Encouraging Physical Therapy Compliance with a Hands-Off Mobile Robot

Rachel Gockley
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213
rachelg@cs.cmu.edu

Maja Matarić
University of Southern California
RTH 407,sq Mailcode 781
941 West 37th Place
Los Angeles, CA
mataric@usc.edu

ABSTRACT

This paper presents results toward our ongoing research program into hands-off assistive human-robot interaction [6]. Our work has focused on applications of socially assistive robotics in health care and education, where human supervision can be significantly augmented and complemented by intelligent machines. In this paper, we focus on the role of *embodiment*, empirically addressing the question: "In what ways can the robot's physical embodiment be used effectively to positively influence human task-related behavior?" We hypothesized that users' personalities would correlate with their preferences of robot behavior expression. To test this hypothesis, we implemented an autonomous mobile robot aimed at the role of a monitoring and encouragement system for stroke patient rehabilitation. We performed a pilot study that indicates that the presence and behavior of the robot can influence how well people comply with their physical therapy.

Categories and Subject Descriptors

K.4.2 [**Computers and Society**]: social issues; J.4 [**Social and Behavioral Sciences**]: psychology; I.2.9 [**Artificial Intelligence**]: robotics

General Terms

Human Factors, Experimentation

Keywords

Human-robot interaction, embodiment, social robots, psychology, physical therapy, stroke recovery

1. INTRODUCTION

Assistive robotics, which encompasses rehabilitation robotics, is a rapidly growing research area, with the majority of work being aimed at hands-on assistance, in which the robot physically aids the patient. Recovery post-stroke is one of the largest application domains, where various hands-on systems have been developed [3, 4, 14, 15]. However, other research indicates that this kind of robotic assistance in therapy may not provide any strong benefit over traditional therapy methods [9]. Additionally, having a robot physically touching a person raises many safety concerns.

Human Robot Interaction '06 Salt Lake City, Utah USA

Figure 1: The Pioneer 2-DX mobile robot used in this study, with a SICK scanning laser range finder mounted on top.

In contrast, our focus is on non-contact (hands-off) assistive human robot interaction, with the role of providing inherently safe robot technologies that aid patient training and recovery through social, rather than physical, interaction. Our previous research [5, 10] indicates that the encouragement offered by a robot during hands-off therapy assistance may convince people to exercise more in various recovery contexts. However, since our robots do not physically aid their users, it is important to examine the role of the robots' *physical embodiment*.

The effects of embodiment on human-robot interaction are still largely unknown [16]. Most studies designed to test such effects have focused on differences between a robot and a video of the same robot [2] or between fairly static representations of the robot and on-screen agent [11]. In the work presented here, we examine the role of the robot's physical expression in the context of assistive interaction. Specifically, we are interested in the aspects of embodiment that cannot be simulated by an on-screen agent, such as shared physical context and physical movement around the user.

To ground the above research question in a real-world context, we are investigating how a robot can be used to encourage patient compliance with therapy, such as in recovery from stroke. Toward that end, we have designed a system that allows us test hypotheses about human responses to interactions with an assistive embodied robot. Shown in Figure 1, the robot is capable of sensing a patient's location and motion and offering support in the form of vocalizations and movement. The system is designed not only to monitor the patient's progress, but also to serve as a friendly companion during the rehabilitation. In this study, we focus on the robot's physical expression through movement, its correlation with user personality, and its effectiveness in motivating exercise.

2. SYSTEM DESIGN

The robot system, shown in Figure 1, was largely unchanged from our previous hands-off assistive stroke rehabilitation work [5]. The system consists of an ActivMedia Pioneer 2-DX mobile robot, augmented with a SICK LMS200 scanning laser range finder. The range finder allows the robot to navigate safely and to track and locate a patient wearing a reflective fiducial on his or her calf. The fiducial allows the robot to find the user robustly, without having to rely on any sophisticated vision processing.

We have found that people tend to anthropomorphize the pan-tilt-zoom camera when mounted on a robot, so we removed the camera for the context of this project. The robot itself is almost entirely *non-biomemetic*—while the location of the laser may evoke vague notions of a "head," the robot bears little resemblance to a human or animal. We believe that this constitutes an interesting research platform for that very reason: we can interpret results in terms of the effects of the robot itself, rather than in terms of the users' anthropomorphized interpretations of the robot.

The assistive system also includes a motion-capture setup worn by the user and used by the robot to detect and track the user's arm movement. We used the motion capture sensors developed in our prior work [12]. Each sensor is a small, light-weight, self-contained inertial measurement unit. The patient wears one (or more) of the sensors on his or her upper arm. The robot communicates with the motion sensors wirelessly in real time. Each sensor can report up to three degrees of freedom at about 33Hz. An example output from a subject's arm movement is shown in Fig. 2.

Detecting the user's arm activity is an important part of the robot's assistive role, because repetitive movement is one of the most effective methods of regaining arm function after a stroke [17, 18]. In designing the experiment, we modeled functional repetitive arm tasks, such as shelving magazines, that provide repetition in the context of everyday activity.

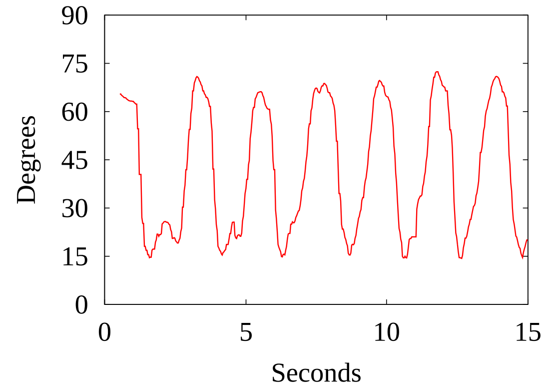

Figure 2: Example data from a motion sensor attached to a person's upper arm. As the person moves his arm up and down, the sensor reports the arm's angle in real-time.

3. INTERACTION DESIGN

In designing the interaction between the user/patient and the robot, we focused on having the robot use its physical embodiment, i.e., space and movement, rather than verbal prompting and encouragement. We chose two dimensions of physical expression: proxemics and level of engagement. Each is described in turn.

3.1 Proxemics

Proxemics, a concept introduced by Hall, is a measure of space around a person [8]. Hall differentiated between four distances, each with a "close" and a "far" phase: intimate, personal, social, and public. Intimate distance generally involves physical contact; one tends to feel uncomfortable if a stranger enters one's intimate space. Personal distance can be thought of as "arm's length" away, and is the distance at which friends may stand to chat. Social distance corresponds to the distance assumed when conducting business and may extend as far as 4 meters, and public distance extends outward from there.

Since intimate distance would risk the robot colliding with the person and public distance is beyond the robot's accurate sensing capabilities, we chose to consider only personal and social distances in this work. In the personal space condition, the robot attempted to stay within the range of 0.5–1.0 meters from the person. In the social space condition, the robot stayed approximately 1.5–2.0 meters from the person.

We believe that people are likely to differ in their preferences for how close the robot may approach them. That is, some people may prefer to have the robot closer, acting as a more intimate source of encouragement, while others may feel that a robot too near by is imposing. We postulate that these preferences are strongly influenced by people's personality traits; for example, less extroverted people tend to have less acceptance of electronic monitoring [19], and thus might prefer the robot to remain farther away.

3.2 Engagement

In addition to distance, we also considered how engaged with the participant the robot appeared to be. We based our concept of "engagement" on the idea of mimicking the person's movements. Bailenson found that people tend to prefer and be more persuaded by agents who mimic their movements at a short delay [1]. Thus, in order to appear highly engaged, the robot moved forward and back following the speed of the user's arm motion. To prevent the user from directly teleoperating the robot, the robot's movements lagged behind the participant's by a random delay of 1–10 seconds. This delay was randomly changed every 10–30 seconds in order to further disconnect the robot's movements from the exact movements of the user. We used randomness in order to make the robot's movements appear more purposeful and to avoid having users focus on controlling the robot rather than on performing the exercise task.

We designed two levels of engagement. In the "high engagement" condition, the robot behaved as described, while in the "low engagement" condition the robot did not follow the participants' arm movement at all. In both conditions, the robot performed an "encouragement" behavior when the participant paused for longer than 5–10 seconds. The encouragement behavior consisted of the robot briefly turning back and forth in place over a 60-degree arc. Thus, even in the low engagement condition, we intended for the robot to move at least occasionally.

4. HYPOTHESES

Based on the above discussion, we formed the following hypotheses:

1. All participants will prefer, and will exercise longer with, the highly engaged robot.
2. Participants will vary in their proxemics preference, based on their personality profiles. In particular, we hypothesized that more extroverted participants will prefer the robot to stay closer.

5. PROCEDURE

Following the robot design discussed above, we conducted a pilot study to test our hypotheses. The study had a 2 (proxemics) x 2 (engagement) design. Dependent variables included the personality of the participant (self-reported), time on task, accuracy in compliance, and the participant's self-ratings of several aspects of the task, including the robot.

At the start of each experiment, the participant was given a personality survey consisting of the 50-question Big Five Inventory [7], the 10-question Technology Readiness Index [13], and basic demographic information.

The experimenter then informed each participant that the study was intended to test an exercise monitoring system with and without robotic augmentation. Each participant was then asked to perform three exercise tasks, with breaks in between. As noted above, the tasks were all designed to be functional exercises that are similar to tasks used during standard stroke recovery therapy. We designed three different tasks to help maintain participant interest. The tasks were as follows:

1. Moving wooden pencils, one at a time, from one bin into another. The two bins were placed approximately one meter apart on a desk.
2. Lifting magazines from a desktop onto a raised shelf, one at a time.
3. Flipping through the pages of a newspaper.

Each task was designed to be open-ended; participants were told that once they finished the task (e.g., moved all the pencils or magazines or got to the end of the newspaper), they could reverse/repeat the process (i.e., move the pencils/magazines back, etc.). Specifically, the experimenter asked each participant to "repeat this process until you feel that you have exercised your arm enough at this time." Participants could press a button on the table to signal when they wished to stop. If a participant continued exercising longer than six minutes in any condition, the experimenter ended the trial.

The experimenter then fitted the participant with a motion sensor and a laser fiducial. As discussed above, a motion sensor was strapped to the participant's upper arm, and a reflective laser fiducial was strapped around the participant's lower leg, both using velcro straps.

Because this was only a pilot study, we did not have enough participants to run a fully random design. Rather, the three tasks were consistently presented in the order given above (pencils, magazines, newspaper), and each participant was randomly assigned two of the four possible robot conditions (2 proxemics and 2 engagement), selected such that all possible combinations of conditions might be tested. Additionally, we also randomized which of the three tasks served as the control (no robot) condition for each participant.

After each task, the experimenter had the participant complete a brief survey. In the control condition, the survey consisted of only two questions, asking the participant to rate the degree to which he or she stopped due to boredom and due to tiredness. In the robot conditions, the survey contained an additional 10 questions regarding the participant's opinion of the robot, including several questions designed to verify our manipulations. All questions were presented on a 5-point scale ranging from "strongly disagree" to "strongly agree."

After the participant had completed all three tasks, the experimenter presented a short debriefing.

6. RESULTS

We had 11 participants (8 male and 3 female) for our pilot study, most of whom were in the 20- to 30-year age group. All participants had finished at least two years of university study, and had fairly high understanding of robotics and of technology in general (all had at least 12.5/20 on the

Table 1: Average time spent and exercise chunks (semi-completion of a task, such as shelving all magazines) performed by participants in each robot condition. Standard deviation is shown in parentheses. In the control condition, participants averaged 250.4(80.1) seconds and 1.10(0.50) chunks. All run times were capped at six minutes (360s).

Proxemics	**Engagement**				*Total*	
	High		**Low**			
	Seconds	Chunks	Seconds	Chunks	Seconds	Chunks
Personal	273.4(104.6)	1.11(0.36)	277.7(93.7)	1.17(0.21)	275.7(93.6)	1.14(0.27)
Social	258.3(104.5)	1.04(0.43)	232.8(101.7)	1.06(0.33)	246.7(98.7)	1.05(0.37)
Total	265.2(99.5)	1.07(0.38)	257.3(95.3)	1.12(0.26)	261.2(95.1)	1.09(0.32)

Figure 3: The experimental setup. The shown participant is performing task 1 (pencils) with the robot at a social distance. The laser fiducial is visible on the participant's leg. The motion sensor is mounted on the left upper arm, and thus not visible in this image.

TRI scale). Since each participant experienced two robot conditions, we had 5–6 participants in each combination of proxemics and engagement conditions. Due to the small number of participants and the large number of measures, we are unable at this time to run full factorial analyses to determine all possible interactions between measures. We focus instead on correlations and interactions of interest, and do not in general report on non-significant effects.

We designed two of the survey questions as manipulation checks. To test the proxemics condition, we asked participants to rate, on a scale of 1 ("strongly disagree") to 5 ("strongly agree"), whether the robot ever "came too close." While few participants agreed with the statement (overall $M = 1.72$), an analysis of variance (ANOVA) indicated a significant difference between answers across the two conditions, with ratings of the robot in "personal space" higher (personal $M = 2.09$, social $M = 1.36$; $F[1, 20] = 5.08$, $p = 0.04$).

To test the engagement condition, we asked participants to rate whether the robot "appeared to be engaged with me." Unfortunately, a similar analysis indicated that this manipulation was not successful ($M = 2.86$, $F[1, 20] = 0.83$, n.s.), though overall participants rated the highly engaged robot as slightly more engaging (high $M = 3.09$, low $M = 2.64$).

We also analyzed the "engagement" survey response with an ANOVA, modeling the main effects of both the proxemics and the engagement conditions and their interaction. While there were no significant main effects, the interaction of variables was significant ($F[3, 22] = 4.87$, $p = 0.04$). Contrast tests showed that, when the robot was in "personal space," participants did not differentiate between engagement conditions (overall $M = 2.88$, $F[1, 18] = 0.73$, n.s.). However, when the robot was in "social space," participants rated the highly engaged robot as significantly more engaged than the low-engagement robot (high $M = 3.50$, low $M = 2.00$; $F[1, 18] = 5.13$, $p = 0.04$). No other survey questions differed significantly across any combination of proxemics or engagement conditions.

Nine of the 33 trials were terminated by the experimenter, as the participants in those trials had not stopped by the six minute cut-off. Those trials were spread across conditions and tasks, except that three occurred in the personal space and low engagement condition. Including cut-off runs, participants spent similar amounts of time on each of the three tasks (approximately 4.5 minutes, $F[2, 30] = 0.09$, n.s.). However, the number of arm movements, such as transferring a single pencil or magazine, differed significantly across tasks; participants performed more movements in Task 1 than in either of the other two tasks ($F[2, 30] = 6.25$, $p = 0.005$). We observed that participants tended to perform each task in "chunks," such as transferring all pencils from one basket and back, shelving all of the magazines, and flipping through all pages of the newspaper. That is, though we designed the tasks to be open-ended, each task had a point at which the task was "completed," such as when all magazines were on the shelf (or back off again), and many participants chose to stop at such a point. The number of arm movements constituting such a task segment varied across tasks, but participants consistently tended to perform repetitions in multiples of those numbers. Analyzed in terms of these "chunks," then, the three tasks did not differ significantly (approximately 1.1 chunks per task, $F[2, 30] = 0.16$, n.s.). Average times and chunk repetitions are reported in Table 1.

The control condition allowed us to analyze whether the mere presence of the robot had any effect on participant performance. We looked at both seconds spent and number of "chunks" performed each as an ANOVA modeling the main effects of participant and presence of the robot and the interaction effect of the two. These models predicted approximately 82% and 91% of the variance in seconds and chunk repetitions, respectively. Participants differed significantly in terms of how many chunks they performed ($F[10, 11] = 9.05$, $p < 0.001$) and marginally in terms of seconds ($F[10, 11] = 2.26$, $p = 0.1$). That is, the amount of exercise performed varied from person to person, as might be expected.

While the presence of the robot did not have a significant main effect in either case, we found a significant interaction effect between participant and the presence of the robot on the number of chunks performed ($F[10, 11] = 2.74$, $p = 0.06$). In particular, about one-third of participants exercised more when the robot was present and one-third exercised less. We did not find any direct correlation between any personality measures and the amount of exercise performed by participants.

We also analyzed the effects of the proxemics and engagement conditions on participant performance. Unfortunately, an ANOVA indicated that neither condition (nor their interaction) had a significant effect on either the seconds spent or the number of chunks performed in the robot condition. However, we did find strong correlations between participants' performance and their ratings of various aspects of the robot. Participants who saw the robot as more engaging tended to exercise longer ($r = 0.49$, $p = 0.02$) and perform more chunks ($r = 0.43$, $p = 0.05$) than participants who rated the robot as less engaging. Additionally, participant awareness of the robot correlated significantly with how many seconds they exercised ($r = 0.64$, $p = 0.002$) and how many chunks they performed ($r = 0.59$, $p = 0.004$). In general, we did not find correlations between participant personality and post-survey responses. However, whether participants felt that the robot "came too close" inversely correlated with participant extroversion ($r = -0.40$, $p = 0.07$).

7. DISCUSSION

In the presented pilot study with limited participation, we have found some support for both of our hypotheses. Although the engagement manipulation was not entirely successful, participants tended to exercise more with the robot when they perceived the robot as engaged, which follows Hypothesis 1. Additionally, we found a strong inverse correlation between participant extroversion and preference for the robot to stay further away, in support of Hypothesis 2. That is, more extroverted participants were less likely to feel that the robot came too close to them.

Another interesting result is the interaction between individual participants and the presence of the robot. We found that some participants exercised more (in terms of both seconds and task "chunks") when the robot was present, others exercised more when the robot was absent, and still others did not seem affected by the robot's presence/absence. Unfortunately, we are not at this time able to map people's behavior to their personality traits, as we had hoped. Further research is needed to understand the mechanism behind people's preferences, and to develop methods for assistive robots to learn the needs of individual users.

Figure 4: A participant in the initial experimental set-up turning around to watch and "play" with the robot.

Several implementational features likely had key impact on our results. We believe that participants did not always recognize the engagement condition for several reasons. The primary reason is that, even in the "low engagement" condition, the robot still moved around a large amount in several cases, resulting from the simple thresholding used to determine when a participant had not moved sufficiently; the "encourage" action was performed more frequently than we had anticipated. Additionally, in most trials, the geometry of the experimental set-up (having the desk against the wall) was such that the robot tended to stay largely behind the participant. While some participants turned to watch the robot (see Fig. 4), one participant remarked that she did not even notice the robot's presence. While we did eventually modify the set-up, we did not have enough remaining participants to note any significant difference.

8. CONCLUSIONS AND FUTURE WORK

We have presented a robotic system intended for non-contact physical therapy assistance, and have used the system as a means of testing hypotheses about the role of robot embodiment. Although the robot offered no vocal encouragement or reminders, its presence and movement had an effect on how well participants complied with the therapy tasks. The exact effects of the robot—whether the participants exercised more or less in both time and accuracy—varied significantly across participants. Overall, participants who thought the robot was engaged with them tended to exercise longer, often watching the robot as they worked. Additionally, we found support for the idea that people's preferences for the robot's behavior vary according to their personality traits; in particular, we found that more extroverted people were more comfortable with the robot entering their personal space.

We consider it highly significant that we found these results, though small, in such a small sample. In particular, our participants generally had a high level of familiarity with robots, and thus were probably less likely to be influenced by them than the general population. Thus we believe that the effects may be more significant with a more representative population and plan to perform more extensive experiments to that effect.

Specifically, our next step is to refine our experimental procedure and run the full study with recovering stroke patients. We believe that this work, in conjunction with our earlier work on hands-off assistive robotics, demonstrates that this technology can be beneficial to stroke survivors as well as in other assistive application domains.

9. ACKNOWLEDGMENTS

This work was funded in part by an NSF Graduate Research Fellowship to the first author and by NSF grants #IIS-032914 and #IIS-0121426. In addition, we would like to thank Jodi Forlizzi and Reid Simmons for their support and assistance in the design of this study.

10. REFERENCES

[1] J. N. Bailenson and N. Yee. Digital chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychological Science*, Aug 2005. In press.

[2] C. Bartneck, J. Reichenbach, and A. van Breemen. In your face, robot! The influence of a character's embodiment on how users perceive its emotional expressions. In *Design and Emotion 2004 Conference*, Ankara, 2004. `http://www.bartneck.de/work/bartneckDE2004.pdf`.

[3] B. R. Brewer, R. Klatzky, and Y. Matsuoka. Feedback distortion to overcome learned nonuse: A system overview. *IEEE Engineering in Medicine and Biology*, pages 1613 – 1616, 2003.

[4] C. G. Burgar, P. S. Lum, P. C. Shor, and M. V. der Loos. Development of robots for rehabilitation therapy: the Palo Alto VA/Stanford experience. *Journal of Rehabilitation Research and Development*, 37(6):639–652, 2000.

[5] J. Eriksson, M. J. Matarić, and C. Winstein. Hands-off assistive robotics for post-stroke arm rehabilitation. In *International Conference on Rehabilitation Robotics*, Chicago, Illinois, Jun 2005.

[6] D. Feil-Seifer and M. J. Matarić. Socially assistive robotics. In *Proceedings of the International Conference on Rehabilitation Robotics (ICORR-05)*, Chicago, IL, Jun 28–Jul 1 2005.

[7] L. R. Goldberg. *A Broad-Bandwidth, Public-Domain, Personality Inventory Measuring the Lower-Level Facets of Several Five-Factor Models*, volume 7, pages 7–28. Tilburg University Press, Tilburg, The Netherlands, 1999.

[8] E. T. Hall. *The Hidden Dimension*. Doubleday, New York, 1966.

[9] L. E. Kahn, M. Averbuch, W. Z. Rymer, and D. J. Reinkensmeyer. Comparison of robot-assisted reaching to free reaching in promoting recovery from chronic stroke. In *Integration of Assistive Technology in the Information Age, Proceedings 7th International Conference on Rehabilitation Robotics*, pages 39–44, Evry, France, Apr 2001. IOS Press, Amsterdam.

[10] K. I. Kang, S. Freedman, and M. J. Matarić. A hands-off physical therapy assistance robot for cardiac patients. In *Proceedings of the IEEE International Conference on Rehabilitation Robotics (ICORR-05)*, Chicago, IL, Jun 28–Jul 1 2005.

[11] C. Kidd. Sociable robots: The role of presence and task in human-robot interaction. Master's thesis, Massachusetts Institute of Technology, Massachusetts, 2003.

[12] N. Miller, O. C. Jenkins, M. Kallmann, and M. J. Matarić. Motion capture from inertial sensing for untethered humanoid teleoperation. In *Proceedings of the IEEE-RAS International Conference on Humanoid Robotics*, Los Angeles, CA, Nov 2004.

[13] A. Parasuraman. Technology readiness index (TRI): A multiple-item scale to measure readiness to embrace new technologies. *Journal of Service Research*, 2(4):307–320, 2000.

[14] D. J. Reinkensmeyer, L. E. Kahn, M. Averbuch, A. McKenna-Cole, B. D. Schmit, and W. Z. Rymer. Understanding and treating arm movement impairment after chronic brain injury: progress with the arm guide. *Journal of Rehabilitation Research and Development*, 37(6):653–662, 2000.

[15] D. J. Reinkensmeyer, C. T. Pang, J. A. Nessler, and C. C. Painter. Web-based tele-rehabilitation of the upper extremity after stroke. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 10(2):102–108, 2002.

[16] S. Thrun. Toward a framework for human-robot interaction. *Human-Computer Interaction*, 19:9–24, 2004.

[17] C. Winstein, D. Rose, S. Tan, L. R, H. Chui, and S. Azen. A randomized controlled comparison of upper-extremity rehabilitation strategies in acute stroke: A pilot study of immediate and long-term outcomes. *Archives of Physical Medicine and Rehabilitation*, 85(4):620–628, Apr 2004.

[18] S. Wolf, S. Blanton, H. Baer, J. Breshears, and A. Butler. Repetitive task practice: a critical review of constraint-induced movement therapy in stroke. *Neurologist*, 8(6):325–338, Nov 2002.

[19] D. Zweig and J. Webster. Personality as a moderator of monitoring acceptance. *Computers in Human Behavior*, 19:479–493, July 2003.