INVITED PAPER

Using Socially Assistive Human–Robot Interaction to Motivate Physical Exercise for Older Adults

A robot designed to engage elderly users in physical exercise is described in this paper; a user study indicates a strong user preference for a relational robot.

By Juan Fasola and Maja J Matarić, *Fellow IEEE*

ABSTRACT | In this paper, we present the design, implementation, and user study evaluation of a socially assistive robot (SAR) system designed to engage elderly users in physical exercise aimed at achieving health benefits and improving quality of life. We discuss our design methodology, which incorporates insights from psychology research in the area of intrinsic motivation, and focuses on maintaining engagement through personalized social interaction. We describe two user studies conducted to test the motivation theory in practice with our system. The first study investigated the role of praise and relational discourse in the exercise system by comparing a relational robot coach to a nonrelational robot coach. The second study evaluated participant preferences regarding user choice in the task scenario. Both studies served to evaluate the feasibility and overall effectiveness of the robot exercise system. The results of both studies are presented; they show a strong user preference for the relational over the nonrelational robot in terms of enjoyableness, companionship, and as an exercise coach, varying user preferences regarding choice, and high user ratings of the system across multiple metrics. The outcomes of the presented user studies, brought together, support the motivational capabilities of the robot, and demonstrate the viability and usefulness of the system in motivating exercise in elderly users.

KEYWORDS | Exercise therapy; human–robot interaction; intrinsic motivation; quality of life technology; socially assistive robotics

Manuscript received April 8, 2011; revised November 25, 2011 and February 20, 2012; accepted March 22, 2012. Date of publication July 9, 2012; date of current version July 13, 2012. This work was supported by a grant from the Robert Wood Johnson Foundation's Pioneer Portfolio through its national program, "Health Games Research: Advancing Effectiveness of Interactive Games for Health," and the National Science Foundation under Grants IIS-0713697, CNS-0709296, and IIS-1117279.
The authors are with the University of Southern California, Los Angeles, CA 90089 USA (e-mail: fasola@usc.edu; mataric@usc.edu).

Digital Object Identifier: 10.1109/JPROC.2012.2200539

I. 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 fivefold, according to recent estimates [1], and the shortfall of nurses is already becoming an issue [2]–[4]. Regular physical exercise has been shown to be effective at maintaining and improving the overall health of elderly individuals [8]–[11]. Physical fitness is associated with higher functioning in the executive control processes [5], 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]–[16]. Among the many healthcare services that will need to be provided, physical exercise therapy, social interaction, and companionship can be addressed by socially assistive robotics technology.

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 [54]. Previous SAR work from our research laboratory includes systems that were developed and tested for 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].

This paper focuses on the design methodology, implementation details, and user study evaluations of a SAR system that aims to motivate and engage elderly users in physical exercise as well as social interaction to help address the physical and cognitive healthcare needs of the growing elderly population. SAR systems equipped with such 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.

The rest of this paper is organized as follows. In the next section, we discuss related work in the area of assistive robotics for the elderly. Section III presents our SAR system approach and design methodologies. Section IV introduces our SAR humanoid robot platform along with the implementation details of our SAR exercise system. In Sections V and VI, we discuss two user studies conducted with our system to investigate and evaluate the effects of different motivational techniques, to test system effectiveness, and to obtain user feedback. We conclude the paper with a summary of the key research contributions of this work.

II. RELATED WORK

A. Robots for the Elderly

The literature that addresses the area of assistive robotics for the elderly is limited. Representative work includes robots that focus on providing assistance for functional needs, such as mobility aids and navigational guides. Dubowsky *et al.* developed a robotic cane/walker device designed to help individuals by functioning as a mobility aid that provides physical support when walking as well as guidance and health monitoring of a user's basic vital signs [29]. Montemerlo *et al.* designed and pilot tested a robot that escorts elderly individuals in an assisted living facility, reminds them of their scheduled appointments, and provides informational content such as weather forecasts [30].

Researchers have also investigated the use of robots to help address the social and emotional needs of the elderly, including reducing depression and increasing social interaction with peers. Wada *et al.* studied the psychological effects of a stuffed seal robot, Paro, used to engage seniors at a day service center. The study found that Paro, which was always accompanied by a human handler, was able to consistently improve the moods of elderly participants who had spent time petting it and engaging with it over the course of a six-week period [31]. Kidd *et al.* used Paro in another study that found it to be useful as a catalyst for social interaction. They observed that seniors who participated with the robot in a group were more likely to interact socially with each other when the robot was present and powered on, than when it was powered off or absent [32].

Perhaps the most related robotic system for the elderly to our SAR exercise system is the work of Matsusaka *et al.*, who developed an exercise demonstrator robot, TAIZO, to aid human demonstrators teaching simple arm exercises to a training group [52]. However, this robot was not autonomous: it was controlled via key input or voice by the lead human demonstrator, and did not have any sensors for perceiving the users. Hence, the system did not provide any real-time feedback, active guidance, or personalized training.

B. Social Agent Coaches

Social agents that aim to assist individuals in health-related tasks such as physical exercise have been developed in both the human–computer interaction (HCI) and human–robot interaction (HRI) communities. 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 dialog [33]. Kidd and Breazeal developed a tabletop robot to serve as a daily weight-loss advisor, which interacted through a touchscreen interface, tracked user progress and the user–robot relationship state over time, and was tested in a six-week field study with participants at home [51]. 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 with the exception of French's work, in the activity being monitored (physical exercise). However, our system differs from all 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 for the system: the social interaction between the robot and the user is not only useful for maintaining user engagement and influencing intrinsic motivation, but is also an instrumental necessity in achieving the physical exercise task.

III. SAR APPROACH

In designing our system to help address the physical exercise needs of the elderly population, we followed the design methodology which asserts that the SAR agent must possess: 1) the ability to influence the user's intrinsic motivation to perform the task; and 2) the ability to personalize the social interaction to maintain user engagement in the task and build trust in the task-based human–robot relationship. The following elaborates on the importance of both of these qualities in the context of providing healthcare interventions, as well as details how each was incorporated into our SAR exercise system.

A. Intrinsic Motivation

Motivation is a fundamental tool in establishing adherence to a therapy regimen or task scenario and in promoting behavior change. There are two forms of motivation: intrinsic motivation, which comes from within a person, and extrinsic motivation, which comes from sources external to a person. Extrinsic motivation, though effective for short-term task compliance, has been shown to be less effective than intrinsic motivation for long-term task compliance and behavior change [35].

Intrinsic motivation, however, can be, and often is, affected by external factors. In a task scenario, the instructor (a SAR, in our case) can impact the user's intrinsic motivation through verbal feedback. Praise, for example, is considered a form of positive feedback and has the potential to increase the user's intrinsic motivation for performing the task, whereas criticism, a form of negative feedback, tends to negatively impact the user's intrinsic motivation [38], [39]. The effect of positive feedback, however, is closely tied to the user's own perceived competence at the task. Once the user believes he is competent at the task, additional praise no longer affects his intrinsic motivation. Our SAR exercise system provides positive feedback to the user in the form of praise upon correct completion of the given exercises, and never gives negative feedback so as to avoid diminishing intrinsic motivation to engage in the exercise task.

Indirect competition, wherein the user is challenged to compete against an ideal outcome, has also been shown to increase user enjoyment on an otherwise noncompetitive task [36]. For example, when the user is shown her high score on the task, her intrinsic motivation for the task tends to increase, as she strives to better her previous performance. Thus, in a task scenario, it is important that the task instructor continually report to the user his/her performance scores during the task, for motivational purposes. Our robot exercise instructor implements this strategy by reporting the user's personal high scores during two of the three exercise games played.

Verbal feedback provided to the user by the instructor certainly plays an important role in task-based motivation, but the task itself and how it is presented to the user perhaps plays an even more significant role. Csikszentmihalyi's research suggests that "when one engages in an optimally challenging activity with respect to one's capacities, there is a maximal probability for task-involved enjoyment or flow" [37]. He also states that intrinsically motivated activities are those characterized by enjoyment. Simply put, people are "intrinsically motivated under conditions of optimal challenge" [40]. If a task is below the optimal challenge level, it is too easy for the user and results in boredom. Alternatively, if the task is above the optimal challenge level, it is too hard and causes the user to get anxious or frustrated. Therefore, an instructor that oversees user performance in a task scenario must be able to continually adjust the task to meet the appropriate needs of the user in order to increase or maintain intrinsic motivation to perform the task. We have incorporated these guidelines for achieving the optimal challenge level for the user into our SAR exercise system. For example, the exercise games are changed at regular intervals to prevent the user from getting bored or frustrated with any one of them. In addition, the Memory game, discussed in the next section, challenges the user with progressively more difficult exercise sequences based on the user's performance level.

Another task characteristic with the potential to influence user enjoyment is the incorporation of direct user input. Studies have shown that tasks that support user autonomy and self-determination lead to increased intrinsic motivation, self-esteem, creativity, and other related variables among the participants [42], all of which are important for achieving task adherence and long-term behavior change. Self-determination, represented in the task in the form of choice of activity [41], choice of difficulty level [42], and choice of rewards [43], has been shown to either increase or be less detrimental to intrinsic motivation than similar task conditions that do not involve choice. In the context of our SAR exercise system, user choice is a very interesting research question and one that we investigated with a user study to test the role of choice in the exercise scenario. The study design and results are presented in Section VI.

B. Social Interaction and Personalization

Many social intricacies contribute to the foundation of a meaningful relationship, both in HCI (as detailed by Bickmore and Picard [33]) and in HRI. These include empathy, humor, references to mutual knowledge, continuity behaviors, politeness, and trust, among others. We place great importance on these relationship building tools; therefore, we integrated each, in one form or another, into the social interaction component of our robot exercise instructor.

Our primary focus was on eliminating the perceived repetitiveness of the robot's verbal instructions/comments. We believe that if the robot is perceived by the user as repetitive and hence predictable, this can lead to a decrease in the perception of the robot's intelligence by the user, and ultimately to a loss of trust in the robot's helpfulness in motivating exercise. We therefore placed special attention on adding variety to the robot's utterances. Toward this end, the robot always drew from a list of phrases that emphasized the same point when speaking to the user, choosing randomly at run time. For example, there were more than ten different ways in which the robot could praise the user (e.g., "Awesome!," "Nice job!," "Fantastic!"). Furthermore, if the robot did need to repeat itself exactly, for example when providing the same feedback comment during one of the exercise games, it added filler words to the given phrase, such as the user's name or the word "try" or both (e.g., "Try to raise your left arm," "John, raise your left arm").

Adding the user's name to the interaction dialog was an important part of our system design, not only to add variability, but also for its relationship building effect [33]. The robot always used the user's name at the first greeting, and also when bidding farewell at the end of a session. Having the robot refer to the user by name is an important part of personalizing the interaction, along with providing direct feedback specific to the individual user's performance level and performance history during the games, and referencing mutual knowledge. Our SAR exercise system introduced continuity by having the robot refer to previous sessions with the user upon introduction, reference planned future sessions at the end of interaction, and refer to past exercise performance, such as when reporting previous high scores.

Fig. 1. ***Exercise setup with user and robot facing each other.***

IV. ROBOT EXERCISE SYSTEM

In this section, we present the design and implementation details of the SAR exercise system, including the motivation behind the types of exercise routines, the humanoid robot platform, the different exercise games, and the robot's visual user arm motion recognition procedure.

A. System Overview

The exercise scenario consists of a SAR whose purpose is to instruct, evaluate, and encourage users to perform simple exercises. The scenario is one-on-one; the robot focuses its attention on the user in order to provide timely and accurate feedback, and to maximize the effectiveness of the exercise session for the user. In the setup, the user is seated in a chair in front of the robot; the user and the robot face each other. A black curtain is used as a backdrop to facilitate the visual perception of the user's arm movements. The exercise setup is shown in Fig. 1.

During the exercise sessions, the robot asks the user to perform simple seated arm gesture exercises. 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. 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]–[11], for their safety as they reduce the possibility of injury due to falling from improper balance [8], [11], and for their health benefits such as improved flexibility [8], [10], muscle strength [8], [10], [11], ability to perform everyday tasks [8], [10], [11], and even memory recall [9].

The user is able to communicate with the robot through a wireless button control interface, the popular Wiimote remote controller, which communicates via Bluetooth with the button labels modified to suit our system. There are two buttons available to 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.

It is important to note that the robot conducts the exercise sessions, evaluates user performance, and gives the user real-time feedback completely autonomously, without human operator intervention at any time during the exercise sessions.

B. Robot Platform

To address the role of the robot's physical embodiment, we used 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): six DOF arms (x2), one DOF gripping hands (x2), two DOF pan/tilt neck, one DOF expressive eyebrows, and a two DOF expressive mouth. The robot is shown in Fig. 2.

A standard USB camera is located at the waist of the robot, and used to capture the user's arm movements during the exercise interaction, allowing the robot to provide appropriate performance feedback to the user.

The robot's speech is generated by the commercially available NeoSpeech text-to-speech engine [49] and a speaker on the robot outputs the synthesized voice to the

Fig. 2. ***Robot platform used in the experiments.***

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.

C. Exercise Games

Three exercise games are available in our system: the Workout game, the Imitation game, and the Memory game. During an exercise session, the user is given the opportunity to play all three games, and often each more than once within the session duration. The following is a description of each game in detail.

1) Workout Game: In this game, the robot fills the role of a traditional exercise instructor by demonstrating the arm exercises with its own 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"). In monitoring user performance, the robot compares the user's current arm angles as detected by the vision module to those of the specified goal arm angles to determine performance accuracy. The comparison procedure is robust to user fatigue and variations in range of motion; it relies more on the user's current hand positions and forearm angles than on the absolute differences between the user's arm angles and the target angles.

2) Imitation Game: In this game, the roles of the user and the robot from the Workout game are reversed; 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.

As 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 and general commentary. For example, if the robot detects that the user is not moving, it encourages the user to create new gestures by saying, for instance, "Mary, try and come up with your own gestures and I'll imitate you." In addition, the robot makes general comments about the game or the user, such as "You're a good instructor, Mary" or "This is my favorite game, thanks for the workout."

3) Memory Game: In this game, the user is challenged to learn a sequence of different arm gestures. The goal of the game is for the user to try and memorize ever-longer sequences, 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 any skill level.

The robot starts out by showing the first two positions of the sequence and asks the user to perform them while it provides feedback. Once the user has successfully repeated the first two gestures with the help of the robot, the user is asked to repeat the sequence again from the beginning, this time without demonstration or verbal feedback from the robot. Once the gestures are completed without help, the robot shows the next two gestures in the sequence, and the user is again asked to perform the entire sequence from the beginning (now four gestures in length). As the user continues to successfully memorize all shown gestures, the robot continues to show the user two more (six, then eight gestures, and so on), and the game progresses in difficulty.

The robot helps the user to keep track of the sequence by counting along with each correct gesture, and reminding the user of the poses when it detects errors (e.g., "Oh, that's too bad! Here is gesture five again"). The robot also reports to the user his/her current high score (i.e., the number of gestures remembered correctly) in an attempt to motivate improvement upon past performance.

D. Vision Module

In order to monitor user performance and provide accurate feedback during the exercise routines, the robot must be able to recognize the user's arm gestures. To accomplish this, we developed a vision module that recognizes the user's arm gestures/poses in real time, with minimal requirements for the surrounding environment and none for the user.

Several different approaches have been developed to accomplish tracking of human motion, both in 2-D and 3-D, including skeletonization methods [44], [45], gesture recognition using probabilistic methods [46], and color-based tracking [47], among others. We opted to create an arm pose recognition system that takes advantage of our simplified exercise setup in order to achieve real-time results without imposing any markers on the user.

To simplify visual recognition of the user, a black curtain was used to provide a static and contrasting background for fast segmentation of the user's head and hands, the most important features of the arm pose recognition task, independent of the user's skin tone.

The arm pose recognition algorithm works by first segmenting the original grayscale camera frame into a black and white image by applying a single threshold over the image; white pixels are assumed to form part of the user's body. The algorithm determines the final arm angles after localizing the user's hand and elbow locations using a heuristic procedure which takes as input the extrema points of the segmentation image. Example detection results are displayed in Fig. 3. Additional details regarding the visual recognition procedure can be found in [28].

The development of the SAR exercise system and visual recognition procedure predated the availability of the Microsoft Kinect [50]. Future implementations of the system will utilize Kinect-type 3-D vision technology and do away with the curtain and the planar limits of the motions. Nevertheless, the 2-D nature of the exercises was not noted as an issue by any of the participants in our user studies.

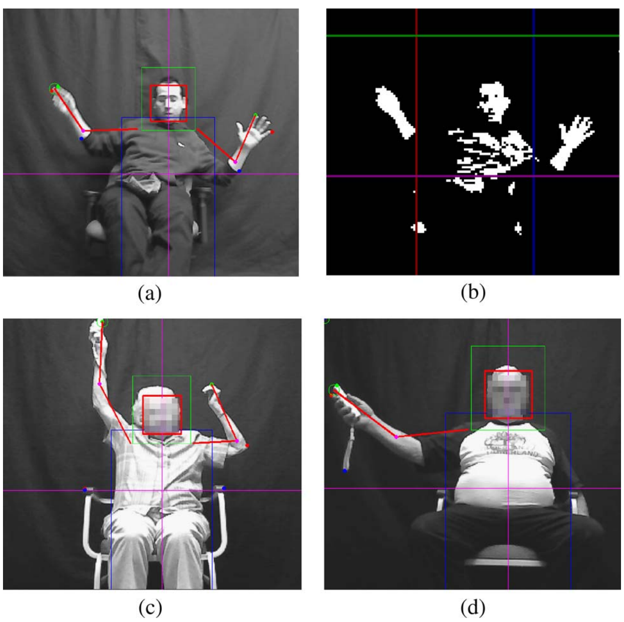

Fig. 3. *(a), (c), and (d) Example face and arm angle detection results superimposed over original grayscale camera frames. (b) Segmented image of camera frame shown in (a).*

V. MOTIVATION STUDY I: PRAISE AND RELATIONAL DISCOURSE EFFECTS

We designed and conducted an intrinsic motivation study to investigate the role of praise and relational discourse (politeness, humor, empathy, etc.) in the robot exercise system. Toward that end, the study compared the effectiveness and participant evaluations of two different coaching styles used by our system to motivate elderly users to engage in physical exercise. This section discusses the study methods employed, the subjective and objective measures that were evaluated, and the outcomes of the study and system evaluation with elderly participants.

A. Study Design

The study consisted of two conditions, relational and nonrelational, to explore the effects of praise and communicative relationship-building techniques on a user's intrinsic motivation to engage in the exercise task with the SAR coach. The study design was within subject; participants saw both conditions, one after the other, and the order of appearance of the conditions was counterbalanced among the participants. Each condition lasted 10 min, totaling 20 min of interaction, with surveys being administered after both sessions to capture participant perceptions of each study condition independently. The following describes the two conditions in greater detail.

1) Relational Condition: In this condition, the SAR exercise coach employs all of the social interaction and personalization approaches described in Section III. Specifically, the robot always gives the user praise upon correct completion of a given exercise gesture (an example of positive feedback) and provides reassurance in the case of failure (an example of empathy). The robot also displays continuity behaviors (e.g., by referencing past experiences with the user), humor, and refers to the user by name, all with the purpose of encouraging an increase in the user's intrinsic motivation to engage in the exercise session.

2) Nonrelational Condition: In this condition, the SAR exercise coach guides the exercise session by providing instructional feedback as needed (e.g., user score, demonstration of gestures, verbal feedback during gesture attempts, etc.), but does not employ explicit relationship building discourse of any kind. Specifically, the robot does not provide positive feedback (e.g., praise) in the case of successful user completion of an exercise gesture, nor does it demonstrate empathy (e.g., reassurance) in the case of user failure. The SAR coach also does not display continuity behaviors, humor, or refer to the user by name. This condition represents the baseline condition of our SAR exercise system, wherein the robot coach does not employ any explicit motivational techniques to encourage an increase in the user's intrinsic motivation to engage in the task.

B. Participant Statistics

We recruited elderly individuals to participate in the study through a partnership with be.group, an organization of senior living communities in Southern California, using flyers and word-of-mouth. Thirteen participants responded and successfully completed both conditions of the study. The sample population consisted of 12 female participants (92%) and one male participant (8%). Participants' ages ranged from 77 to 92, and the average age was 83 (S.D. = 5.28). Half of the participants ($n = 7$) engaged in the relational condition in the first session, whereas the other half ($n = 6$) engaged first in the nonrelational condition.

C. Measures

Survey data were collected at the end of the first and second sessions in order to analyze participant evaluations of the robot and of the interaction with the exercise system in both conditions. The same evaluation surveys were used for each session to allow for objective comparison between the two conditions.

In addition to these evaluation measures, at the end of the last exercise session, we administered one final survey asking the participants to directly compare the two study conditions (labeled "first" and "second") according to ten evaluation categories. This survey allowed us to obtain a general sense of the participants' preferences regarding the different SAR approaches and hence gauge their respective motivational capabilities.

Objective measures were also collected to evaluate user performance and compliance in the exercise task.

The following describes the specific evaluation measures captured in the postsession surveys, and the objective measures captured during the exercise sessions.

1) Evaluation of Interaction: Two dependent measures were used to evaluate the interaction with the robot exercise system. The first measure was the *enjoyableness of the interaction*, collected from participant assessments of the interaction according to six adjectives: enjoyable; interesting; fun; satisfying; entertaining; boring; and exciting (Cronbach's $\alpha = 0.93$). Participants were asked to rate how well each adjective described the interaction on a ten-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 enjoyableness of the interaction was measured to gain insight into the user's motivation level to engage in the task, because, as Csikszentmihalyi states, intrinsically motivating activities are characterized by enjoyment [37].

The second measure was the perceived *value or usefulness of the interaction*. Participants were asked to evaluate how well each of the following four adjectives described the interaction: useful; beneficial; valuable; and helpful (Cronbach's $\alpha = 0.95$). The same ten-point scale anchored by "describes very poorly" (1) and "describes very well" (10) was used in the evaluation. The perceived usefulness of the system was measured to estimate user acceptance and trust of the system in helping to achieve the desired health goals, which is necessary for the system to be successful in the long term.

2) Evaluation of Robot: The *companionship of the robot* was measured based on participant responses to nine ten-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 $\alpha = 0.86$). These questions were derived from the Companion Animal Bonding Scale of Poresky *et al.* [53]. The companionship of the robot was measured to assess potential user acceptance of the robot as an in-home companion, thereby demonstrating the capability of the system toward uses in independent living/aging-in-place.

To assess the perceptions of the capabilities of the system in motivating exercise, we measured participant evaluations of the robot as an *exercise coach*. Participant evaluations of the robot as an exercise coach were gathered from a combination of the participants' reported level of agreement toward two coaching-related statements, and responses to three additional questions. The two statements and three questions were, respectively: I think Bandit is a good exercise coach; I think Bandit is a good motivator of exercise; 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 $\alpha = 0.88$). The two statements were rated on a ten-point scale anchored by "very strongly disagree" (1) and "very strongly agree" (10), and the three question items were each measured according to a ten-point scale anchored by "not at all" (1) and "very much" (10).

To quantify the effectiveness of the robot's social capabilities, we measured the *social presence of the robot*. Social presence is defined as the feeling that mediates how people respond to social agents; it strongly influences the relative success of a social interaction [20]. In essence, the greater the social presence of the robot, the more likely the interaction is to be successful. The social presence of the robot was measured by a ten-point scale anchored by "not at all" (1) and "very much" (10) using questionnaire items established 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 $\alpha = 0.82$).

3) Direct Comparison of Conditions: The ten evaluation categories assessed by the direct comparison survey, which asked participants to choose between the first or second exercise sessions, were as follows: *enjoy more; more useful; better at motivating exercise; prefer to exercise with; more frustrating; more boring; more interesting; more intelligent; more entertaining; choice from now on*. Analysis of the direct-comparison data serves primarily to support and confirm the results obtained from the within-subjects analysis of the dependent measures across study conditions.

4) User Performance Measures: To help assess the effectiveness of the SAR exercise system in motivating exercise among the participants, we collected nine different objective measures during the exercise sessions regarding user performance and compliance in the exercise task. Most of the objective measures were captured during the Workout game, wherein the robot guides the interaction similar to a traditional exercise coach. These measures include the *average time to gesture completion* (from the moment the robot demonstrates the gesture, to successful user completion of the gesture), *number of seconds per exercise completed, number of failed exercises, number of movement prompts* by the robot to the user due to lack of arm movement, and *feedback percentage*. The feedback percentage measure refers to the fraction of gestures, out of the total given, where the robot needed to provide verbal feedback to the user regarding arm positions in order to help guide the user to correct gesture completion.

We also recorded the *maximum score* over all sessions, *average maximum score* among users, and *average time per gesture attempt* in the Memory game. For the Imitation game, the only measure captured was again the *number of movement prompts* by the robot due to lack of user arm movement.

D. Hypotheses

Based on the related research on the positive effects of praise and relational discourse on intrinsic motivation, discussed in Section III, seven hypotheses were established for this study.

- *Hypothesis 1*: Participants will evaluate the enjoyableness of their interaction with the relational robot more positively than their interaction with the nonrelational robot.
- *Hypothesis 2*: Participants will evaluate the usefulness of their interaction with the relational robot more positively than their interaction with the nonrelational robot.
- *Hypothesis 3*: Participants will evaluate the companionship of the relational robot more positively than that of the nonrelational robot.
- *Hypothesis 4*: Participants will evaluate the relational robot more positively as an exercise coach than the nonrelational robot.
- *Hypothesis 5*: There will be no significant difference between participant evaluations of the social presence of the relational robot and nonrelational robot. The reasoning behind this hypothesis is that people's sense of social presence is largely determined by the embodiment type and perceived intelligence of the social agent, which is assumed to be more or less equal in the two robot conditions.
- *Hypothesis 6*: Participants will report a clear preference for the relational robot over the nonrelational robot when asked to compare both exercise sessions directly.
- *Hypothesis 7*: There will be no significant difference in participant exercise performance when interacting with either the relational or nonrelational robot. This hypothesis is based on the assumption that, due to the short-term nature of the study and novelty of the system, performance measures will be approximately equal between robot conditions.

E. Results

1) Evaluation of Interaction Results: Participants who engaged with the relational robot in their first session, rated the nonrelational condition on average 22% lower than the relational condition in terms of enjoyment ($M_R = 7.5$ versus $M_{NR} = 5.9$), and 23% lower in terms of usefulness ($M_R = 7.5$ versus $M_{NR} = 5.8$). Similarly, the participants who instead engaged with the nonrelational robot in their first session also expressed a greater preference for interacting with the relational robot by rating the relational condition on average 10% higher than the nonrelational condition in terms of enjoyment ($M_{NR} = 7.6$ versus $M_R = 8.4$), and 7% higher in terms of usefulness ($M_{NR} = 7.5$ versus $M_R = 8.0$).

Altogether, 85% of the participants (11 of 13) rated the relational condition higher than the nonrelational condition in terms of enjoyment, and 77% of the participants (10 of 13) rated the relational condition higher in terms of usefulness than the nonrelational condition.

To test for significant differences among the participant evaluations of the study conditions, we performed a Wilcoxon signed-rank test on the data to analyze matched pairs from the sample population's evaluations of both study conditions according to the dependent measures. Supporting Hypothesis 1, the results show that the participants evaluated the interaction with the relational robot as significantly more enjoyable/entertaining than the interaction with the nonrelational robot ($W[12] = 4$, $p < .005$), and as somewhat more valuable/useful than the interaction with the nonrelational robot, although not to a significant degree ($W[12] = 15.5$, $p < 0.10$), hence Hypothesis 2 was not supported by the data. For illustration purposes, Fig. 4(a) shows the average participant ratings of

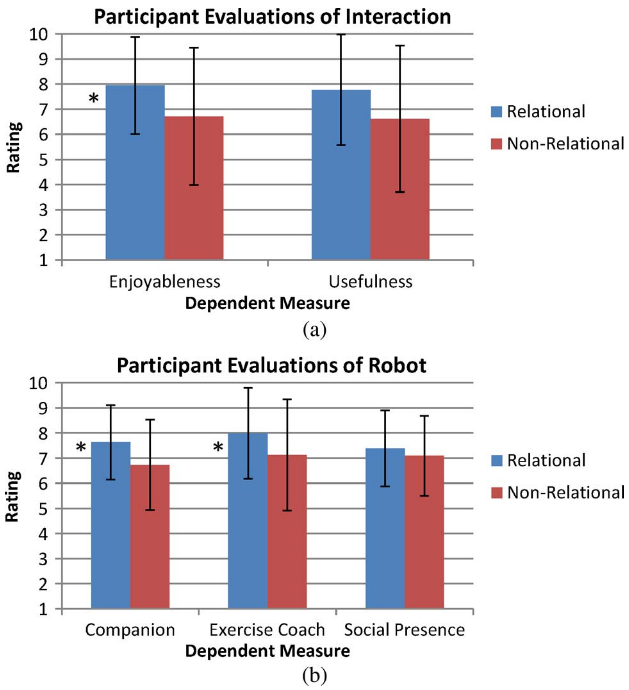

Fig. 4. ***(a) Plot of participant evaluations of the interaction, in terms of enjoyableness and usefulness, for both study conditions; (b) plot of participant evaluations of the robot (as a companion, exercise coach, and level of social presence) for both study conditions. Note: significant differences are marked by asterisks (*).***

the enjoyableness and usefulness of the interaction for both study conditions.

2) Evaluation of Robot Results: Participants who engaged in the relational condition in their first session rated the nonrelational robot on average 11% lower than the relational robot in terms of companionship ($M_R = 7.4$ versus $M_{NR} = 6.5$), 11% lower as an exercise coach ($M_R = 7.7$ versus $M_{NR} = 6.9$), and 1% lower in terms of social presence ($M_R = 7.2$ versus $M_{NR} = 7.1$). Greater positive scores for the relational robot were also reported by the participants who instead engaged first in the nonrelational condition, having rated the relational robot on average 14% higher than the nonrelational robot in terms of companionship ($M_{NR} = 6.9$ versus $M_R = 7.9$), 10% higher as an exercise coach ($M_{NR} = 7.4$ versus $M_R = 8.2$), and 8% higher in terms of social presence ($M_{NR} = 6.9$ versus $M_R = 7.5$).

Altogether, 77% of the participants (10 of 13) rated the relational robot higher than the nonrelational robot in terms of companionship, 77% of the participants (10 of 13) rated the relational robot more positively as an exercise coach, and the comparative ratings of social presence between the robot conditions were approximately equal, as 54% of participants (7 of 13) reported higher social presence for the relational robot.

We again analyzed the data to test for significant differences among participant evaluations across the two robot conditions by performing a Wilcoxon signed-rank test. The results show that the participants rated the relational robot as a significantly better companion than the nonrelational robot ($W[13] = 14, p < .05$), supporting Hypothesis 3, and as a significantly better exercise coach than the nonrelational robot ($W[11] = 7, p < .02$), in support of Hypothesis 4. As expected, there was no significant difference in the participant evaluations of social presence between both robot conditions ($W[12] = 28.5, p > 0.2$), confirming Hypothesis 5, with both robots receiving equally high ratings. The average participant ratings of both robot conditions for all three dependent measures are shown in Fig. 4(b).

3) Direct Comparison Results: At the end of the final exercise session, participants were asked to directly compare both robot conditions with respect to ten different evaluation categories; results are provided in Table 1. It is important to note that the study conditions were labeled as "first session" and "second session" on the survey. These labels would correspond to either the relational condition or nonrelational condition, depending on the order of the conditions in which each participant engaged, and were chosen to avoid any potential bias in the survey items.

The results support Hypothesis 6 by demonstrating that, regardless of the order of condition presentation, the participants expressed a strong preference for the relational robot over the nonrelational robot. Specifically, the relational robot received 82% of the positive trait votes versus 16% for the nonrelational robot, with the remaining 2% shared equally between them. Other notable results include the high number of participants who rated the relational robot as more enjoyable (10 votes, 77%), better at motivating exercise (11 votes, 85%), more useful (11 votes, 85%), and the robot they would choose to exercise with in the future (11 votes, 85%). In contrast, the nonrelational robot received a high number of votes for being more

Table 1 Participant Responses to Direct Comparison Survey Items

	Relational	*Non-Relational*	*Both Equal*
Enjoy More	10 (77%)	3 (23%)	0 (0%)
More Intelligent	11 (85%)	2 (15%)	0 (0%)
More Useful	11 (85%)	2 (15%)	0 (0%)
Prefer to Exercise with	11 (85%)	2 (15%)	0 (0%)
Better at Motivating	11 (85%)	2 (15%)	0 (0%)
More Frustrating	3 (23%)	10 (77%)	0 (0%)
More Boring	2 (15%)	10 (77%)	1 (8%)
More Interesting	10 (77%)	2 (15%)	1 (8%)
More Entertaining	10 (77%)	2 (15%)	1 (8%)
Choice from now on	11 (85%)	2 (15%)	0 (0%)

Table 2 Participant Exercise Performance Statistics

Objective Measure	*Avg.(std.)*
Time to Gesture Completion (seconds)	2.46 (0.70)
Seconds per Exercise	5.21 (1.00)
Feedback Percentage	7.4% (4.8%)
Number of Failed Gestures	0
Number of Movement Prompts $_W$	0
Maximum Score	6
Average Maximum Score	3.08 (1.12)
Time per Gesture Attempt (seconds)	8.57 (4.11)
Number of Movement Prompts $_I$	0.26 (0.53)

frustrating (10 votes, 77%) and more boring (10 votes, 77%) than the relational robot.

4) User Exercise Performance Statistics: The collected statistics regarding participant performance in the exercise task were very encouraging as they demonstrated a consistently high level of user exercise performance and compliance with the exercise task. As expected, and in support of Hypothesis 7, there were no significant differences found in participant performance between the two study conditions, with both conditions reporting equally high performance among the participants. For example, the average gesture completion time for participants in the relational condition was 2.45 s (S.D. $= 0.65$), compared to 2.46 s (S.D. $= 0.78$) for participants in the nonrelational condition ($W[13] = 37, p > 0.2$). Given the lack of significant difference in user performance between the two conditions, the statistics presented in this section refer to the participant performance across all exercise sessions of the study.

User compliance and performance in the Workout game were high. The average gesture completion time was 2.46 s (S.D. $= 0.70$), and the overall exercise performance averaged 5.21 s per exercise (S.D. $= 1.0$), which also includes time taken for verbal praise, feedback, and score reporting from the robot. The low percentage of necessary corrective feedback, averaging 7.4%, zero failures, and zero movement prompts during the entire study, are all very encouraging results, as they suggest that the participants were consistently motivated to do well on the exercises throughout the interaction.

A summary of all statistics regarding user performance, including those from the Memory and Imitation games, can be found in Table 2.

F. Discussion

The results of the study show a strong user preference for the relational robot over the nonrelational robot, demonstrating the positive effects of praise and relational discourse in a healthcare task-oriented HRI scenario, and supporting all of our hypotheses with the exception of Hypothesis 2, which missed reaching significance by a small margin. Participants rated the relational robot significantly higher than the nonrelational robot in terms of enjoyableness, companionship, and as an exercise coach. Comments made by participants after the study further illustrate the positive response to the relational robot, including "It's nice to hear your name, it's personal. I felt more positive reinforcement," and from another participant "The robot encourages you, compliments you; that goes a long way." These results provide significant insight into how people respond to SARs, and confirm the positive influence that praise and relational discourse have on intrinsic motivation. These are of particular importance for the healthcare domain, where effectiveness in social interaction, relationship building, and gaining user acceptance and trust are all necessary in ultimately achieving the desired health outcomes of the therapeutic interventions.

The effectiveness of the SAR exercise system was also demonstrated by the outcomes of the study. Not only did the participants rate the interaction with our robot coach as highly enjoyable/entertaining, suggesting they were intrinsically motivated to engage in the exercise task, but they also consistently engaged in physical exercise throughout the interaction, as demonstrated by the gathered user performance statistics. These results are very encouraging, as they clearly show that the system was successful in motivating elderly users to engage in physical exercise, thereby confirming its effectiveness and achieving the primary goal of the system.

VI. MOTIVATION STUDY II: USER CHOICE AND SELF-DETERMINATION

As discussed in Section III, allowing the user to gain a sense of self-determination within a task, for example from choice of activity, has been shown to increase or be less detrimental to intrinsic motivation when compared to similar task conditions that do not involve choice [41], [42]. To investigate the role of choice and user autonomy in influencing user intrinsic motivation in the robot exercise system, as well as to further test and validate the effectiveness of our system, we conducted a second user study with elderly participants.

A. Study Design

The study consisted of two conditions, choice and no choice, designed to test user preferences regarding choice of activity. The conditions differed only in the manner in which the three exercise games (Workout, Imitation, Memory) were chosen during the exercise sessions. As in the first study, the design was within subject; each participant engaged in both conditions one after the other, with the order of appearance counterbalanced among the participants. Each condition lasted 10 min, totaling 20 min of interaction. The following are descriptions of each condition in greater detail.

1) Choice Condition: In this condition, the user is given the choice of which game to play at specific points in the interaction. The robot prompts the user to press the "Yes" button upon hearing the desired game, and then calls out the names of each of the three game choices. After the user has made a choice, the chosen game is played for a duration ranging from 1 to 2 min in length. Then, the robot asks the user if he would like to play a different game. Depending on the user's response, the robot either continues playing the same game for another 1–2 minutes, or prompts the user again to choose the game to play next.

2) No Choice Condition: In this condition, the robot chooses which of the three games to play at the specified game change intervals (every 1–2 minutes). The robot always changes games, to try to minimize any user frustration, as in this condition the robot is unaware of the user's game preferences. For simplicity, in this condition, the robot always chooses to first play the Workout game, followed by the Imitation and then Memory games, then cycles through them again in the same order.

B. Participant Statistics

We recruited elderly individuals to participate in the study again through our partnership with the be.group senior living organization. Eleven individuals participated in the first trial of the study, which was subsequently expanded to include thirteen additional participants. Therefore, a total of 24 participants were recruited and successfully completed both conditions of the study. Half of the participants engaged in the choice condition in their first session, whereas the other half engaged first in the no choice condition. The sample population consisted of 19 female participants (79%) and five male participants (21%). Participants' ages ranged from 68 to 89, and the average age was 77 (S.D. = 5.76).

C. Measures

As in the first study, survey data were collected at the end of the first and second sessions in order to analyze participant evaluations of the interaction with the exercise system in both conditions. The same evaluation surveys were used for each session to allow for objective comparison between the two conditions.

We administered an additional questionnaire at the end of the last session, asking the participants about their preferences regarding choice in the exercise system, in addition to various other opinion items for further evaluation of the exercise system.

The following describes the specific evaluation measures captured in the postsession surveys.

1) Evaluation of Interaction: The two dependent measures used to evaluate the interaction with the robot exercise system were the same as in the previous study, namely the *enjoyableness of the interaction*, and the perceived *value or usefulness of the interaction*. The ratings scales and survey items for each measure also remained the same.

2) User Preferences Regarding Choice: Three questionnaire items were used to assess participant preferences and opinions regarding choice in the exercise system (direct user input in choosing the exercise games). The first item asked participants to state their *session preference*, labeled as "first" or "second," which referred to either the choice or no choice conditions, depending on each participant's session ordering. The ordinal labels were again chosen, as in the previous study, to avoid any bias in the survey item. The second item asked the participants about *user choice*, specifically whether they preferred to choose the exercise games to be played, or whether they preferred to let the robot choose instead. This question is similar to the first item but in more direct terms. Last, the third item asked the participants about *added enjoyment due to user choice*, specifically asking whether having the ability to choose which game to play added to their enjoyment of the interaction.

3) Evaluation of SAR System: The last seven questionnaire items were used to obtain additional feedback on the user perceptions of and feelings toward the SAR exercise system. The first four of these items asked participants to rate, respectively: their perception of the *robot's intelligence*, their perception of the *robot's helpfulness*, the *level of importance* they put on their participation in the exercise sessions with the robot, and their *mood* in general during the exercise sessions. The rating scales were five-point Likert scales, anchored by "not at all" (1) and "very" (5) (e.g., "not at all intelligent" and "very intelligent"). The question regarding user mood during the sessions contained a modified scale, where the mood options ranged from "irritated/frustrated" (1) to "happy/joyful" (5), with the medium range being "normal" (3). Participants were also asked to report their *favorite game, least favorite game*, and to state their choice of the *robot description* that best fit among four available options: companion, exercise instructor, game conductor, none of these.

D. Hypotheses

Based on the related research on the positive effects of user choice and autonomy on intrinsic motivation, discussed in Section III, five hypotheses were established for this study.

- *Hypothesis 1*: Participants will evaluate the enjoyableness of their interaction in the choice condition more positively than their interaction in the no choice condition.
- *Hypothesis 2*: Participants will evaluate the usefulness of their interaction in the choice condition more positively than their interaction in the no choice condition.

- *Hypothesis 3*: Participants will report a clear preference for the choice condition over the no choice condition when asked to compare both exercise sessions directly.
- *Hypothesis 4*: Participants will report a clear preference for choosing the exercise games themselves, as opposed to having the robot choose which games to play during the interaction.
- *Hypothesis 5*: Participants will report feeling an increase in the enjoyment of the exercise task when given the opportunity to choose which games to play during the interaction.

E. Results

1) Evaluation of Interaction Results: The evaluation of interaction survey items was introduced after the first trial of the study, therefore the results presented here for the two interaction measures were analyzed from the data gathered solely from the 13 participants of the expanded study. Nevertheless, all other survey results presented were gathered from all 24 participants of the study.

Participants who engaged in the choice condition in the first session rated the no choice condition on average 7% higher than the choice condition in terms of enjoyment ($M_C = 7.5$ versus $M_{NC} = 8.0$), and 2% lower in terms of usefulness ($M_C = 8.7$ versus $M_{NC} = 8.5$). In slight contrast, the participants who instead engaged in the no choice condition in their first session rated the choice condition on average 4% higher than the no choice condition in terms of enjoyment ($M_{NC} = 8.5$ versus $M_C = 8.8$), and 5% higher in terms of usefulness ($M_{NC} = 9.0$ versus $M_C = 9.5$).

Altogether, there was no clear participant preference for one condition over the other, as 62% of the participants (8 of 13) rated the no choice condition higher than the choice condition in terms of enjoyment, and 62% of the participants (8 of 13) rated the choice condition higher in terms of usefulness than the no choice condition.

We performed a Wilcoxon signed-rank test on the data and found no significant differences between participant evaluations of the two study conditions, neither with respect to the enjoyableness ($W[13] = 28.5, p > 0.2$), nor the usefulness of the interaction ($W[13] = 30.5, p > 0.2$). Thus, Hypotheses 1 and 2 were not supported by the data. Nevertheless, participant ratings for the enjoyableness ($M = 8.18$, S.D. $= 1.67$) and usefulness ($M = 8.95$, S.D. $= 1.63$) of the interaction across both conditions were very positive, with scores even higher than those seen in the previous study. These high evaluations of the SAR exercise system further illustrate the effectiveness of the system in instructing and motivating elderly users to exercise.

2) User Preferences Regarding Choice Results: The survey results regarding session preference indicated that 42% of the participants (10 of 24) preferred the no choice condition, 33% of the participants (8 of 24) preferred the choice condition, and 25% of the participants (6 of 24) expressed no preference for one condition over the other. Fig. 5(a) plots the participants' stated preferences of study conditions. The varied participant condition preferences indicate no clear preference for one over the other, and thus Hypotheses 3 was not supported. Concerning user choice in the exercise system, 62% of participants (15 of 24) reported preferring to let the robot choose the games to play, with the remaining 38% of participants (9 of 24) preferring to choose the games themselves. The slight preference among participants for having the robot choose countered the reasoning of Hypothesis 4, which was not supported.

It is interesting to note that even though most participants preferred letting the robot decide which games to play, almost all of the participants, 92% (22 of 24),

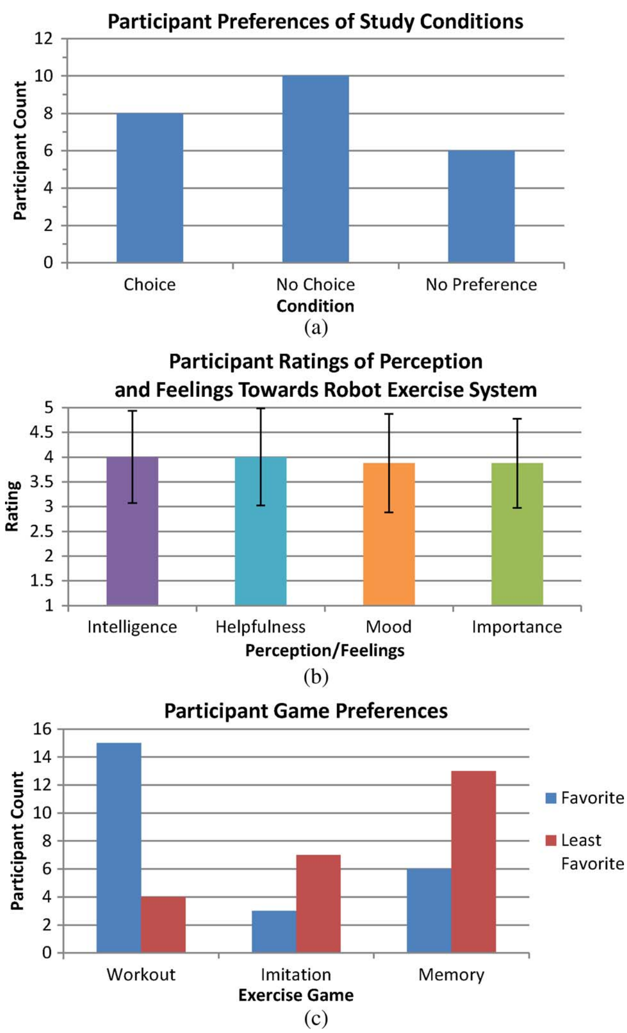

Fig. 5. ***Graphs of: (a) the participants' preferences of study condition; (b) the participants' ratings in response to survey questions on their perception of the robot's intelligence, helpfulness, their mood during sessions, and how important the sessions were to them; and (c) the participants' preferences of exercise game.***

reported increased enjoyment of the task when given the opportunity to choose the exercise game to play. This result supports Hypothesis 5 and is consistent with the literature on the effects of user choice on intrinsic motivation [41], [42].

3) Evaluation of SAR System Results: The results of the survey questions regarding participant perceptions and feelings toward the SAR exercise system are very encouraging; the participants rated the robot highly in terms of intelligence ($M = 4.0, \text{S.D.} = 0.93$) and helpfulness ($M = 4.0, \text{S.D.} = 0.97$), attributed a moderately high level of importance to the exercise sessions ($M = 3.87$, $\text{S.D.} = 0.89$), and reported their mood throughout the sessions to be normal-to-moderately pleased ($M = 3.87$, $\text{S.D.} = 0.99$). These results are important because positive user perceptions of the agent's intelligence and helpfulness are a key part of establishing trust in the human–robot relationship. This, along with positive user mood and user-attributed importance to the therapeutic task, are in turn important for establishing and maintaining user intrinsic motivation. These are all key components for achieving long-term success in any SAR setting. An illustration of the results is shown in Fig. 5(b).

Regarding the exercise games, the participants largely favored (62%, 15 of 24) the Workout game over the others, wherein the robot serves as a traditional exercise coach, with the Memory game being chosen most often as the participants' least favorite game (54%, 13 of 24). Fig. 5(c) summarizes the participants' game preferences.

The description most chosen by the participants as the best fit for the robot was that of an exercise instructor (67%, 16 of 24), not surprisingly, as opposed to that of a game conductor (25%, 6 of 24) or companion (8%, 2 of 24). While all of the descriptions represent characteristics of the robot in one form or another, the primary selection of an exercise coach by the participants illustrates the perception of the robot as an agent that they can trust and that is capable of helping, rather than simply entertaining.

F. Discussion

The results of the study showed no clear preference for one condition over the other, as the user enjoyment level of the interaction was reported to be equally high for both conditions, with or without user choice of activity. The high participant evaluations regarding the enjoyableness and usefulness of the interaction, the intelligence and helpfulness of the robot, and positive user mood and attributed importance to the exercise sessions, further validate the SAR system's effectiveness in motivating elderly users to engage in physical exercise.

The relatively mixed condition preferences among participants, or rather the lack of clear preference for the choice condition, seem somewhat counterintuitive given the positive effect that choice and user autonomy have been shown to have on task-based enjoyment [41], [42]. One possible explanation for the mixed preferences may be that, since the robot's role in the interaction was that of an exercise instructor, some participants might have felt it was the robot's duty to determine the exercise regimen, and hence were comfortable relinquishing the choice of exercise games. Another possible explanation may be that the enjoyment derived from choosing the games did not outweigh the enjoyment derived from relaxation due to the reduced responsibility of not having to choose the games. Both explanations seem plausible, as some of the participants reported preferring the robot to have the "responsibility" of steering the task. A third explanation may be that, given the short-term nature of the study, some participants may have needed more experience with the robot system before they felt confident enough to make task-based decisions themselves.

It is interesting to note that, even though the condition preferences were varied and nearly half of participants preferred letting the robot decide which games to play, all participants at one point or another during the study took advantage of having greater control in the choice condition. Specifically, when given the option by the robot to change games, all participants at some point either chose to continue playing the same game they were playing, or chose to avoid playing a game they did not want to play. Neither of these cases could occur in the no choice condition, as the robot was unaware of the user's current game preferences.

This observation speaks to the value of user preference within the task scenario, suggesting that a hybrid approach that includes both user and robot decision making, personalized and tuned automatically for each user, might ultimately be the best solution for achieving a fluid and enjoyable task interaction for all users. For example, for users who prefer greater robot responsibility and input in the SAR-based task, the robot can recommend the "best" choices given the current task conditions and situation, giving the user an informed choice. Alternatively, for users who prefer greater control only once they've gained enough experience, the robot can initially make all task-based decisions until the user is ready and confident in making choices. For users who have a clear preference regarding who should make task-based decisions during interaction, the chosen strategy can be implemented continually throughout the sessions. Clearly, no single fixed user-choice strategy is appropriate for all users; users have varied preferences regarding choice, and those preferences may even change over time. Therefore, it is important that the strategy employed in SAR systems regarding user choice and autonomy be continually adapted to the specific user engaged in the interaction, thus personalizing the therapeutic intervention.

VII. CONCLUSION

In this paper, we have presented the design methodology, implementation, and evaluation of a SAR that is capable of interacting with elderly users and engaging them in physical exercise in a seated aerobic exercise scenario. Methods for influencing an individual's intrinsic motivation to perform a task, including verbal praise, relational discourse, and user choice, were implemented and evaluated in two separate user studies with elderly participants.

The results of the first motivation study showed a strong participant preference for the relational robot over the nonrelational robot in terms of enjoyableness of the interaction, companionship, and as an exercise coach, in addition to demonstrating similar evaluations of both robots in terms of usefulness of interaction and social presence. These results illustrate the positive effects of motivational relationship building techniques, namely praise and relational discourse, on participant perceptions of the social agent and interaction in a health-related task scenario, and ultimately on user intrinsic motivation to engage in the task. The results of the second motivation study showed varying participant preferences regarding user choice within the exercise system, suggesting the need for customizable interactions automatically tailored to accommodate the personal preferences of the individual users.

The SAR exercise system was very well received, as demonstrated by both user studies, with high participant evaluations regarding the enjoyableness and usefulness of the interaction, companionship, social presence, intelligence, and helpfulness of the robot coach, and the positive mood and attributed importance of the exercise sessions. The system was also found to be effective in motivating consistent physical exercise throughout the interaction, according to various objective measures, including average gesture completion time, seconds per exercise, and feedback percentage.

The overall acceptance of the SAR exercise system by elderly users, as evidenced by the outcomes of two user studies evaluating the motivation capabilities and effectiveness of the system, is very encouraging and illustrates the potential of the system to help the elderly population to engage in physical exercise to achieve beneficial health outcomes, to facilitate independent living, and ultimately to improve quality of life. ■

Acknowledgment

The authors would like to thank be.group (formerly Southern California Presbyterian Homes), its staff, and participating residents, with special thanks to A. Mehta, K. Reed, A. Escobedo, and M. Anderson for their help in the preparation and execution of the robot exercise system studies.

REFERENCES

[1] U.S. Census Bureau, U.S. Census Bureau Report (Issue Brief No. CB09-97), 2009. [Online]. Available: http://www.census.gov/population/international/

[2] American Association of Colleges of Nursing, Nursing Shortage Fact Sheet, 2010. [Online]. Available: http://www.aacn.nche.edu/media/FactSheets/NursingShortage.htm

[3] American Health Care Association, Summary of 2007 AHCA Survey Nursing Staff Vacancy and Turnover in Nursing Facilities, Jul. 2008.

[4] P. Buerhaus, "Current and future state of the US nursing workforce," *J. Amer. Med. Assoc.*, vol. 300, no. 20, pp. 2422–2424, 2008.

[5] S. Colcombe and A. Kramer, "Fitness effects on the cognitive function of older adults," *Psychol. Sci.*, vol. 14, pp. 125–130, 2003.

[6] S. J. Colcombe, A. F. Kramer, K. I. Erickson, P. Scalf, E. McAuley, N. J. Cohen, A. Webb, G. J. Jerome, D. X. Marquez, and S. Elavsky, "Cardiovascular fitness, cortical plasticity, and aging," in *Proc. Nat. Acad. Sci. USA*, 2004, vol. 101, pp. 3316–3321.

[7] W. Spirduso and P. Clifford, "Replication of age and physical activity effects on reaction and movement time," *J. Gerontol.*, vol. 33, pp. 26–30, 1978.

[8] E.E. Baum, D. Jarjoura, A. E. Polen, D. Faur, and G. Rutecki, "Effectiveness of a group exercise program in a long-term care facility: A randomized pilot trial," *J. Amer. Med. Directors Assoc.*, vol. 4, pp. 74–80, 2003.

[9] D. Dawe and R. Moore-Orr, "Low-intensity, range-of-motion exercise: Invaluable nursing care for elderly patients," *J. Adv. Nursing*, vol. 21, pp. 675–681, 1995.

[10] M. D. McMurdo and L. M. Rennie, "A controlled trial of exercise by residents of old people's homes," *Age and Ageing*, vol. 22, pp. 11–15, 1993.

[11] V. Thomas and P. Hageman, "Can neuromuscular strength and function in people with dementia be rehabilitated using resistance-exercise training? Results from a preliminary intervention study," *J. Gerontol. A, Biol. Sci. Med. Sci.*, vol. 58, pp. M746–M751, 2003.

[12] Z. B. Moak and A. Agrawal, "The association between perceived interpersonal social support and physical and mental health: Results from the national epidemiological survey on alcohol and related conditions," *J. Public Health*, vol. 32, pp. 191–201, 2010.

[13] L. K. George, D. G. Blazer, D. C. Hughes, and N. Fowler, "Social support and the outcome of major depression," *British J. Psychiatry*, vol. 154, pp. 478–485, 1989.

[14] E. Paykel, "Life events, social support and depression," *Acta Psychiatrica Scandinavica*, vol. 89, pp. 50–58, 1994.

[15] E. Stice, J. Ragan, and P. Randall, "Prospective relations between social support and depression: Differential direction of effects for parent and peer support?" *J. Abnormal Psychol.*, vol. 113, pp. 155–159, 2004.

[16] S. A. Stansfeld, G. S. Rael, J. Head, M. Shipley, and M. Marmot, "Social support and psychiatric sickness absence: A prospective study of British civil servants," *Psychol. Med.*, vol. 27, pp. 35–48, 1997.

[17] J. Wainer, D. J. Feil-Seifer, D. A. Shell, and M. J. Matarić, "Embodiment and Human-Robot Interaction: A task-based perspective," in *Proc. IEEE Int. Workshop Robot Human Interactive Commun.*, Jeju Island, Korea, Aug. 2007, pp. 872–877.

[18] A. Powers, S. Kiesler, S. Fussell, and C. Torrey, "Comparing a computer agent with a humanoid robot," in *Proc. ACM/IEEE Int. Conf. Human-Robot Interaction*, New York, 2007, pp. 145–152.

[19] C. Bartneck, "Interacting with an embodied emotional character," in *Proc. Int. Conf. Designing Pleasurable Products Interfaces*, New York, 2003, pp. 55–60.

[20] K. M. Lee, *Presence, Explicated. Communication Theory*, vol. 14, no. 1, pp. 27–50, 2004.

[21] Y. Jung and K. M. Lee, "Effects of physical embodiment on social presence of social robots," in *Proc. Presence*, 2004, pp. 80–87.

[22] A. Tapus, C. Tapus, and M. Matarić, "The use of socially assistive robots in the design of intelligent cognitive therapies for people with dementia," in *Proc. IEEE Int. Conf. Rehabil. Robot.*, Kyoto, Japan, 2009, pp. 924–929.

[23] M. J. Matarić, J. Eriksson, D. J. Feil-Seifer, and C. J. Winstein, "Socially assistive robotics for post-stroke rehabilitation," *J. Neuro Eng. Rehabil.*, vol. 4, no. 5, Feb. 2007.

[24] A. Tapus, C. Tapus, and M. J. Matarić, "User-robot personality matching and robot behavior adaptation for post-stroke rehabilitation therapy," *Intell. Service Robot.*, vol. 1, no. 2, pp. 169–183, Apr. 2008.

[25] D. J. Feil-Seifer and M. J. Matarić, "Towards the integration of socially assistive robots into the lives of children with ASD," presented at the 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," in *Proc. Int. Conf. Human-Robot Interaction*, Osaka, Japan, Mar. 2010, pp. 143–144Poster Paper.

[27] J. Fasola and M. J. Matarić, "Robot motivator: Increasing user enjoyment and performance on a physical/cognitive task," in *Proc. IEEE Int. Conf. Develop. Learn.*, Ann Arbor, MI, Aug. 2010, pp. 274–279.

[28] J. Fasola and M. J. Matarić, "Robot exercise instructor: A socially assistive robot system to

monitor and encourage physical exercise for the elderly," in *Proc. 19th IEEE Int. Symp. Robot Human Interactive Commun.*, Viareggio, Italy, Sep. 2010, pp. 416–421.

[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 *Proc. IEEE Int. Conf. Robot. Autom.*, Apr. 2000, vol. 1, pp. 570–576.

[30] M. Montemerlo, J. Pineau, N. Roy, S. Thrun, and V. Verma, "Experiences with a mobile robotic guide for the elderly," in *Proc. AAAI Nat. Conf. Artif. Intell.*, Edmonton, AB, Canada, 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 *Proc. Int. Conf. Intell. Robots Syst.*, Oct. 2002, vol. 2, pp. 1152–1157.

[32] C. Kidd, W. Taggart, and S. Turkle, "A sociable robot to encourage social interaction among the elderly," in *Proc. Int. Conf. Robot. Autom.*, Orlando, FL, May 2006, pp. 3972–3976.

[33] T. W. Bickmore and R. W. Picard, "Establishing and maintaining long-term human-computer relationships," *ACM Trans. Comput.-Human Interaction*, vol. 12, no. 2, pp. 293–327, Jun. 2005.

[34] B. French, D. Tyamagundlu, D. Siewiorek, A. Smailagic, and D. Ding, "Towards a virtual coach for manual wheelchair users," in *Proc. Int. IEEE Symp. Wearable Comput.*, Sep. 2008, pp. 77–80.

[35] R. A. Dienstbier and G. K. Leak, "Effects of monetary reward on maintenance of weight loss: An extension of the overjustification effect," presented at the Amer. Psychol. Assoc. Conv., Washington, DC, 1976.

[36] R. S. Weinberg and J. Ragan, "Effects of competition, success/failure, and sex on intrinsic motivation," *Res. Quart.*, vol. 50, pp. 503–510, 1979.

[37] M. Csikszentmihalyi, *Beyond Boredom and Anxiety*. San Francisco, CA: Jossey-Bass, 1975.

[38] R. J. Vallerand and G. Reid, "On the causal effects of perceived competence on intrinsic motivation: A test of cognitive evaluation theory," *J. Sport Psychol.*, vol. 6, pp. 94–102, 1984.

[39] R. J. Vallerand, "Effect of differential amounts of positive verbal feedback on the intrinsic motivation of male hockey players," *J. Sport Psychol.*, vol. 5, pp. 100–107, 1983.

[40] E. Deci and R. Ryan, *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Plenum, 1985, pp. 29, 318, 322.

[41] M. Zuckerman, J. Porac, D. Lathin, R. Smith, and E. L. Deci, "On the importance of self-determination for intrinsically motivated behavior," *Personality Social Psychol. Bull.*, vol. 4, pp. 443–446, 1978.

[42] C. D. Fisher, "The effects of personal control, competence, and extrinsic reward systems on intrinsic motivation," *Organizational Behav. Human Performance*, vol. 21, pp. 273–288, 1978.

[43] R. B. Margolis and C. R. Mynatt, "The effects of self and externally administered reward on high base rate behavior," Bowling Green State Univ., unpublished manuscript, 1979.

[44] O. C. Jenkins, C. Chu, and M. J. Matarić, "Nonlinear spherical shells for approximate principal curves skeletonization," Univ. Southern California Ctr. Robot. Embedded Syst., Tech. Rep. CRES-04-004, 2004.

[45] H. Fujiyoshi and A. Lipton, "Real-time human motion analysis by image skeletonization," in *Proc. Workshop Appl. Comput. Vis.*, Princeton, NJ, Oct. 1998, pp. 15–21.

[46] S. Waldherr, S. Thrun, R. Romero, and D. Margaritis, "Template-based recognition of pose and motion gestures on a mobile robot," in *Proc. Nat. Conf. Artif. Intell.*, Madison, WI, 1998, pp. 977–982.

[47] C. R. Wren, A. Azarbayejani, T. Darrell, and A. P. Pentland, "Pfinder: Realtime tracking of the human body," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 780–785, Jul. 1997.

[48] J. Wainer, D. J. Feil-Seifer, D. A. Shell, and M. J. Matarić, "The role of physical embodiment in human-robot interaction," in *Proc. IEEE Int. Workshop Robot Human Interactive Commun.*, Hatfield, U.K., Sep. 2006, pp. 117–122.

[49] NeoSpeech Text-to-Speech, 2009. [Online]. Available: http://www.neospeech.com

[50] Microsoft Kinect, 2010. [Online]. Available: http://www.xbox.com/kinect

[51] C. D. Kidd and C. Breazeal, "Robots at home: Understanding long-term human-robot interaction," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2008, pp. 3230–3235.

[52] Y. Matsusaka, H. Fujii, T. Okano, and I. Hara, "Health exercise demonstration robot TAIZO and effects of using voice command in robot-human collaborative demonstration," in *Proc. 18th IEEE Int. Symp. Robot Human Interactive Commun.*, Sep.–Oct. 2009, pp. 472–477.

[53] R. H. Poresky, C. Hendrix, J. E. Mosier, and M. Samuelson, "Companion animal bonding scale: Internal reliability and construct validity," *Psychol. Rep.*, vol. 60, pp. 743–746, 1987.

[54] D. J. Feil-Seifer and M. J. Matarić, "Defining socially assistive robotics," in *Proc. Int. Conf. Rehabil. Robot.*, Chicago, IL, Jun. 2005, pp. 465–468.

ABOUT THE AUTHORS

Juan Fasola received the B.S. degree in computer science and the M.S. degree in computer science from Carnegie Mellon University (CMU), Pittsburgh, PA, in 2006 and 2007, respectively. He is currently working toward the Ph.D. degree in computer science as a member of the Interaction Lab at the University of Southern California (USC), Los Angeles.

His research interests are in human-robot interaction, specifically in the area of socially assistive robotics, and robot learning by demonstration. His current research includes incorporating intrinsic motivation theory into the design and implementation of a socially assistive robot exercise system for the elderly, and developing a methodology for autonomous in-home service robots to learn tasks from nontechnical users through natural communication. He has also studied multirobot coordination and computer vision. He is actively involved in K-12 robotics education outreach and has mentored many graduate and undergraduate students in robotics research projects.

Mr. Fasola has received the Northrop Grumman student scholarship for exceptional research as a student (2006–2007), and his work on the development of a robot exercise coach for the elderly was nominated for the best paper award at the 19th IEEE International Symposium in Robot and Human Interactive Communication (Ro-Man 2010).

Maja J Matarić (Fellow, IEEE) received the B.S. degree in computer science from the University of Kansas, Lawrence, in 1987 and the M.S. degree in computer science and the Ph.D. degree in computer science and AI from the Massachusetts Institute of Technology (MIT), Cambridge, in 1990 and 1994, respectively.

She is a Professor of Computer Science, Neuroscience, and Pediatrics at the University of Southern California (USC), Los Angeles, Founding Director of the USC Center for Robotics and Embedded Systems, Co-Director of the Robotics Research Lab, and Vice Dean for Research in the Viterbi School of Engineering. Her research into socially assistive robotics aims to endow robots with the ability to help people through noncontact assistance in convalescence, rehabilitation, training, and education. Her lab is developing robot-assisted therapies for children with autism, stroke survivors, Alzheimer's patients, and other populations (see http://robotics.usc.edu/interaction/).

Prof. Matarić is a Fellow of the American Association for the Advancement of Science (AAAS) and the Association for the Advancement of Artificial Intelligence (AAAI). She is a recipient of the Presidential Award for Excellence in Science, Mathematics and Engineering Mentoring, and the Okawa Foundation, NSF Career, MIT TR35 Innovation, and IEEE Robotics and Automation Society Early Career Awards. She is featured in the documentary *Me & Isaac Newton*, in the 2009 *New Yorker* article "Robots that care," and is one of five *LA Times Magazine* 2010 Visionaries. She is Co-Editor-in-Chief of a major journal and associate editor of three others, has published extensively, serves on a number of advisory boards, and is active in K-12 outreach.