Evaluating Human-Robot Interaction Using a Robot Exercise Instructor at a Senior Living Community

Lundy Lewis[1(✉)], Ted Metzler[2], and Linda Cook[2]

[1] Computer Technology, Southern New Hampshire University, Manchester, NH, USA
l.lewis@snhu.edu

[2] School of Nursing, Oklahoma City University, Oklahoma City, OK, USA
{tmetzler,lcook}@okcu.edu

Abstract. A NAO humanoid robot and a Paro animaloid robot are taken to a senior living community for a study in human-robot interaction. The humanoid robot was programmed to perform autonomously (i) a warm-up routine in which the robot directs the participants to ask it to perform various tasks and (ii) an exercise routine in which the robot invites the participants to participate in various physical exercises. The Paro robot is then passed around among the participants. The participants included six elderly residents, three nurses/caregivers, and two administrators. The elderly residents are categorized with respect to cognitive awareness and physical capability. We tabulated video data to measure several dimensions of human-robot interaction among these diverse participant groups. Our findings suggest that while senior residents moderately accept the robots and nurses and administrators are enthusiastic about them, more work needs be done on the auditory capability of the robots.

Keywords: Elder care · Quality of life · Socially assistive robot · Exercise · Intelligent architecture · Human-Robot interaction

1 Introduction

It is well-known that the number of elders will increase dramatically over the next few decades while the number of care-givers will decrease, thus resulting in a social problem to which researchers, entrepreneurs, and governments are currently seeking possible solutions. Reasons for the impending problem include (i) a better understanding of healthy habits such as food intake and physical/mental exercise, (ii) advances in medicine and treatments of physical ailments such as hip and knee replacements, (iii) a better understanding of supportive environments such as mobility enhancement and transportation, (iv) fewer offspring from the current generation of potential parents, and (v) better social security for the elderly, among other reasons. Current trends among elders are aging-in-place in the elderly's familiar home or alternatively joining a senior living community. Fewer seniors elect to live with their offspring.

N. Kubota et al. (Eds.): ICIRA 2016, Part II, LNAI 9835, pp. 15–25, 2016.
DOI: 10.1007/978-3-319-43518-3_2

Nonetheless, elderly people will have difficulty performing routine tasks due to natural mental and physical decline. Robots show promise to assist the elderly with these tasks and thus help improve their quality of life. As an example, a tele-presence robot may allow a person to visit friends, relatives, and sights of interest remotely, or conversely allow friends and relatives to visit the elderly person. Simple virtual robots may remind the elderly of events or tasks such as medicine-taking. Humanoid and animaloid robots may provide companionship and conversation, or perhaps advice and therapy.

Some researchers are experimenting with humanoid robots as motivators for physical exercise [1]. Similarly, the pilot study described here involves a session with the humanoid robot NAO serving as an exercise coach on the premises of the Golden Oaks Senior Living Community in Oklahoma, USA [2] on October 23rd 2015. The goals of the study were to assess the acceptability of the robot among senior residents, caregivers, and administrators. The robot was programmed to be interactive and autonomous. The interaction is both verbal and physical, in which the robot reacts to the subjects' verbal input and the subjects react to the robot's verbal directives and physical demonstrations. It first provided an introductory, warm-up routine and then proceeded to lead the group through exercises focusing on leg, arm, feet, hands, neck, eyes, and full-body.

The authors described the details of the pilot study in a prior workshop paper [3] wherein we discussed the importance of context in studies on the acceptance of robots as social assistants for the elderly [4], our iterative methodology (Soft Systems Methodology [5, 6]), design principles, implementation of behaviors into the NAO robot, and a preliminary analysis of the results. In this paper we provide a further in-depth analysis of video data with respect to human-robot interaction (HRI). In Sect. 2 we discuss the parameters of the study, including the environment, the participants, and design concepts. In Sect. 3 we describe the architecture and implementation of the routines in the NAO Choregraphe development environment. In Sect. 4 we review a preliminary analysis of the study [3], discuss our subsequent analysis on human and robot responsiveness during the study, and our analysis of focus sessions with all participants. The paper concludes with comparisons with other work in the field and plans for further study.

2 The Parameters of the Study

2.1 The Environment and Participants

The subjects under study included six elderly residents, three nurses, and two administrators. The elderly residents were selected by the nurses and administrators with the goal of having seniors with varying cognitive and physical deficits. All eleven subjects participated in the exercises. Figure 1 shows the environment for the study. Table 1 characterizes each of the six elderly residents R*n*.

Fig. 1. Robot coach-to-senior interaction during an arm exercise

Table 1. Characterization of the senior residents

	Cognitive awareness	Physical capability	Physical dependencies	Stamina & determination
R1	Low to moderate	Low	Walker	Low
R2	High	High	Wheelchair	High
R3	High	High	Wheelchair	High
R4	Very high	Very high	None	Very high
R5	Very low	Very low	Wheelchair	Very low
R6	Very high	Moderate	None	Very high

2.2 Method of Assessment

To assess the participants' engagement with the robot and the acceptability of the robot, we established several high-level design concepts as listed below, based on work reported in [1]. Next, we posited observable visual characteristics of the participants and mapped these characteristics into the design concepts. By video analysis, we simply counted and tabulated the observable characteristics.

Motivating. The degree to which the participants enjoy the engagement with the robot, indicated by observables *Attentive*, *Smiling/Laughter*, *Head-Nodding*, *Leaning Forward*, and *Intensity*.

Fluid and Highly Interactive. The degree to which the participants and the robot react to each other including verbal and physical reactions, indicated by observables *Cooperation*, *Imitation*, *Initiation*, *Intensity*, and *Participation*.

Intelligent. The degree to which participants perceive the robot as an entity that can be trusted as a competent human being rather than a toy, indicated by observables *Cooperation*, *Imitation*, *Initiation*, *Intensity*, and *Participation*.

Task-Driven. The degree to which participants perceive the robot as working towards a goal via various relevant tasks, indicated by observables *Cooperation*, *Imitation*, *Initiation*, and *Participation*.

Since our study will be carried out over multiple sessions where robot behaviors and routines will evolve over the sessions, the robot itself needs to be portable and programmable in a reasonable amount of time. Thus, we add two additional design concepts that are technology-driven:

Programmability. The degree to which the robot can be re-programmed in light of new suggestions and ideas offered by the participants, to be determined in future iterations of exercise therapy as the study unfolds.

Portability. The degree to which the robot and accompanying apparatus can be transported to different locations and set up in a reasonable amount of time without technical difficulties, to be determined in future iterations of exercise therapy as the study continues.

3 Robot Platform, Software, and Behavior Programming

The robot platform is the humanoid NAO robot. See Fig. 1. The technical specifications of the robot can be viewed at [7]. The robot is portable and offers several methods of behavior programming: (i) visual programming, (ii) Python and C++ programming, (iii) posture and gesture programming, and (iv) conversational programming [3]. These methods support the technology-driven design concepts of Programmability. The robot is capable also of random fluid gestures when listening or speaking to users, thus contributing to the design concepts Motivating, Fluid and Highly Interactive, and Intelligent.

The robot behaviors were implemented as intelligent agents and the overall design is based on the subsumption architecture [8]. The agents reside on levels such that higher levels subsume lower levels. If higher-level agents cannot function, the lower-level agents may continue to function.

For the exercise behavior, we developed (i) an introduction agent whereby the robot describes explains what he plans to do, (ii) a conversational agent to interact with the participants, (iii) a multiplicity of agents representing specific forms of exercise, and (iv) a transition agent that completes a specific exercise and initiates conversation with the participants to determine the next form of exercise. The agents communicate via message-passing between the input and outnodes nodes of the agents. The agents execute in linear fashion rather than distributive, collaborative fashion. Future behaviors may require collaborative decision-making. Figure 2 illustrates the exercise routine in accordance with the agent-based, subsumption architecture. Portions of the figure a circled in red for purposes of explanation.

The introduction agent on level 2 is passive. It speaks to the participants about the exercise agenda: “OK, let’s do some exercises. We will exercise our head, arms, legs, hands, eyes, and feet. When you are ready we will try Tai Chi. First, let’s get in the sitting position like I am now. Ready? OK, what would you like to do first? Head, arms, legs, hands, eyes, or feet?” After the introduction, control is passed to the conversational agent.

The conversational agent on level 1 has a reactive component and a deliberative component. The reactive component listens for audio input from participants and tries to detect a word it understands. The deliberative component is a simple rule-based system whereby the detection of a word routes control to a particular exercise. If the conversational agent doesn’t detect an audio input within 5 s, it selects a random exercise and passes control to the selected exercise agent on level 0. If that doesn’t work, a robot operator may start an exercise agent manually by clicking on the first box comprising the agent. In this way, the operator may take control of the robot in tele-operated fashion rather than abort the session. In our study we did not have to resort to tele-operation the robot, although it is conceivable that future instances might require tele-operation. It is comforting to know that this feature is available although one would hope not to have to use it.

The eight rows in level 0 comprise the exercise agents for specific forms of exercise, save row 7 row which is a “stop” agent. As an example, we have drawn a red circle around the behaviors comprising the leg exercise agent in row 3. The seven boxes in this row instruct the robot to perform the following verbal and physical communication with the participants: “Let’s stand and then squat down, but be careful”. Try it only if you feel comfortable. Here we go. Watch me. Stand. (The robot slowly moves from the sitting position to the standing position). Squat. (Robot slowly moves to a squatting position.) Back up. (Robot moves to the standing position, and so on).

The order of the exercises from top to bottom in Fig. 2 are head, arms, legs, Tai Chi, hands, eyes, stop, feet. The order in which the exercises are programmed is insignificant since exercises will be determined by the conversational agent on level 1. One can see that other forms of exercise are rather easily programmed and embedded into the existing structure, thus contributing to the design concept Programmability. A harder challenge would be to program the robot to gauge the success, progress, and/or motivation of individual participants and to give such individuals special treatment and advice. This feature is a subject for future research. It is yet to be determined whether such a feature is well-advised.

The transition agent on level 0 concludes each exercise with a compliment for the participants by randomly saying some synonym for “good,” e.g. ‘very good’, ‘excellent’, ‘beautiful’, ‘cool’, ‘very cool’, ‘most cool’, ‘splendid’, or ‘wonderful’, then saying “Let’s sit back and relax (at which point the robot goes into the sitting position)”, then saying “What would you like to do next? Head, arms, legs, hands, eyes, feet, or Tai Chi?”, and finally passing control back to the conversational agent.

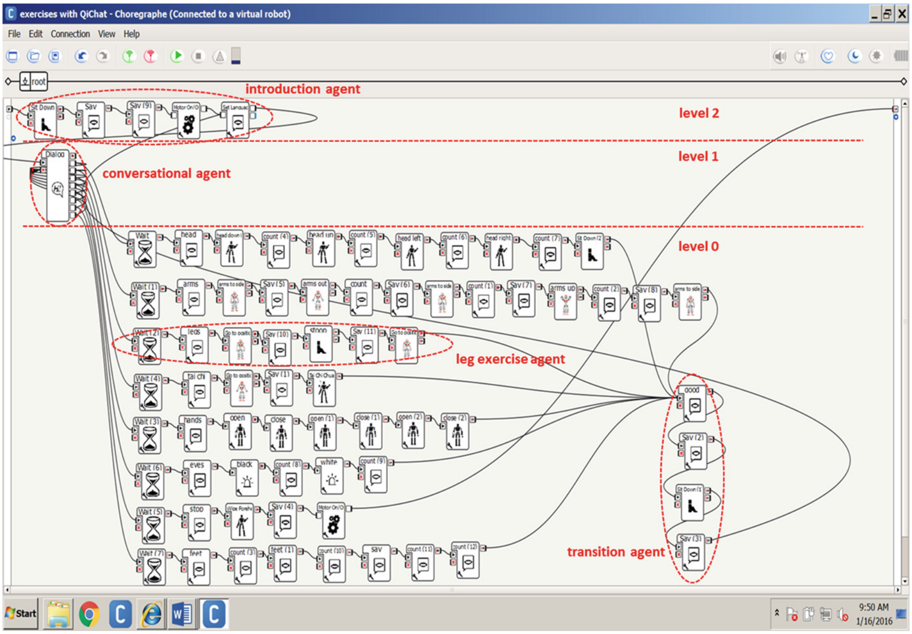

Fig. 2. Structure and implementation of the agents in the exercise routine [3]

4 Results of the Study

4.1 Preliminary Results of the Pilot Study

We designed three activities for the participants: (i) a warm-up routine with NAO, (ii) an exercise routine with NAO, and (iii) a session with Paro. Paro is robotic model of a baby harp seal and is used for pet therapy. The total amount of time spent with the participants was 36 min. Of those 36 min, the warm-up comprised 8 min, the exercise routine 15 min, and the session with Paro 9 min, totaling 32 min of interaction with the robots.

Each participant was scored on a scale of 1 to 10 for each observable discussed in Sect. 2.2 above. An additional category "Approval" was a subjective assessment by the authors. The "Total Engagement" score reflects the number of engagement points out of 100. Table 2 shows the preliminary results in raw form for the time spent interacting with the robots.

The average total engagement score of residents is 53 and the average score of staff is 71.4, suggesting that residents are slightly more than ambivalent in their approval of the robot, while nurses and administrators approve. Visual inspection shows that three of the residents highly approve, one is rather ambivalent, and two do not appear to approve – R1 and R5. We note that the low scores of R1 and R5 correlate with their mental and physical characteristics as shown in Table 1.

The evidence thus far invites a number of questions and suggests some lessons learned that will influence our future work. Inasmuch as the results show a stronger

Table 2. Preliminary results

	R1	R2	R3	R4	R5	R6	S1	S2	S3	S4	S5
Attentive	7	10	10	8	10	10	10	10	10	10	10
Smiling & laughter	1	9	10	2	3	10	9	9	10	8	8
Head-nodding	0	0	3	0	0	10	0	3	10	0	0
Leaning forward	0	0	0	0	0	6	0	0	7	0	0
Cooperation	2	9	10	6	4	9	8	8	8	7	7
Imitation	0	10	9	7	0	7	8	8	8	7	7
Initiation	0	3	2	9	0	9	8	8	8	0	7
Intensity	0	10	10	7	0	9	8	8	8	8	8
Participation	3	10	10	7	1	9	10	10	10	10	10
Approval	1	9	9	5	4	9	9	10	10	8	9
Total engagement	14	70	73	51	22	88	70	74	89	58	66

approval of the robots by staff than by residents, a number of potential factors causing the difference are possible. The difference may reflect one or more of the following kinds of differences separating the two groups: prior technological experience; experience with art forms such as movies involving robots; and the role differences in the Golden Oaks environment.

Based on the mapping of our observables into the design concepts, the first four items reflect pleasure and fun while the remaining items reflect actual work. With the exception of R1 and R5, one can see by visual inspection of the table that the participants were enjoying the engagement with the robot. One can see also that all participants scored favorably with respect to Cooperation, Imitation, Intensity, and Participation. These observations are encouraging.

4.2 An Analysis of Responsiveness by the Participants and by the Robot

In a second analysis of the video data we wished to examine more closely the responsiveness of the participants to the robot's directives and also the responsiveness of the robot to the participants' directives. For this study we considered the activities with NAO since Paro does not issue directives. During the warm-up routine and the exercise routine with NAO there were a total of 70 directives issued by the robot and 23 directives issued by the participants.

Table 3 shows the number of times when at least one participant responded to the robot's verbal directives. During the warm-up routine the robot asked the audience to ask him to do something, where each participant had a "cheat sheet" listing the things the robot could do: sing, recite a poem, tell a joke, wave, chill out, exercise, dance, or stop. During the exercise routine the robot asked the participants what exercise they would like to do next, also with a cheat sheet: head, arms, legs, Tai Chi, hands, eyes, feet, or stop. Sometimes a participant would respond immediately without prompting, sometimes only when prompted or encouraged by a nurse or administrator, and sometimes a participant would respond too quickly, i.e. before the robot finished a directive, in which case the robot did not hear the response. The numbers in Table 3 show that the HRI along this dimension was very good, where the participants improved from the first activity to the second. The fact that the participants responded too quickly 4 times during the exercise routine might suggest increased comfort with the robot.

Table 3. Participants' verbal response to robot directives during both routines (27 total)

	Warm-up routine	Exercise routine
Without prompting	7	14
With prompting	2	0
Too quickly	0	4
No response	0	0

For each exercise during the exercise routine, the robot issued additional directives instructing the participants to perform some sort of bodily behavior. We considered the response of the participants to be positive if the majority of them performed the behavior. For example, during the arm exercise the robot issued the following six directives (refer to Fig. 1): "Let's stand up with our arms down, like this" – "now let's raise our arms just a little and count like this 1 2 3 4 5 … good" – "now let them rest to the side again like this 1 2 3 4 5 … good" – "now raise them high and count to five like this 1 2 3 4 5 … good" – "now back to the side … beautiful" – "let's sit back and relax". Table 4 shows that the participants followed all exercise directives perfectly without prompting. This result is quite encouraging.

Table 4. Participants' exercise response to robot directives during the exercise routine (43 total)

	Warm-up routine	Exercise routine
Without prompting	N/A	43
With prompting	N/A	0
No response	N/A	0

Next we measured how the robot responded to directives from the participants, where sometimes the robot understood the speaker perfectly without the speaker having to repeat the directive, sometimes only when the speaker repeated a directive, and sometimes the robot entirely misunderstood the directive. For example, twice the robot misunderstood "feet" for "hands," once "legs" for "eyes", once "legs" for "hands," once "Tai Chi" for "arms," and once the robot heard "sing" without a directive at all. Table 5 shows that the robot understood a directive correctly without repetition roughly one third of the time, needed repeating a third of the time, and misunderstood the directive roughly a third of the time. Clearly the robot, the participants, and/or the environment need to be examined to improve the auditory capability of the robot. On the other hand we noted that participants laughed profusely when the robot misunderstood a directive, which may have contributed to the smiling/laughter item in our preliminary analysis.

Table 5. Robot's response to human directives during both routines (23 total)

	Warm-up routine	Exercise routine
Without repetition	3	5
With repetition	5	4
Misunderstood	1	5
No response	0	0

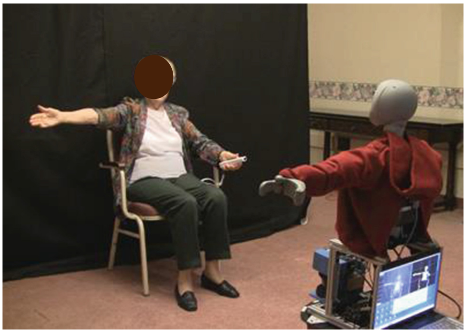

Fig. 3. Coach-to-senior robot at USC [1]

Fig. 4. RoboCoach [9]

4.3 Subsequent Focus Sessions with the Participants

The subsequent focus sessions with each group provided a wealth of suggestions to consider. The discussion among the nursing participants was fruitful, producing positive, enthusiastic, and creative discussion of the potential of robots on the premises. The group displayed sophistication concerning HRI, independently raising subjects and questions that one finds in HRI literature. Among the nurses, prior direct experience with a telepresence robot allowing hospital physicians to visit patients remotely was mentioned. A nurse's suggestion of robots possibly mimicking family members of Alzheimer residents provoked discussion or related ethical concerns. Another nurse participant introduced discussion of privacy issues, e.g. wondering whether robots watching residents would generate objections.

Ideas from administrative participants began with a marketing agent's speculation that the current generation of residents might show acceptance of robot presence rather different from the impending generation of new residents from the "baby boomer" generation. The marketing agent also commented on the psychological value of robots that could recognize individual residents and address them by name. The facility director and the marketing agent displayed further HRI sophistication in observations concerning the ability of robots to perform repetitive tasks for long periods of time without showing human frustration about wasting time. Referring to the facility's dining hall, the director added that a robot capable of taking orders at the tables and transmitting them wirelessly to the kitchen could save time and be "very helpful."

The resident group clearly enjoyed talking about the robot as well, evidenced by humor and laughter. Suggestions from the residents included talking to the robot, asking questions, and playing with it as one would play with a toy. One resident expressed a desire to have a robot to replace his wife.

5 Conclusion

The results of our first session of a humanoid robot leading exercises at a senior living community encourage us to continue our work. Plus, the results of other research groups on the same topic are motivating. Figure 3 shows the exercise coach developed at the University of Southern California (USC) in the United States and Fig. 4 shows the same by a research group at Ngee Ann Polytechnic (NAP) in Singapore.

The primary difference between the USC study and our study are (i) our study involves robot-to-group interaction whereas the USC study involves robot-to-individual-senior interaction one at a time, (ii) our robot communicates with participants verbally and physically in mutual interaction whereas the USC robot communicates via a remote clicker, and (iii) our study faces the complexities and dynamics of senior living in their real environment whereas the USC study is more of a lab experiment. Further, we question whether the USC robot satisfies the programmability and portability design concepts posited in our study. Nonetheless, the results of the USC study are largely consistent with those of our study. With respect to the NAP robot, Fig. 4 shows that the environment and the robot-to-group interaction is similar to that in our study. However, to our knowledge there is no published work that discusses the study in detail from a experimental/scientific point of view. For example, we cannot determine whether the robot is tele-operated or autonomous. In our study, the robot is completely autonomous.

For the next sessions at the facility we plan to (i) improve the conversational behavior of the robot to make it more personal [10], (ii) improve the responsiveness of the robot to the participants' directives, (iii) improve our study evaluation criteria, (iv) program the robot so that it can learn to attach names to faces and thus identify participants by name, (v) obtain more complete profile information on the participants in order to compare diverse population groups, (vi) add more exercise routines, e.g. a simple meditation exercise, and (vii) carefully select viable candidates with varying degrees of cognitive and physical abilities. For example, a better demographic profile of the participants (age, gender, level of education, previous technological experience) and clear selection criteria for the participants could make a good case for the integration of socially assistive robots into the lives of older people with varying cognitive and physical impairments. Future research will focus on these issues.

It is difficult to assess or predict the participants' motivation over time with multiple sessions based on the one experiment described here. Our future research will address this question. Future research will study also the effect of participants' characteristics with respect to personality, background, and individual preferences. We plan to tabulate and compare the response results in Tables 3, 4 and 5 for each participant. It is likely that reaction patterns will vary with each individual. If the pattern could be classified based on the type of personal characteristic, the result would be useful. Finally, our analysis focusses mostly on results for NAO. Further analysis will include results for PARO.

References

1. Juan Fasola, J., Matarić, M.: A socially assistive robot exercise coach for the elderly. J. Hum.-Robot Interact. **2**(2), 3–32 (2013)
2. www.goldenoaks.com. Accessed 19 Apr 2016
3. Lewis, L., Metzler, T., Cook, L.: Results of a pilot study with a robot instructor for group exercise at a senior living community. In: Bajo, J., et al. (eds.) PAAMS 2016 Workshops. CCIS, vol. 616, pp. 3–14. Springer, Heidelberg (2016). doi:10.1007/978-3-319-39387-2_1
4. Talaei-Khoei, A., Lewis, L., Talaei-Khoei, T., Ghapanchi, A.: Seniors' perspectives on perceived transfer effects of assistive robots in elderly care: a capability approach analysis. In: International Conference on Information Systems, Fort Worth, Texas (2015)
5. Checkland, P., Poulter, J.: Learning for Action: A Short Definitive Account of Soft Systems Methodology, and Its Use for Practitioners Teachers and Students. Wiley, Chichester (2007)
6. Checkland, P.: Soft systems methodology: a thirty year retrospective. Syst. Res. Behav. Sci. **17**(1), 11–58 (2000)
7. doc.aldebaran.com/2-1/family/robots/index_robots.html#all-robots. Accessed 20 May 2016
8. Brooks, R.: A robust layered control system for a mobile robot. IEEE J. Robot. Autom. **2**(1), 14–23 (1986)
9. www.np.edu.sg/sg50/events/Pages/20150101_npevents_robocoach.aspx. Accessed 19 Apr 2016
10. Lewis, L.: Using narrative with avatars and robots to enhance elder care (Chapter 14). In: Ghapanchi, A., Tavana, M., Talaei-Khoeif, A. (eds.) Healthcare Informatics and Analytics: Emerging Issues and Trends. IGI Global, Pennsylvania (2014)

Springer

http://www.springer.com/978-3-319-43517-6

Intelligent Robotics and Applications
9th International Conference, ICIRA 2016, Tokyo, Japan,
August 22-24, 2016, Proceedings, Part II
Kubota, N.; Kiguchi, K.; Liu, H.; Obo, T. (Eds.)
2016, XXIII, 493 p. 321 illus., Softcover
ISBN: 978-3-319-43517-6