

Motivational Effects of Acknowledging Feedback from a Socially Assistive Robot

Sebastian Schneider and Franz Kummert

CITEC, Bielefeld University,
Inspiration 1, 33602 Bielefeld, Germany
{sebschne, franz}@techfak.uni-bielefeld.de
<http://aiweb.techfak.uni-bielefeld.de>

Abstract. Preventing diseases of affluence is one of the major challenges for our future society. Recently, robots have been introduced as support for people on dieting or rehabilitation tasks. In our current work, we are investigating how the companionship and acknowledgement of a socially assistive robot (SAR) can influence the user to persist longer on a plank-ing task. We conducted a 2 (acknowledgement vs. no-acknowledgment) x 2 (instructing vs. exercising together) x 1 (baseline) study with 96 subjects. We observed a motivational gain if the robot is exercising together with the user or if the robot is giving acknowledging feedback. However, we could not find an increase in motivation if the robot is showing both behaviors. We attribute the later finding to ceiling effects and discuss why we could not find an additional performance gain. Moreover, we highlight implications for SAR researchers developing robots to motivate people to extend exercising duration.

Keywords: Human-Robot Interaction, Socially Assistive Robots

1 Introduction

Recently, assistive robots have been designed to support people on different rehabilitation and cognitive tasks by the robot's presence and assistance [1, 2]. Moreover, virtual agents and socially assistive robots (SAR) have been introduced as a useful tool to promote a healthy lifestyle, support for dieting or exercising besides conservative methods [1, 3, 4]. Therefore, different approaches from sociology or psychology have already been exploited (i.e. goal-setting, empathy, backstory or personalization [2, 5–7]). One other important interactional aspect that sport instructors are regularly incorporating into exercising is acknowledgement. Acknowledgement is a special type of feedback, which is used to appreciate the current state of exercising which in turn motivates the trainee to keep up on a task. It is positive reinforcing and does not compare the current performance to other persons or to previous sessions. While other types of feedback (i.e. positive, negative, comparative or corrective feedback) have been studied in Human-Robot Interaction (HRI) [8–10], the quantitative effects of acknowledgement while working out with a robot have not been compared to a

baseline measurement. Because most of the published studies comparing different types of feedback are lacking comparative results to baseline studies with no robot or no feedback, it is difficult to distinguish the true motivational effects of SAR.

In previous works, we have investigated how the mere presence of the robot affects the motivation to exercise compared to a robot that exercises together with the persons. Our findings conclude that the users experience a motivational gain when the robot is working out co-actively. However, if the system is only instructing, then the presence is not perceived as useful. Hence, we investigate in this paper how acknowledgment of the robot can increase the user’s motivation to exercise. Furthermore, we are interested in whether the acknowledgment is sufficient or the system needs to actively work out with the user. The result of this research can be useful to decide whether a SAR should work out along with the person or whether having acknowledgement is enough to motivate the user to exercise longer. For our investigation we examined four conditions: a robot exercise instructor (RI), a robot exercise instructor which is giving acknowledgement (RIF), a robot exercise companion (RC) and a robot exercise companion which is giving acknowledgement (RCF). We compare all conditions against a baseline condition without a SAR (IC). Our hypotheses are that:

- H1: people exercise longer in the RCF condition compared to all other conditions.
- H2: people exercise longer in the RIF condition than in the IC and RI condition.

This manuscript is organized as follows. The next section gives the reader an overview of existing literature in the field of motivational feedback from SARs for exercising. Afterwards, we will describe the study design to investigate our hypotheses and the system design we used. These sections are followed by the results and a discussion.

2 Related Work

In the past different robots assisting users on tasks have been reported. In this section we want to briefly review some approaches and results. Different to other investigations, we do not want to question which kind of feedback is most useful for the user. We want to answer the general question how acknowledgement can influence the user’s motivation in general.

The preference for a relational vs. a non-relational robot-coach have been investigated in [1]. The study shows that users have a preference for relational feedback. Relational is used as the robot’s capabilities to exploit all of its social interaction and personalization approaches. Thus, the robot always gives the user praise upon correct completion of an exercise, it provides reassurance in case of failures, refers to the user by its name, references past experiences and uses humor. In the non-relational condition the robot coach gives instructional feedback but does not employ any relationship building. The authors of this study used an in-between subject design to evaluate the different relational styles of the robot. They did not find any differences in the exercise performance based the

relation or non-relational robot. In our previous work we have investigated the differences of performance-based vs. non-performance-based feedback for users doing a cognitive task [11]. Personalized feedback from robotic tutors has been investigated in [2]. Both works support that individualized feedback based on the user’s performance can increase the tutor’s effectiveness. A different study compares self-comparative vs. other-comparative to non-comparative feedback in a push-button task[9]. However, the authors could not find evidence for any of their hypotheses. Positive, negative and neutral feedback of a robot or human instructor has been studied in [12]. Study subjects have a preference for positive feedback from a robot. However, they could not find any feedback preference from a human interaction partner.

Compared to the presented related work, we are focusing on the effects of acknowledgement of a social robot while exercising. Furthermore, we compare our results to a baseline condition where no robot is present at all. Thus, the results of our study let us conclude about the importance of acknowledgement for the user’s motivation to exercise longer as well as how it is moderated by the robot’s style of companionship.

3 Study Design

Our study is inspired by a research comparing the effects of working out with a virtual avatar with different degrees of human-likeness [13]. To allow possible comparisons, we wanted to replicate the study as close as possible. However, we needed to include some changes in the study design due to the actual robot agents. Therefore, we changed the exercises from forearm planks to full planks due to the robots limited degree of freedom.

3.1 Experimental Design and Participants

Participants (n=95) were in one of 5 conditions (independent condition (IC), robot instructor condition (RI), robot companion condition (RC), robot instructor feedback condition (RIF) and robot companion feedback condition (RCF)) with around 18 participants in each condition. Participants were mostly students (f/m = 44/51, age M=25.4 years, SD=5.6) from our university acquired by flyer distributed on the campus. They received seven Euros as monetary compensation. Three participants from the IC had to be excluded. One was an outlier persisting much less during the first block compared to all other participants. Two other persons had to be excluded because they were doing the exercises wrong. One participant from the RI and two participants from the RCF condition had to be excluded from the survey evaluation because the data were missing. However, we still could analyze their performance data.

3.2 Exercises and Conditions

Participants had to do two blocks of five isometric exercises each (see Fig. 1). Participants in the IC condition did the exercises in both blocks alone. In all other

conditions the participants did the exercises alone in the first block and with the humanoid robot Nao¹ in the second block. During the RI condition the robot was announcing the exercises the user had to do, as well as how long the break is. While the user was exercising the robot was standing in front of the person observing her/him while exercising and showing some idle behavior. After the user has finished an exercise s/he received a general encouraging feedback. In the RC condition the robot was showing the same behavior as in the RI condition. But instead of just standing in front of the participant the robot was exercising together with the user and also held the isometric exercises. However, the robot always hold it a bit longer than the participant. In the feedback conditions (RIF, RCF) the robot was giving some acknowledgment during each exercise. This feedback was generated based on the performance from the first block. After three quarter of the time they hold the exercises during Block 1 the robot gave some acknowledgement ('you are doing great', 'keep on', 'this is the last exercise, keep on the good work', etc.). We used three quarter of the previous time as a threshold because a preliminary analysis showed that most of the people will stop doing an exercise after three quarter of the time from the first block in the second block. The system was implemented using a state machine coordination for distributed systems, a Kinect for user activity recognitions as well as features from the NaoQi API. We neglect a detailed preliminary analysis and an extensive description of our system due to the limitation of this paper.

3.3 Procedures

Participants arrived at the lab, read and signed a consent form which informs them that they will be recorded during the whole time of the experiment. They watched a short video of Nao demonstrating the five exercises. Afterwards, they were brought to a fitting room to change clothes and strap on a heart rate belt. They were instructed to do each exercise as long as they can and if they could not persist any longer to stand up immediately, rate their perceived exertion on a prepared scale and wait for thirty seconds before continuing with the next exercise. They were guided to the lab and told to start after they have waited for a short time, so that the experimenter can check that the recording is working properly. Then the participants did each exercise alone in the lab while the experimenter observed them from a different room and took the participant's times for each exercise. The participants completed Block 1 (each exercise once). Afterwards, the participants had a ten minute break where they were offered a glass of water and had to answer a self-efficacy belief scale about the exercises. After the break participants in the IC condition were told

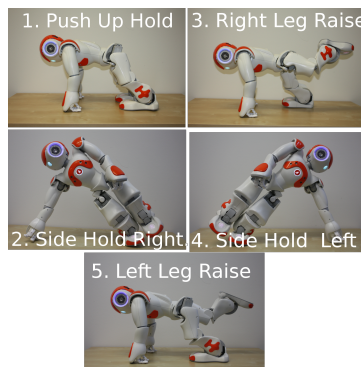


Fig. 1. The five isometric planking exercises.

¹ <https://www.aldebaran.com/en>

the average time they held the planks and that they would complete the same set of exercises again (Block 2). In every condition the participants were not told that they had to do a second block of exercises until they had finished the first block.

In the robot conditions participants were told that they will do the same set of exercises again but that this time a robot will be present. They were instructed to follow the guidance of the robot through the session and that they are exercising together from now on (RC). The experimenter told the participant the true average time they held the planks, like in the IC condition, but gave them a false information on how long the robot can persist the exercises. They were always told a number which is forty percent higher than their average time. This creates an unfavorable comparison which leads to greater effects and was adopted from the previous study [13].

Again the experimenter did not enter the room together with the participant. In all robot conditions, the participant and robot had a short interaction phase. During this phase the robot told them its name (Nao), hometown (Paris) and hobbies (gardening, reading) and waited for a short time to give the human participant a chance to also share his/her personal information. This was done because prior research showed that people treat agents more like humans when there was an initial verbal interaction between them [7]. After completing Block 2 the robot thanked the participants for their participation, told them that they are allowed to leave the room and that it needs to rest a bit. After leaving the room the participants completed a questionnaire, were debriefed and received a monetary compensation. The whole procedure took about 45 minutes to one hour.

3.4 Measures

Persistence Persistence was the number of seconds a plank was held from the moment participants moved into position they quit. Block scores were calculated using total average seconds held on all five exercises.

Perceived Exertion Perceived exertion was measured using the Borg rating scale ([14]). The scale goes from 6 to 20 (6: 'no exertion at all', 20: 'maximal exertion'). The participants were asked to rate their exertion immediately after each exercise.

Negative Attitudes Towards Robots Scale For the RIF and RIC, we also assessed the negative attitudes towards robots scale [15].

Perception of the Partner Participants in the partnered condition completed questionnaires asking to rate the perception of the robot. To investigate the human-likeness, we asked the participant to rate the robot based on the God-speed questionnaire (5 point-based differential scale, [16]).

In the feedback conditions we also asked the participants to rate the information quality, the cooperation, the openness to influence as well as the team perception on a five-point Likert-scale [17].

Physical Training Enjoyment We assessed the physical training enjoyment the users had using the Physical Activity Enjoyment Scale ([18]). We used the average value of all items as overall enjoyment score. Furthermore, we asked how much time they spent exercising per week and their intention to train tomorrow for at least 30 minutes.

4 Results

We conducted an analysis of variance (ANOVA) to find differences in the enjoyment, performance on Block 1, intention to exercise, perceived exertion and general amount of time spent doing sports. We found no difference in persistence on Block 1 ($P = 0.47$), exercising per week ($P = 0.39$), enjoyment ($P = 0.467$), perceived exertion on Block 1 ($P = 0.106$) and Block 2 ($P = 0.397$). However, we found a difference in the intention to exercise for at least 30 minutes on the following day ($F_{4,84} = 3.453, p < 0.05$). Pairwise comparisons using t tests with pooled SD and Bonferroni correction revealed a differences between the RCF ($M=4.77, SD=.42$) and RI ($M=3.6, SD=1.37$) condition ($p < 0.05$).

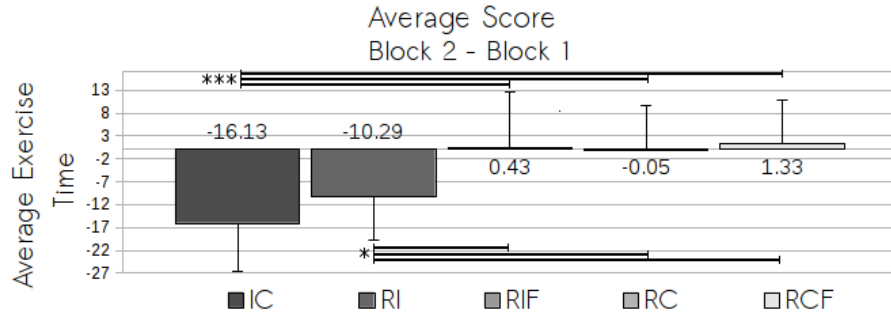


Fig. 2. Differences between Block 2 and Block 1 (*: $p < 0.05$; ***: $p < 0.0001$).

Persistence As a primary dependent variable we used the average difference persistence time between the two blocks (Block 2 - Block 1). This approach controls for individual differences in strength and fitness and shows possible changes in persistence. The results obtained for the average block score of Block 2 subtracted with the average block score of Block 1 are shown in Figure 2. An 5 (conditions) \times 1 (persistence) analysis of variance (ANOVA) on the difference scores showed a significant main effect for the conditions ($F_{4,80} = 9.927, p < 0.001$). A pairwise comparison using t tests with pooled SD and Bonferroni adjustment revealed significant differences between IC and RC ($p < 0.0001$), IC and RCF ($p < 0.0001$), IC and RIF ($p < 0.0001$), RC and RI ($p < 0.05$), RCF and RI ($p < 0.05$) and RI and RIF ($p < 0.05$) (see Fig. 2). These results hint that acknowledgement has a positive effect on the user's exercising performance. Since, we instructed participants that they will interact with a robot and we have not told them that a human is tele-operating the robot, it shows that people are

sensitive to acknowledgement from a clearly technical system. This shows that interactive motivational capabilities can be exploited from technical systems to enhance human exercise duration. This is particularly important in light of the results that participants exercised longer without enjoying the training less or feeling more exerted. Nevertheless, we did not find the same motivational gain in the companion conditions (RC vs. RCF). Subjects in this RCF condition did not persisted significantly longer compared to the RC condition or the RIF condition. Based on the feedback from the interviews we suppose that this is due to ceiling effects based on the picked exercises. Participants mentioned in post-interviews that they would have liked to hold the exercise longer, but they were not able to persist any longer due to too much load on the wrists.

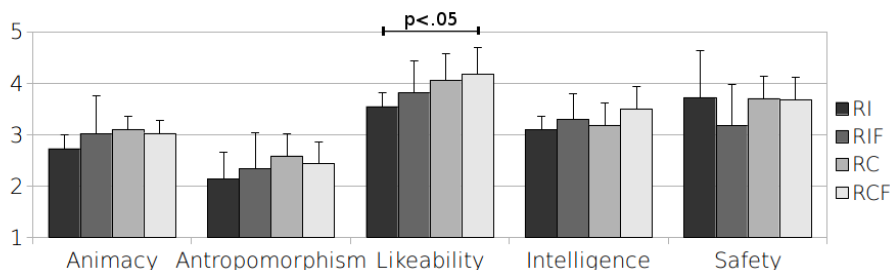


Fig. 3. Results of the Godspeed questionnaire.

Perception of the Partner We conducted several 4 (condition) x 1 (sub-scale) ANOVAs to find differences between the ratings on the Godspeed questionnaire for the different conditions (see Fig. 3). We did not find differences for the ratings of animacy ($P=.21$), anthropomorphism ($P=.25$), intelligence ($P=.13$) and safety ($P=.13$). However, we found significant main effects for likeability ($F_{3,67} = 3.956, p < .05$). A pairwise comparison using t tests with pooled SD and bonferroni adjustment revealed significant differences on the likeability scales between RCF and RI ($p < .05$) and a tendency between RC and RI ($p = .08$).

While we could not find evidence for higher performance in the companion conditions, these results show that people liked the robot in the companion conditions more than in the instructor conditions. Higher ratings on the likeability scales could be beneficial for long-term investigations, because they could positively influence the user’s long-term motivation to engage in an interaction.

Further Results We analyzed the differences between the ratings for team perception, perceived information, cooperation, openness to influence and NARS (see Fig. 4). A Welch Two Sample t-test between the RIF and RCF conditions revealed no differences for NARS ($t(28.781) = 1.425, p = 0.16$), openness ($t(33.942) = -0.82468, p = 0.4153$) and cooperation ($t(33.783) = -1.2039, p = 0.237$). However, we found a difference for the team perception ($t(33.576) = -2.2821, p < 0.05$) and a tendency for the perceived information quality ($t(33.983) = -2.0284, p = 0.05$)

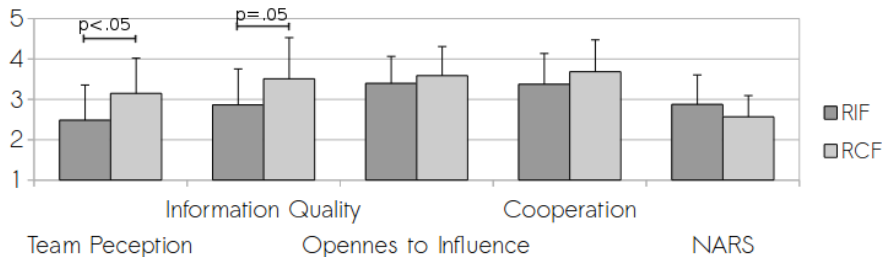


Fig. 4. Results for team perception, information quality, cooperation, openness and NARS ratings.

It is hardly surprising that participants had higher feelings for working out together, but it is remarkable that they rated the information quality as higher because the speech output was exactly the same in both conditions. Thus, co-actively working out could increase the user's trust in the system quality.

5 Discussion

We have presented a study on the effects of acknowledgement from a SAR on the user's exercise performance. We compared acknowledgement-feedback vs. non-acknowledgement-feedback for a robot exercise instructor and companion. To the best of our knowledge, our results are the first to compare the presence, the role and the acknowledgement of a robot against a baseline condition during an exercising task.

Users exercising co-actively with the robot had a significant performance gain compared to being instructed by the robot or exercising alone. However, if the participants receive additional acknowledgement of a robot we found a performance gain compared to conditions where the robot is only instructing the user. Thus, our hypothesis H2 is supported due to the evidence that participants had a significant performance gain in the instructor conditions if the robot is giving feedback (RI vs. RIF).

However, we could not support our hypothesis H1. We did not detect the same performance gain between co-actively exercising with a robot (RC) or working out with a robot and receiving additional acknowledgement (RCF). We suppose that this is due to a ceiling effect caused by the selection of the exercises. Nevertheless, we can not fully conclude on this issue. Therefore, more investigations are needed that explore different kind of exercises and measure how difficult the participants experience the exercises in relation to their task performance.

While we could not detect any performance gains in the companion conditions, we found that in both companion conditions the users liked the system more. Furthermore, if the robot is exercising along with the subjects, the information quality of the system is rated higher. We assume that this is an important

aspect for long-term HRI and could lead to longer training engagement. Therefore, we plan to run long-term investigations to gain more clarity on this aspect.

Finally we would like to discuss the limitations of this paper. First, we used only isolated abdominal plank exercises. Therefore, our future studies will also include a set of dynamic exercises (e.g. squats, push ups). Second, we have not assessed the utility of the robot against other technological devices (e.g. smart phone based applications that provide feedback). We suppose that anthropomorphic technology will lead to higher exercise adherence and performance than non-anthropomorphic devices. The degree of human-likeness of a virtual agent has already been shown to influence exercising duration [13]. However, we need to conduct more research that evaluates the utility of a robot companion by comparing it to other devices and also to human exercising partners.

In summary, our results show that a SAR should at least do one of two things when interacting with a user: It should either exercise along or give encouraging acknowledgement if it is only instructing an exercise. This research is important for all researchers in the field of SAR who want to build robots to motivate people to workout. The take-home message is: Robots working out co-actively with humans are more motivating and lead to higher performances, however if the robot is not able to exercise along with the user, due to some physical limitation, it is also sufficient to give the user acknowledgement in order to have the same motivational effect. This result allows us to follow our line of research where we intend to build a SAR that is instructing users on a full body weight workout. Our future work also includes to make all the questionnaires, data and programs available to allow other researchers to repeat our study as well as an in-depth analysis of further qualitative and quantitative data we have gathered during our experiments. This includes video material, heart rate data and self-efficacy beliefs. The analysis of these multi-modal data might reveal some new insights for our attempts to build robots that motivate people to work out more.

Acknowledgments This research was funded by grants from the Cluster of Excellence Cognitive Interaction Technology 'CITEC' (EXC 277), Bielefeld University.

References

1. Juan Fasola and Maja J Matarić. Using socially assistive human-robot interaction to motivate physical exercise for older adults. *Proceedings of the IEEE*, 100(8):2512–2526, 2012.
2. Daniel Leyzberg, Samuel Spaulding, and Brian Scassellati. Personalizing robot tutors to individuals' learning differences. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, pages 423–430. ACM, 2014.
3. Timothy Bickmore and Justine Cassell. Relational agents: a model and implementation of building user trust. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 396–403. ACM, 2001.
4. Cory D Kidd and Cynthia Breazeal. Robots at home: Understanding long-term human-robot interaction. In *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pages 3230–3235. IEEE, 2008.

5. Iolanda Leite, Ginevra Castellano, André Pereira, Carlos Martinho, and Ana Paiva. Empathic robots for long-term interaction. *International Journal of Social Robotics*, 6(3):329–341, 2014.
6. Rachel Gockley, Allison Bruce, Jodi Forlizzi, Marek Michalowski, Anne Mundell, Stephanie Rosenthal, Brennan Sellner, Reid Simmons, Kevin Snipes, Alan C Schultz, et al. Designing robots for long-term social interaction. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1338–1343. IEEE, 2005.
7. Timothy W Bickmore and Rosalind W Picard. Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 12(2):293–327, 2005.
8. Jaap Ham and Cees JH Midden. A persuasive robot to stimulate energy conservation: the influence of positive and negative social feedback and task similarity on energy-consumption behavior. *International Journal of Social Robotics*, 6(2):163–171, 2014.
9. Katelyn Swift-Spong, Elaine Short, Eric Wade, and Maja J Mataric. Effects of comparative feedback from a socially assistive robot on self-efficacy in post-stroke rehabilitation. In *Rehabilitation Robotics (ICORR), 2015 IEEE International Conference on*, pages 764–769. IEEE, 2015.
10. Luise Süssenbach, Nina Riether, Sebastian Schneider, Ingmar Berger, Franz Kummert, Ingo Lütkebohle, and Karola Pitsch. A robot as fitness companion: towards an interactive action-based motivation model. 2014.
11. Sebastian Schneider, Nina Riether, Ingmar Berger, and Franz Kummert. How socially assistive robots supporting on cognitive tasks perform. In *Proceedings of the 50th Anniversary Convention of the AISB*, 2014.
12. Eunil Park, Ki Joon Kim, and Angel P Del Pobil. The effects of a robot instructors positive vs. negative feedbacks on attraction and acceptance towards the robot in classroom. In *Social robotics*, pages 135–141. Springer, 2011.
13. Deborah L Feltz, Samuel T Forlenza, Brian Winn, and Norbert L Kerr. Cyber buddy is better than no buddy: A test of the köhler motivation effect in exergames. *GAMES FOR HEALTH: Research, Development, and Clinical Applications*, 3(2):98–105, 2014.
14. Gunnar Borg. *Borg’s perceived exertion and pain scales*. Human kinetics, 1998.
15. Tatsuya Nomura, Takayuki Kanda, and Tomohiro Suzuki. Experimental investigation into influence of negative attitudes toward robots on human–robot interaction. *Ai & Society*, 20(2):138–150, 2006.
16. Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics*, 1(1):71–81, 2009.
17. Clifford Nass, BJ Fogg, and Youngme Moon. Can computers be teammates? *International Journal of Human-Computer Studies*, 45(6):669–678, 1996.
18. Deborah Kendzierski and Kenneth J DeCarlo. Physical activity enjoyment scale: Two validation studies. *Journal of Sport & Exercise Psychology*, 1991.