Comparing Physical and Virtual Embodiment in a Socially Assistive Robot Exercise Coach for the Elderly

Juan Fasola and Maja J Matarić
Interaction Lab
Computer Science Department
University of Southern California
Los Angeles, CA, USA
{fasola, mataric}@usc.edu

ABSTRACT
We present a socially assistive robot (SAR) system designed to engage elderly users in simple physical exercises. The role of physical embodiment in our exercise system is explored with a large multi-session user study comparing the effectiveness of a physically embodied robot coach to that of a virtually embodied robot coach (a computer simulation of the same robot). The study also illustrates the effects across the age range, as half of the study was conducted with older adults (n=33) and the other half with young adults (n=33). The results of both studies are presented; they show a strong user preference for the physical robot embodiment over the virtual robot, demonstrate the positive effect that physical embodiment has on participant evaluations of the enjoyment and usefulness of the interaction and of the robot in terms of helpfulness, social attraction, and social presence. In addition, the high overall participant ratings of the interaction and robot in the studies demonstrate the viability and usefulness of the SAR system in motivating exercise among elderly users.

Categories and Subject Descriptors
H.1.2 [**Models and Principles**]: User/Machine Systems – Human factors, Software psychology. H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – Evaluation/methodology, Theory and methods.

General Terms
Performance, Design, Experimentation, Human Factors.

Keywords
Socially Assistive Robotics, Exercise, Embodiment, Elderly

1. INTRODUCTION
The aging population is increasing the demand for healthcare services worldwide. By the year 2050, the number of people over the age of 85 will increase three-fold, according to recent estimates [1], and the shortfall of nurses is already becoming an issue [2], [3], [4]. Regular physical exercise has been shown to be effective at maintaining and improving the overall health of elderly individuals [8], [9], [10], [11]. Physical fitness is associated with higher functioning in the executive control processes [5], and is correlated with less atrophy of frontal cortex regions [6] and with improved reaction times [7] compared with the sedentary. Social interaction, and specifically high perceived interpersonal social support, has also been shown to have a positive impact on general mental and physical wellbeing [12], in addition to reducing the likelihood of depression [13], [14], [15], [16]. Thus, the availability of physical exercise therapy, social interaction, and companionship will be critical for the growing elderly population; socially assistive robotics has the potential to help to address this need.

A socially assistive robot (SAR) is a system that employs hands-off interaction strategies, including the use of speech, facial expressions, and communicative gestures, to provide assistance in accordance with the particular healthcare context. Previous SAR work from our research laboratory includes systems that were developed and tested with stroke patients [23], [24], Alzheimer's patients [22], children with autism spectrum disorder [25], [26], as well as healthy adults [27] and healthy elderly adults [28]. SAR systems equipped with motivational, social, and therapeutic capabilities have the potential to facilitate elderly individuals to live independently in their own homes, to enhance their quality of life, and to improve their overall health. In this paper, we present a socially assistive robot system that aims to motivate and engage elderly users in simple physical exercise.

The role of physical embodiment is an interesting and fundamental topic in human-machine interaction. Various studies have observed that people respond differently to physically embodied agents than they do to virtually embodied agents, computer-simulated or otherwise [17], [18], [19], [20], [21], [35]. To study the effect of embodiment in our SAR exercise system, we conducted a large, multiple session user study with both older adults and young adults, in order to study the effectiveness of a physically embodied robot coach compared to that of a virtually embodied robot coach (a computer-simulation of the same robot). The results of both studies show a strong preference among the participants for the physically embodied robot over the virtually embodied robot; they also demonstrate the positive effect that physical embodiment has on participant evaluations of both the interaction and robot.

In the next section we discuss related work in the area of social agent coaches, and embodiment effects. Section 3 presents our SAR system approach and exercise interaction. Section 4 describes our robot embodiment comparison study design, and

Sections 5 and 6 present the results of our embodiment comparison user studies with older adults and young adults, respectively. We conclude the paper with a summary of the key research contributions of this work.

2. RELATED WORK

2.1 Social Agent Coaches

Social agents that aim to assist individuals in health-related tasks such as physical exercise have also been developed in the human-computer interaction (HCI) community. Bickmore and Picard developed a computer-based virtual relational agent that served as a daily exercise advisor by engaging the user in conversation and providing educational information about walking for exercise, asking about the user's daily activity levels, tracking user progress over time while giving feedback, and engaging the user in relational dialogue [33]. French et al. designed and explored the use of a virtual coach to assist manual wheelchair drivers by providing advice and guidance to help users avoid hazardous forms of locomotion [34]. These systems are similar to our SAR exercise system in the manner in which they provide feedback (from a social agent), and in the case of Bickmore's and Picard's work, in the activity being monitored (physical exercise). However, our system differs from theirs in that the agent, a robot in our case, not only provides active guidance, feedback, and task monitoring, but is also directly responsible for instructing and steering the task. Hence, our agent is both an administrator and active participant in the health-related activity, resulting in a unique characteristic of the system: the social interaction between the robot and user is not only useful for maintaining user engagement and influencing intrinsic motivation, but is also needed to achieve the physical exercise task.

2.2 Embodiment Work

Wainer et al. showed that healthy adult participants engaging in a physical/cognitive task, a Towers of Hanoi tabletop game, reported a strong preference for a physically embodiment SAR system over similar video-only agents in terms of appeal, its perceptiveness, watchfulness, helpfulness, and enjoyableness [17], [35]. Powers et al. compared interactions between robots and similar computer-simulated agents that engaged participants in a conversation about basic health habits and found that participants rated the robots as more helpful, more lifelike, and possessing more positive personality traits than the computer-based agents [18]. Bartneck conducted a study comparing the effectiveness of an emotionally expressive robot, eMuu, with its screen character version that engaged users in a simple negotiation task, and found that participants exerted more effort and received higher task scores when interacting with the physical eMuu than with the simulated eMuu [19]. Jung and Lee also demonstrated the positive effects of physical embodiment in relation to interactions with both a Sony Aibo robot and an anthropomorphic dancing robot, April [21]. Bainbridge et al. found that users in a book-moving task were more likely to fulfill an unusual request and afford for more personal space when interacting with a physically present robot than when interacting with a live video feed of the same robot on a computer screen [36].

Physically embodied agents appear to possess what Lee refers to as "social presence" [20] to a greater extent than virtually embodied agents. Social presence mediates how people respond to both embodied and disembodied agents and (strongly) influences the relative success of the social interaction.

It is thus important to note that the embodiment type of a socially assistive agent influences its effectiveness in social interaction, relationship building, gaining user acceptance and trust, and ultimately achieving the desired health outcomes of therapeutic interventions. For these reasons, we explored the role of physical embodiment in our SAR exercise system for the elderly. Further discussion of our embodiment comparison study, along with study results, is presented in Sections 4-6.

3. ROBOT EXERCISE COACH

The robot exercise system consists of a socially assistive robot (physically or virtually embodied) whose purpose is to instruct, evaluate, and encourage users to perform simple exercises. During exercise sessions, the robot asks the user to perform simple seated arm gesture exercises. This type of seated exercise, called "chair exercise" or "chair aerobics", is commonly practiced in senior living facilities and provides grounding for our exercise system. Chair exercises are highly regarded for their accessibility to those with low mobility [8], [9], [10], [11], safety [8], [11], and health benefits such as improved flexibility, muscle strength, ability to perform everyday tasks [8], [10], [11], and even memory recall [9].

3.1 Robot Embodiments

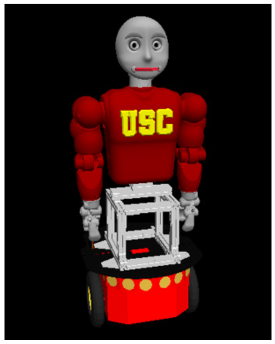
To address the role of the robot's physical embodiment, we use Bandit, a biomimetic anthropomorphic robot platform that consists of a humanoid torso (developed with BlueSky Robotics) mounted on a MobileRobots Pioneer 2DX mobile base. The torso contains 19 controllable degrees of freedom (DOF), which include: 6 DOF arms (x2), 1 DOF gripping hands (x2), 2 DOF pan/tilt neck, 1 DOF expressive eyebrows, and a 2 DOF expressive mouth. A photograph of the physical robot can be seen in Figure 1 (a).

The robot's virtual embodiment consists of a computer simulation of Bandit shown on a 27 inch flat-panel display. The size of the display was chosen to approximate the average size display that would be available in a typical household for use with the robot exercise system, including laptop displays (15 inch), computer monitors (24 inch), and television screens (40 inch). A sample computer simulation image of the virtual robot along with a photograph of the virtual embodiment on the flat-panel display is shown in Figure 1 (b) and (c), respectively.

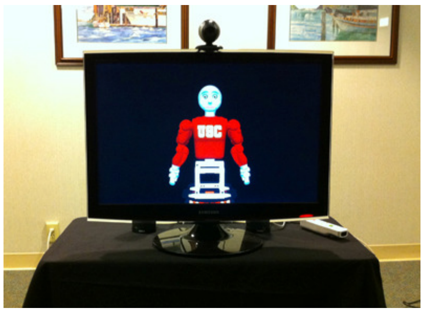
A standard USB camera is used to capture the user's arm movements during exercise interaction, allowing the robot to provide appropriate feedback to the user regarding their performance. For the physical robot embodiment, the camera is placed at the waist of the torso, whereas for the virtual embodiment the camera is attached to the top of the television display. The difference in camera location does not affect the visual recognition accuracy of the user's movements.

The robot's speech is generated by the commercially available NeoSpeech text-to-speech engine and two USB speakers output the synthesized voice to the user. The robot's lip movements are synchronized with the robot's speech so that the lips open at the start and close at the end of spoken utterances.

(a)



(b)



(c)

Figure 1: (a) Physical robot; (b) virtual robot computer simulation; (c) virtual robot on the screen, with camera.

3.2 Interaction Scenario

The one-on-one interaction scenario consists of the user sitting in a chair across from the robot, and each facing the other. In this system, predating the availability of the Kinect, the range of the robot's arm motion in the exercises is restricted to the sides of the body in order to maximize the accuracy of the robot's visual detection of the user's arms upon reproduction. A black curtain is used as a backdrop, also to facilitate the visual perception of the user's arm movements. An updated version of the system will utilize 3D vision via the Kinect, but the 2D nature of the exercises was not noted as an issue by any of the participants in our user studies. An example of the exercise setup and human-robot interaction is shown in Figure 2. A full description of the exercise system and visual recognition procedure is provided in [28]; a summary is provided here for completeness.

The exercise system allows for the user to communicate with the robot through the use of the popular Wiimote wireless Bluetooth button control interface. There are two buttons available for the user to respond to prompts from the robot, labeled "Yes" and "No", and one button for the user to request a rest break at any time during the interaction. The Wiimote is not used for motion sensing of any kind; its role is solely for communication with the robot. It is also important to note that the robot conducts the exercise sessions, evaluates user performance, and gives the user real-time feedback autonomously, without human operator intervention at any time during the exercise sessions.

Four exercise games are available in our system: the Workout game, the Sequence game, the Imitation game, and the Memory game. During an exercise session, the robot chooses the game for the user to play next at pre-specified game change intervals (every one to two minutes). In the Workout game, the robot fills the role of a traditional exercise instructor by demonstrating the arm exercises with its arms, and asking the user to imitate. The robot gives the user feedback in real time, providing corrections when appropriate (e.g., "Raise your left arm and lower your right arm" or "Bend your left forearm inward a little"), and praise in response to each successful imitation (e.g., "Great job!" or "Now you've got the hang of it.").

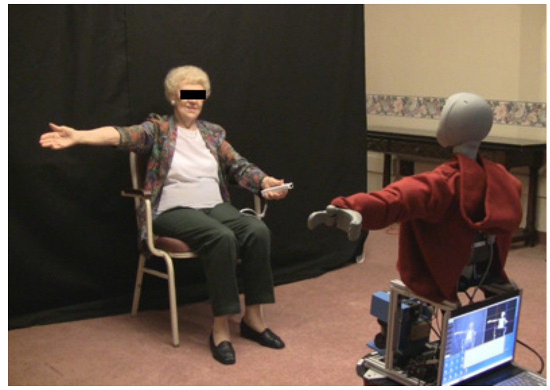


Figure 2: Exercise setup with user and robot

The Sequence game is very similar to the Workout game in that the robot demonstrates arm exercises for the user to repeat. However, instead of showing each gesture for the user to perform only once, the robot shows a sequence of two gestures for the user to repeat a total of three times. The robot keeps verbal count of the number of iterations of the sequence performed in order to guide the user, while providing feedback throughout.

In the Imitation game, the roles of the user and robot are reversed relative to the Workout game: the user becomes the exercise instructor showing the robot what to do. The robot encourages the user to create his/her own arm gesture exercises, and imitates user movements in real time. Since the roles of the interaction are reversed, with the robot relinquishing control of the exercise routine to user, the robot no longer provides instructive feedback on the exercises. However, the robot does continue to speak and engage the user by means of encouragement, general commentary, and movement prompts if necessary.

The goal of the Memory game is for the user to try and memorize an ever-longer sequence of arm gesture poses, and thus compete against his/her own high score. The sequence is determined at the start of the game and does not change for the duration of the game. The arm gesture poses used for each position in the sequence are chosen at random at run time, and there is no inherent limit to the sequence length, thereby making the game challenging for users at all skill levels. As the user successfully memorizes and performs all shown gestures without help from the robot, the robot shows the user two additional gestures to add to the sequence, and hence the game progresses in difficulty.

4. EMBODIMENT COMPARISON STUDY

We designed and conducted an embodiment comparison study to investigate the role of physical embodiment in the robot exercise system. Specifically, the study aimed to compare the effectiveness and participant evaluation of our physical humanoid robot, Bandit, to that of a computer simulation of Bandit shown on a flat-panel display.

4.1 Study Design

Study participants were divided into two groups, physical robot embodiment vs. virtual robot embodiment, and the study consisted of a total of five 20-minute sessions of exercise interaction with the system, conducted over a two-week period.

To analyze the differences between the physical and virtual embodiments in the exercise system, we chose to compare the two conditions with both a between-subjects and within-subjects study design. In addition, a direct comparison between the two conditions, as reported by the participants, was analyzed. The following subsections describe all three robot embodiment comparison methods in detail.

4.1.1 Between-Subjects Analysis

The between-subjects portion of the study constitutes the analysis across both conditions of the data pertinent to the first four exercise sessions, where participants in both groups interacted solely with their designated robot embodiment (physical embodiment group with the physical robot, virtual embodiment group with the virtual robot). Survey data were collected at the end of the first and fourth sessions in order to analyze participant evaluations of the robot and the interaction with the exercise system in both conditions over time. Data gathered from the fourth session post-session surveys were analyzed using a two-tailed independent two-sample T-test assuming unequal variances to test for significant differences among the participant evaluations of the robot and the interaction across both conditions. Survey results from the fourth session were used to perform the comparison analysis as they were less likely to contain scores influenced by the effect of novelty, as opposed to the results obtained from surveys administered after the first session.

4.1.2 Within-Subjects Analysis

The within-subjects portion of the study constitutes the analysis within each condition of the data comparison between the fourth and fifth exercise sessions. Distinct from all previous sessions, in the fifth exercise session participants in both groups interacted with the alternative robot embodiment (physical embodiment group with the virtual robot, virtual embodiment group with the physical robot). The same surveys that were administered after the first and fourth sessions to gather participant evaluations of the robot embodiment and interaction were again given at the end of the fifth session. Data gathered from the fourth and fifth session post-session surveys were analyzed using a two-tailed paired T-test to test for significant differences among the participant evaluations of the robot and the interaction across embodiments within each group. The fourth session survey results were chosen for comparison instead of those from the first session, to minimize any confounding factors (e.g., practice effects due to multiple sessions between evaluations).

4.1.3 Direct Comparison

At the end of the fifth exercise session, after the robot and interaction evaluations were collected, we administered one final survey asking the participants to directly compare the two robot embodiments according to 10 evaluation categories (enjoy more, more useful, better at motivating exercise, etc.). This survey allowed us to gather a general sense of the participants' preferences regarding the physical and virtual embodiments. Analysis of these direct-comparison data serves primarily to support and confirm the results obtained from the between-subjects and within-subjects embodiment comparison analysis.

4.2 Measures

4.2.1 Evaluation of Interaction

There were two dependent measures used to evaluate the interaction with the robot exercise system. The first measure was the *entertainment of the interaction*, collected from participant assessments of the interaction according to six adjectives: enjoyable; interesting; fun; satisfying; entertaining; boring; and exciting (Cronbach's α = .92). Participants were asked to rate how well each adjective described the interaction on a 10-point scale, anchored by "Describes Very Poorly" (1) and "Describes Very Well" (10). Ratings for the adjective "boring" were inverted to keep consistency with the other adjectives that reflect higher scores as being more positive.

The second measure was the perceived *value or usefulness of the interaction*. Similarly, participants were asked to evaluate how well each of the following four adjectives described the interaction: useful; beneficial; valuable; and helpful (Cronbach's α = .96). The same 10-point scale anchored by "Describes Very Poorly" (1) and "Describes Very Well" (10) was used in the evaluation.

4.2.2 Evaluation of Robot

Companionship of the robot was measured from participant responses to nine 10-point semantic differential scales concerning the following robot descriptions: bad/good; not loving/loving; not friendly/friendly; not cuddly/cuddly; cold/warm; unpleasant/ pleasant; cruel/kind; bitter/sweet; and distant/close (Cronbach's α = .89). These questions were derived from the Companion Animal Bonding Scale of Poresky et al. [37].

Participants evaluated the *usefulness of the robot* by rating four robot characteristics on a 10-point scale: useful; valuable; beneficial; and helpful (Cronbach's α = .96). The rating scale was anchored by "Not at all" (1) and "Absolutely" (10). Using the same rating scale, the *intelligence of the robot* was measured according to the following four adjectives: competent; clever; intelligent; and smart (Cronbach's α = .93).

To help capture the robot's social attributes, we measured both the *social attraction towards the robot* and the *social presence of the robot*. Social attraction was measured by a modified version of the Interpersonal Attraction Scale of McCroskey and McCain [38]. Participants reported their level of agreement towards the following four statements: I think Bandit could be a friend of mine; I think I could spend a good time with Bandit; I could establish a personal relationship with Bandit; I would like to spend more time with Bandit (Cronbach's α = .88). The statements were rated on a 7-point scale anchored by "Very Strongly Disagree" (1) and "Very Strongly Agree" (7). Social presence was measured by a 10-point scale anchored by "Not at all" (1) and "Very much" (10) with questions from Jung and Lee [21] (e.g., While you were exercising with Bandit, how much did you feel as if you were interacting with an intelligent being?) (Cronbach's α = .87).

To assess the perceptions of the exercise capabilities of the system, we measured participant evaluations of the robot as an *exercise partner*. Four items were used, each measured according to a 10-point scale anchored by "Not at all" (1) and "Very much"

(10); they were: How much did you enjoy exercising with Bandit?; How likely would you be to recommend Bandit as an exercise partner to your friends?; How much would you like to exercise with Bandit in the future?; How much have you been motivated to exercise while interacting with Bandit? (Cronbach's $\alpha = .93$).

4.3 Hypotheses

Based on the related research on embodiment effects, discussed in Section 2, five hypotheses were established for this study:

Hypothesis 1: Participants will evaluate the entertainment of their interaction with the physical robot more positively than their interaction with the virtual robot.

Hypothesis 2: Participants will evaluate the value/usefulness of their interaction with the physical robot more positively than their interaction with the virtual robot.

Hypothesis 3: Participants will evaluate the usefulness/helpfulness of the physical robot more positively than the virtual robot.

Hypothesis 4: Participants will be more socially attracted to the physical robot than the virtual robot.

Hypothesis 5: Participants will experience a greater sense of social presence when interacting with the physical robot than when interacting with the virtual robot.

Hypothesis 6: Participants will report a clear preference for the physical robot over the virtual robot when asked to compare both embodiments directly.

5. RESULTS

5.1 Participant Statistics

We recruited elderly individuals to participate in the study through a partnership with the be.group, an organization of senior living communities in Southern California, through flyers and word-of-mouth. We offered a $50 Target gift card to those willing to participate in all five sessions of the study. Thirty-seven people responded, of whom four were omitted due to inconsistent/incorrect answers to survey questions used to identify questionable survey results. Thus, we were left with 33 participants whose data were analyzed. Half of the participants were placed in the physical robot group (n = 16), and the other half were placed in the virtual robot group (n = 17). The sample population consisted of 27 female participants (82%) and 6 male participants (18%). Participants' ages ranged from 68-88, and the average age was 76 (S.D. = 6.32).

5.2 Between-Subjects Comparison

A two-tailed independent two-sample T-test was performed on the survey data following the fourth exercise session, to compare participant evaluations of the robot embodiments as well as their interactions with them across the two study groups. Table 1 provides the complete set of between-subjects comparison results.

Consistent with Hypothesis 1, the participants evaluated the interaction with the physical robot embodiment as more enjoyable/entertaining than the interaction with the virtual robot embodiment (t[31]=2.29, p<.03). Hypothesis 2 was supported by the data as well, as the participants evaluated the interaction with the physical robot as more valuable/useful than the interaction with the virtual robot (t[29]=2.72, p=.01).

Regarding the evaluations of both robot embodiments, the participants rated the physical robot as more useful/helpful than the virtual robot (t[31]=2.66, p=.01), consistent with Hypothesis 3. The participants also rated the physical robot as more socially attractive than the virtual robot (t[30]=2.09, p<.05), supporting Hypothesis 4. Concerning social presence, the data were consistent with Hypothesis 5, as the participants reported feeling a stronger sense of social presence with the physical robot than that felt with the virtual robot (t[23]=2.59, p<.02).

Evaluations of the non-hypothesis-testing system performance measures were also favorable to the physical robot, as the participants rated the physical robot as somewhat more of a companion (t[30]=1.81, p<.08), more intelligent (t[31]=1.96, p<.06), and a moderately better exercise partner (t[31]=1.87, p=.07) than the virtual robot.

5.3 Within-Subjects Comparison

To test for significant differences between the participant evaluations of the robot and the interaction for both embodiments within each study group, we used a two-tailed paired T-test to analyze the data gathered from the fourth and fifth session post-session surveys. See Table 1 for all within-subjects results.

The data corresponding to the participants in the virtual embodiment group agree with the results of the between-subjects comparison in their support for Hypotheses 1-5. The participants found the interaction with the physical robot embodiment more enjoyable/entertaining (t[16]=-2.02, p<.06), and more valuable/useful (t[16]=-2.86, p =.01) than the interaction with the virtual robot embodiment. In addition, the participants in the virtual robot group evaluated the physical robot more positively than the virtual robot in terms of social attraction to a degree (t[15]=-1.79, p=.09), social presence (t[16]=-2.10, p=.05), companionship (t[16]=-2.32, p<.05), intelligence (t[16]=-2.85, p=.01), and as an exercise partner (t[16]=-2.78, p<.02).

Data from the physical embodiment group support Hypothesis 5, as the participants felt greater social presence with the physical robot than with the virtual robot (t[15]=3.70, p=.002), and are consistent with Hypotheses 2 and 4 though not significantly, as the participants reported the interaction with the physical robot to be more valuable/useful than the interaction with the virtual robot (t[15]=1.79, p=.09), and felt more socially attracted to the physical robot as well (t[15]=1.99, p<.07). Hypotheses 1 and 3 were not supported by the data from the physical embodiment group, possibly due to positive carryover effects.

5.4 Direct Comparison

At the end of the final exercise session, participants were asked to directly compare both embodiments with respect to 10 different evaluation categories; results are provided in Table 2.

Supporting Hypothesis 6, the direct comparison results show a clear preference for the physical robot embodiment over the virtual embodiment, with the physical robot receiving 81% of the positive trait votes vs. 19% for the virtual robot among participants who chose one embodiment over the other, representing 85% of the sample. Other notable results include the high number of participants who rated the physical robot as more enjoyable (25 votes, 76%), a better exercise partner (27 votes, 82%), more interesting (23 votes, 70%), more useful (21 votes,

Table 1: Results of between-subjects and within-subjects data comparison for all n=33 participants

Dependent Measure	Physical Robot	Virtual Robot	PP. (4th)	PV. (5th)	VV. (4th)	VP. (5th)
Interaction Evaluation	*Between-Subjects Analysis*		*Within-Subjects Analysis*			
Enjoyable/Entertaining	7.51 (1.77)*	6.00 (2.01)	7.51 (1.77)	6.94 (2.21)	6.00 (2.01)	7.11 (2.35)+
Valuable/Useful	8.14 (1.66)*	6.19 (2.39)	8.14 (1.66)	7.70 (2.13)++	6.19 (2.39)	7.51 (2.26)*
Robot Evaluation						
Useful/Helpful	8.11 (1.98)*	6.26 (1.98)	8.11 (1.98)	8.28 (1.61)	6.26 (1.98)	7.44 (2.48)*
Social Attraction	4.70 (1.40)*	3.61 (1.54)	4.70 (1.40)	4.31 (1.43)+	3.61 (1.54)	4.36 (1.58)++
Social Presence	7.88 (0.94)*	6.47 (2.01)	7.88 (0.94)	6.98 (0.97)**	6.47 (2.01)	7.22 (1.66)+
Companion	7.48 (2.07)++	6.23 (1.84)	7.48 (2.07)	7.12 (1.94)	6.23 (1.84)	7.42 (1.87)*
Intelligence	8.17 (2.02)+	6.76 (2.09)	8.17 (2.02)	7.61 (1.54)	6.76 (2.09)	7.79 (2.66)*
Exercise Partner	7.18 (2.17)++	5.76 (2.18)	7.18 (2.17)	6.95 (1.60)	5.76 (2.18)	7.01 (2.16)*

Note: The numbers shown are the mean and the standard deviation in parenthesis of participant evaluations.
PP, PV = Physical robot group evaluating physical robot, virtual robot; VV, VP= Virtual robot group evaluating virtual robot, physical robot.
++p<.10, +p<.06, *p<.05, **p<.01, ***p<.001

64%), better at motivating exercise (22 votes, 67%), and as the embodiment they would choose to continue exercising with after the study (25 votes, 76%).

Table 2: Participant responses to the direct comparison survey

	Physical	*Virtual*	*Both Equal*
Enjoy More	25 (76%)	6 (18%)	2 (6%)
More Intelligent	13 (40%)	6 (18%)	14 (42%)
More Useful	21 (64%)	7 (21%)	5 (15%)
Prefer to Exercise with	27 (82%)	4 (12%)	2 (6%)
Better at Motivating	22 (67%)	4 (12%)	7 (21%)
More Frustrating	10 (30%)	14 (43%)	9 (27%)
More Boring	4 (12%)	17 (52%)	12 (36%)
More Interesting	23 (70%)	5 (15%)	5 (15%)
More Entertaining	25 (76%)	4 (12%)	4 (12%)
Choice from now on	25 (76%)	7 (21%)	1 (3%)

6. EMBODIMENT STUDY EXPANSION

In an effort to 1) increase the size of the sample population overall, 2) balance the gender distribution within the overall sample, and 3) observe the differences if any between age groups, we expanded the embodiment comparison study by recruiting young adults from our university's campus.

6.1 Participant Statistics

Recruitment procedures, with the addition of email, and study guidelines remained the same, with participants being offered a $50 Target gift card to complete all five sessions of the study. Thirty-four people responded, of whom one was omitted as an outlier. Thus, data from 33 participants were analyzed. Half of the participants were placed in the physical embodiment group (n=16), and the other half were placed in the virtual embodiment group (n=17). This sample population consisted of 6 female (18%) and 27 male (82%) participants. Participants' ages ranged from 20-33, with an average age of 24 (S.D. = 3.14).

6.2 Results

The within-subjects comparison results showed an overwhelming preference for the physical robot embodiment over the virtual robot embodiment, as the data from both participant groups support Hypotheses 1-5, although to a lesser extent in the virtual robot group for Hypotheses 2 and 5, as the differences were pronounced though not quite reaching significance.

Direct comparison results were also strongly in favor of the physical robot, consistent with Hypothesis 6, with the physical robot coach receiving 86% of the votes among positive trait items vs. 14% for the virtual robot coach among the participants who chose one embodiment over the other (97% of the sample).

The between-subjects embodiment comparison results, however, were not found to be significant. One possible explanation for this disparity with regard to the within-subjects and direct comparison results is that the participants, being young adults and recognizing that the SAR system was originally designed for the elderly and not their own demographic, may have been more generous in their initial evaluations of the system regardless of embodiment; however, the scores were rectified after the fifth session, as is evident in the data.

6.3 Combined Results

We combined the study populations of both the older and young adults to create a larger combined sample population of 66 participants, with ages ranging from 20-88, to obtain a more general illustration regarding the effects of embodiment. Each participant engaged in five study sessions, yielding a total of 330 sessions of one-on-one interaction with our SAR exercise system. Approximately half of the participants were placed in the physical robot group (n=32), with the other half placed in the virtual robot group (n=34). There were 33 female participants (50%) and 33 male participants (50%).

As expected, the combined results show a strong participant preference for the physical robot over the virtual robot, supporting Hypotheses 1-6. The between-subjects results show significance in all but one measure (which misses by a small margin), as the participants rated the physical robot interaction as more enjoyable/entertaining ($t[63]=2.13$, $p<.05$) and more valuable/useful ($t[60]=2.32$, $p<.03$), in addition to evaluating the physical robot as more helpful ($t[64]=2.25$, $p<.03$), more socially

Table 3: Results of between-subjects and within-subjects data comparison for combined study (n=66)

Dependent Measure	Physical Robot	Virtual Robot	PP. (4th)	PV. (5th)	VV. (4th)	VP. (5th)
Interaction Evaluation	*Between-Subjects Analysis*		*Within-Subjects Analysis*			
Enjoyable/Entertaining	7.41 (1.58)*	6.50 (1.85)	7.41 (1.58)	6.69 (1.83)**	6.50 (1.85)	7.50 (1.98)**
Valuable/Useful	7.89 (1.56)*	6.83 (2.13)	7.89 (1.56)	7.30 (1.88)**	6.83 (2.13)	7.80 (1.88)**
Robot Evaluation						
Useful/Helpful	7.92 (1.68)*	6.93 (1.87)	7.92 (1.68)	7.71 (1.61)	6.93 (1.87)	7.83 (2.00)***
Social Attraction	4.62 (1.21)*	3.98 (1.33)	4.62 (1.21)	3.93 (1.35)**	3.98 (1.33)	4.73 (1.30)**
Social Presence	7.58 (1.11)+	6.86 (1.83)	7.58 (1.11)	6.48 (1.38)***	6.86 (1.83)	7.46 (1.81)**
Companion	7.66 (1.59)++	6.93 (1.63)	7.66 (1.59)	6.92 (1.67)*	6.93 (1.63)	7.75 (1.55)**
Intelligence	7.65 (1.77)++	6.85 (2.04)	7.65 (1.77)	6.90 (1.79)*	6.85 (2.04)	7.55 (2.23)**
Exercise Partner	7.25 (1.94)	6.61 (2.20)	7.25 (1.94)	6.40 (1.73)**	6.61 (2.20)	7.50 (1.92)**

++p<.10, +p<.06, *p<.05, **p<.01, ***p<.001

attractive (t[63]=2.02, p<.05), and possessing greater social presence (t[55]=1.95, p<.06). The within-subjects analysis shows results equally positive towards the physical robot embodiment, with both participant groups achieving significance across all dependent measures, with the lone exception being the physical robot group's assessment of both robots' helpfulness. Table 3 contains the full between-subjects and within-subjects results.

The direct comparison results are also strong, with 84% of the positive trait items listed in the survey attributed to the physical robot embodiment by the participants who chose one embodiment over the other, representing 91% of the combined sample population. The full results show that the physical robot received 402 positive votes (76%), the virtual robot 79 positive votes (15%), with 47 positive votes being split among them (9%).

7. CONCLUSIONS

We have presented a socially assistive robot (SAR) system designed to engage elderly users in physical exercise in a seated aerobic exercise scenario. We investigated the role of physical embodiment in our exercise system by conducting a large multi-session user study comparing the effectiveness of a physically embodied robot coach to that of a virtually embodied robot coach (a computer simulation of the same robot), and presented results from interaction with elderly participants (n=33), young adults (n=33), and both populations combined (n=66). Analysis of the results was performed using between-subjects, within-subjects, and direct comparison methods.

The results of the embodiment comparison study, both parts individually and combined, show a strong user preference for the physical robot embodiment over the virtual robot embodiment in our SAR exercise system across age ranges. Consistent with our stated hypotheses, participants reported the interaction with the physical robot as being more enjoyable/entertaining and more valuable/useful than the interaction with the virtual robot. Furthermore, the participants evaluated the physical robot as more helpful/useful, more socially attractive, and felt a greater sense of social presence with the physical robot than with the virtual robot.

These results are consistent with prior related work on embodiment effects in human-agent interaction. The key contributions of our work are in demonstrating the positive effects of physical embodiment with users who were recruited from a target healthcare population for interaction with a system specifically designed to address their healthcare needs (older adults), that the data analyzed were gathered from multiple sessions of interaction to help reduce the effect of novelty which can often be seen in single session studies, and that the effects were demonstrated across age range.

8. ACKNOWLEDGMENTS

We are grateful to the be.group (formerly Southern California Presbyterian Homes) staff and participating residents. Thanks to Kwan Min Lee for providing insights into the study design and survey questions.

9. REFERENCES

[1] Centers for Disease Control and Prevention. (2003) Morbidity and Mortality Weekly Report. Department of Health and Human Services.

[2] American Association of Colleges of Nursing. (2010) Nursing Shortage Fact Sheet.

[3] American Health Care Association. (July 2008) Summary of 2007 AHCA Survey Nursing Staff Vacancy and Turnover In Nursing Facilities.

[4] Buerhaus P. (2008) Current and future state of the US nursing workforce. JAMA 300: 2422.

[5] Colcombe S and Kramer A. (2003) Fitness effects on the cognitive function of older adults. Psychological Science

[6] Colcombe S, Kramer A, Erickson K, et al. (2004) Cardiovascular fitness, cortical plasticity, and aging. Proceedings of the National Academy of Sciences of the United States of America 101: 3316.

[7] Spirduso W and Clifford P. (1978) Replication of age and physical activity effects on reaction and movement time. Journal of Gerontology 33: 26.

[8] Baum E, Jarjoura D, Polen A, et al. (2003) Effectiveness of a group exercise program in a long-term care facility: a randomized pilot trial. Journal of the American Medical Directors Association 4: 74.

[9] Dawe D and Moore-Orr R. (1995) Low-intensity, range-of-motion exercise: invaluable nursing care for elderly patients. Journal of Advanced Nursing 21: 675-681.

[10] McMurdo MD and Rennie LM. (1993) A controlled trial of exercise by residents of old people's homes. Age and Ageing 22: 11.

[11] Thomas V and Hageman P. (2003) Can neuromuscular strength and function in people with dementia be rehabilitated using resistance-exercise training? Results from a preliminary intervention study. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences 58: M746.

[12] Moak ZB and Agrawal A. (2010) The association between perceived interpersonal social support and physical and mental health: results from the national epidemiological survey on alcohol and related conditions. Journal of Public Health 32: 191-201.

[13] George L, Blazer D, Hughes D, et al. (1989) Social support and the outcome of major depression. The British Journal of Psychiatry 154: 478.

[14] Paykel E. (1994) Life events, social support and depression. Acta Psychiatrica Scandinavica 89: 50-58.

[15] Stice E, Ragan J and Randall P. (2004) Prospective relations between social support and depression: Differential direction of effects for parent and peer support? Journal of Abnormal Psychology 113: 155-159.

[16] Stansfeld S, Rael G, Head J, et al. (1997) Social support and psychiatric sickness absence: a prospective study of British civil servants. Psychological Medicine 27: 35-48.

[17] Joshua Wainer, David J. Feil-Seifer, Dylan A. Shell, and Maja J. Matarić. "Embodiment and Human-Robot Interaction: A task-based perspective". In IEEE Proceedings of the International Workshop on Robot and Human Interactive Communication (RO-MAN), pages 872-877, Jeju Island, South Korea, Aug 2007.

[18] Aaron Powers, Sara Kiesler, Susan Fussell, and Cristen Torrey. Comparing a computer agent with a humanoid robot. In Proceedings of the ACM/IEEE international conference on Human-robot interaction (HRI '07). ACM, New York, NY, USA, 145-152, 2007.

[19] Christoph Bartneck. Interacting with an embodied emotional character. In Proceedings of the 2003 international conference on Designing pleasurable products and interfaces (DPPI '03). ACM, New York, NY, USA, 55-60, 2003.

[20] Lee. K. M, 2004. Presence, Explicated. Communication Theory 14 (1), 27-50.

[21] Jung, Y., & Lee, K. M. (2004). Effects of physical embodiment on social presence of social robots. Proceedings of Presence 2004, 80-87.

[22] Tapus A, Tapus C and Mataric M. (2009) The use of socially assistive robots in the design of intelligent cognitive therapies for people with dementia. International Conference on Rehabilitation Robotics. Kyoto, Japan: IEEE, 924-929.

[23] M. J. Matarić, J. Eriksson, D. J. Feil-Seifer, and C. J. Winstein. "Socially Assistive Robotics for Post-Stroke Rehabilitation". In Journal of NeuroEngineering and Rehabilitation, 4(5), Feb 2007.

[24] A. Tapus, C. Tapus, and Maja J Matarić. User-robot personality matching and robot behavior adaptation for post-stroke rehabilitation therapy. Intelligent Service Robotics, 1(2):169–183, April 2008.

[25] D. J. Feil-Seifer and M. J. Matarić. "Towards the integration of socially assistive robots into the lives of children with ASD". Working Notes Human-Robot Interaction Workshop on Societal Impact: How Socially Accepted Robots Can be Integrated in our Society, San Diego, CA, Mar 2009.

[26] D. J. Feil-Seifer and M. J. Matarić. "Using Proxemics to Evaluate Human-Robot Interaction". Poster paper in Proceedings of the International Conference on Human-Robot Interaction (HRI), pages 143-144, Osaka, Japan, 2010

[27] J. Fasola and M. J. Matarić. "Robot Motivator: Increasing User Enjoyment and Performance on a Physical/Cognitive Task". In Proc. IEEE International Conference on Development and Learning (ICDL), MI, USA, 2010.

[28] Fasola, J. and Matarić, M. J. "Robot Exercise Instructor: A Socially Assistive Robot System to Monitor and Encourage Physical Exercise for the Elderly". In 19th IEEE International Symposium in Robot and Human Interactive Communication (Ro-Man 2010), Viareggio, Italy, Sep 2010.

[29] S. Dubowsky, F. Genot, S. Godding, H. Kozono, A. Skwersky, H. Yu, and L. Shen Yu. PAMM - a robotic aid to the elderly for mobility assistance and monitoring. In IEEE International Conference on Robotics and Automation, vol. 1, 570–576, Apr. 2000.

[30] M. Montemerlo, J. Prieau, S. Thrun, and V. Varma. Experiences with a mobile robotics guide for the elderly. In Proc. of the AAAI Nat Conf on Artif Intelligence, Edmonton, Alberta, Aug 2002, pp. 587–592.

[31] K. Wada, T. Shibata, T. Saito, and K. Tanie. Analysis of factors that bring mental effects to elderly people in robot assisted activity. In Proceedings of the International Conference on Intelligent Robots and Systems, vol. 2, 1152–1157, Oct. 2002.

[32] C. Kidd, W. Taggart, and S. Turkle. A sociable robot to encourage social interaction among the elderly. In Int Conf on Robotics and Automation (ICRA), Orlando, USA, 2006.

[33] T.W. Bickmore and R.W. Picard. Establishing and maintaining long-term human-computer relationships. ACM Trans. Comput.-Hum. Interact. 12, 2 (Jun. 2005), 293-327.

[34] French, B., Tyamagundlu, D., Siewiorek, D., Smailagic, A., Ding, D. Towards a Virtual Coach for manual wheelchair users. Proc. of Int'l IEEE Symposium of Wearable Computers, IEEE Computer Society Press, Sept 2008.

[35] Joshua Wainer, David J. Feil-Seifer, Dylan A. Shell, and Maja J. Matarić. "The role of physical embodiment in human-robot interaction". In IEEE RO-MAN, pages 117-122, Hatfield, United Kingdom, Sep 2006.

[36] W.A. Bainbridge, J.W. Hart, E.S. Kim, and B. Scassellati. The Benefits of Interactions with Physically Present Robots over Video-Displayed Agents. I. J. Social Robotics 3(1): 41-52 (2011).

[37] Poresky, R.H., Hendrix, C., Mosier, J.E., Samuelson, M. 1987. Companion animal bonding scale: internal reliability and construct validity. Psychological Reports 60, 743–746.

[38] McCroskey, J.C., McCain, T.A., 1974. The measurement of interpersonalattraction. Speech Monographs 41, 261–266.