–807

60. Reeves B, Nass C (1996) How people treat computers, television, and new media like real people and places. Cambridge University Press, Cambridge

61. Reynolds TJ, Gutman J (1988) Laddering theory, method, analysis, and interpretation. J Advert Res 28(1):11–31

62. Riessman CK (1993) Narrative analysis, vol 30. Sage, Thousand Oaks

63. Roncone A, Pattacini U, Metta G, Natale L (2016) A cartesian 6-dof gaze controller for humanoid robots. In: Proceedings of robotics: science and systems. AnnArbor, Michigan

64. Rösmann C, Feiten W, Wösch T, Hoffmann F, Bertram T (2012) Trajectory modification considering dynamic constraints of autonomous robots. In: 7th German conference on robotics; proceedings of ROBOTIK 2012; VDE, pp 1–6

65. Scassellati B, Admoni H, Matarić M (2012) Robots for use in autism research. Annu Rev Biomed Eng 14:275–294

66. Schneider S, Kümmert F (2016) Exercising with a humanoid companion is more effective than exercising alone. In: 2016 IEEE-RAS 16th international conference on humanoid robots (Humanoids), pp 495–501

67. Scholes S, Mindell J (2012) Physical activity in adults. Health Surv Engl 1(2):1–49

68. Schwarz N, ParkD Knauper B, Sudman S (1998) Cognition, aging and self-reports. Psychology Press, London

69. Shinozawa K, Naya F, Yamato J, Kogure K (2005) Differences in effect of robot and screen agent recommendations on human decision-making. Int J Hum-Comput Stud 62(2):267–279

70. Simao H, Bernardino A (2017) User centered design of an augmented reality gaming platform for active aging in elderly institutions. In: 5th International congress on sport sciences research and technology support (icSPORTS 2017)

71. Skjæret N, Nawaz A, Morat T, Schoene D, Helbostad JL, Vereijken B (2016) Exercise and rehabilitation delivered through exergames in older adults: an integrative review of technologies, safety and efficacy. Int J Med Inf 85(1):1–16

72. Van Doorn J, Mende M, Noble SM, Hulland J, Ostrom AL, Grewal D, Petersen JA (2017) Domo arigato Mr. Roboto: emergence of automated social presence in organizational frontlines and customers' service experiences. J Serv Res 20(1):43–58

73. Venkatesh V, Morris MG, Davis GB, Davis FD (2003) User acceptance of information technology: toward a unified view. MIS Q 27(3):425–478

74. Ventura R, Basiri M,Mateus A,Garcia J, Miraldo P, Santos P, Lima P (2016) A domestic assistive robot developed through robotic competitions. In: IJCAI 2016 Workshop on Autonomous Mobile Service Robots, New York, USA

75. Voiklis J, Kim B, Cusimano C, Malle BF (2016) Moral judgments of human vs. robot agents. In: 2016 25th IEEE international symposium on robot and human interactive communication (RO-MAN), pp 775–780

76. Wattanasoontorn V, Boada I, Garca R, Sbert M (2013) Serious games for health. Entertain Comput 4(4):231–2470

77. WHO (2008) Global tuberculosis control: surveillance, planning, financing, vol 393. WHO report 2008. World Health Organization

78. Zellner A (1962) An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. J Am Stat Assoc 57(298):348–368

Martina Čaić is a Post-Doctoral researcher in the Department of Marketing and Supply Chain Management at Maastricht University and Marie Curie Fellow at the Service Design for Innovation Network (SDIN). In addition, she is a visiting researcher at the Fraunhofer Portugal Research Center for Assistive Information and Communication Solutions. Her research addresses user experiences in value networks, with a particular focus on robotic and ambient assisted living technologies. Specifically, she conducts applied research in healthcare, healthy aging contexts, using service design tools and methods to better understand the adoption of network innovations, as well as measuring user experiences with new technology infusions.

João Avelino is a Ph.D. candidate at the Institute for Systems and Robotics at Instituto Superior Técnico, University of Lisbon. He has a background in electrical and computer engineering, majoring in systems, decision, control & robotics. His research focuses on social robotics and specifically on the detection of social cues and the development of socially aware engagement behaviors that take sensor uncertainty into account.

Dominik Mahr is Professor for Digital Innovation and Marketing at Maastricht University. He is also Scientific Director of the Service Science Factory. His research centers on recent marketing and innovation phenomena including open innovation, virtual social communities, online innovation brokers, customer co-creation, service design, and the development of new services. He has published in high impact journals. Prior to his academic career, Dr. Mahr worked for several years in different management and marketing consultancies, operating in industries such as high tech, automotive, real estate, and insurance.

Gaby Odekerken-Schröder is Professor of Customer-Centric Service Science at Maastricht University, the Netherlands. Her main research interests are service innovation, relationship management, customer loyalty, and service failure and recovery. She is a partner of the European consortium Service Design for Innovation, and her research has been published in *Journal of Marketing, MISQ, Journal of Retailing, Journal of Service Research, Journal of Service Management,* and *Journal of Business Research*, among others.

Alexandre Bernardino is an associate professor in the Department of Electrical and Computer Engineering and senior researcher in the Computer and Robot Vision Laboratory of the Institute for Systems and Robotics at IST, the faculty of engineering of Lisbon University. He has published more than 40 research papers in peer-reviewed journals and contributed more than 100 papers to peer-reviewed conferences in the fields of robotics, vision, and cognitive systems. He is associate editor of the journal *Frontiers in Robotics and AI* and of major robotics conferences (ICRA, IROS). He is a Senior Member of the IEEE and the current chair or the IEEE Portugal Robotics and Automation Chapter. His main research interests focus on the application of computer vision, machine learning, cognitive science, and control theory to advanced robotics and automation systems.