Increasing Personalization in Long-Term Interactions with a Workplace Companion Robot

Rhian C. Preston and Naomi T. Fitter
{prestonr,naomi.fitter}@oregonstate.edu
CoRIS Institute, Oregon State University
Corvallis, OR, USA

ABSTRACT

Modern computer users face health challenges from sitting continually and practicing bad posture. Solutions (e.g., phone applications and life coaches) exist to address these potentially hazardous behaviors, but current solutions either fall into disuse after a short period of time or are unattainable for many potential users. We propose socially assistive workplace robots as a solution with the potential to stay in use for longer than comparable apps while offering easier access than personal coaches. Building on a related past study by our research group, this paper presents an updated workplace companion robot design with more personalization features that we plan to evaluate in an upcoming long-term user study. Data gained from the future deployment will offer additional insights for how to personalize future robot prototypes via online learning. This work can help to inform other researchers with interest in socially assistive robotics, long-term robot deployments, and personalized robot learning.

KEYWORDS

socially assistive robots, break-taking, long-term robot deployment

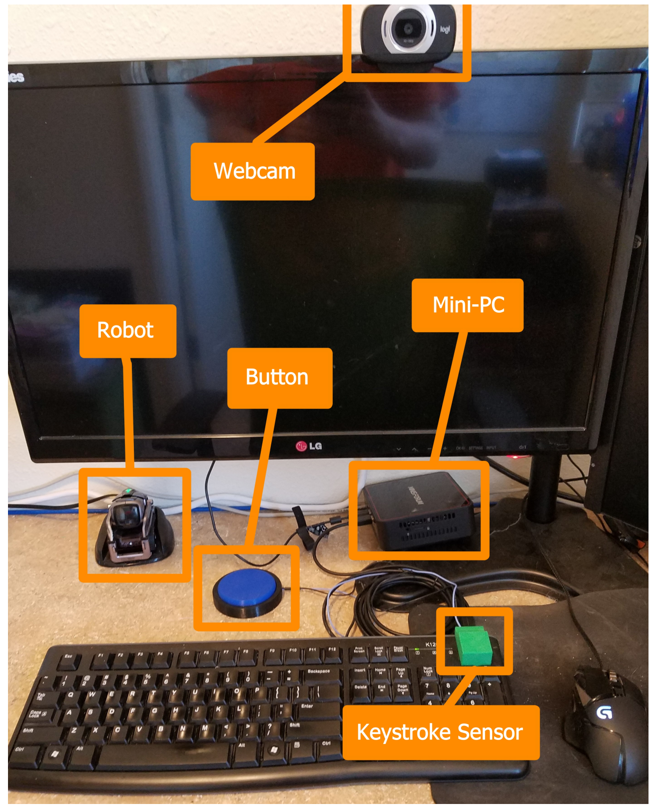

Figure 1: System setup for the tabletop robot. Not shown: seat sensor and wireless chair sensor module.

1 INTRODUCTION

Health challenges specific to sedentary office work (e.g., poor cardiovascular and musculoskeletal health) are becoming increasingly common for computer users of all walks of life [10]. Automated break-taking aids (like computer and phone apps) can yield short-term success in improving health outcomes, but these solutions have yet to demonstrate sustained behavior change results. More resource-intensive and tailored behavior change aids (such as life coaches) can produce long-term changes in break-taking habits [4]; however, these tailored treatments cannot scale up to the needs of a broader population due to cost and logistical complications. Physically embodied robotic systems have the scalability and automation of break-taking support apps while offering a presence and social engagement akin to life coaches. Therefore, we propose a tabletop socially assistive robotic (SAR) system, pictured in Fig. 1, as a suitable middle ground for supporting healthy habits at work.

Past work on the project discussed in this paper has shown that over a few days, participants enjoyed using a break-taking SAR system more than a phone application-like alternative, but there were no significant differences in user productivity or break motivation [15]. Participants also pointed out desired changes in the robotic system to better suit their needs. Work by [8] further showed a relationship between expressive break-taking systems and user responsiveness. Efforts in a related robotics research area (SAR systems in autism therapy) shows that long-term, in-situ engagement with socially assistive robots can yield positive behavioral changes, such as improving social skills in children with developmental delays [13]. In the current paper, we propose a method for studying an improved version of the break-taking treatment from [15] for a time period on the order of the robotic treatment length in [13]. The planned study will help us understand if robotic break-taking aids, in a way similar to robotic social coaches, can promote lasting and positive behavior change.

Our central goal in this work is to understand differences between users' long-term break-taking practices when supported by an embodied SAR system vs. a non-embodied phone app-like alternative. We propose a single-case-style design that will allow us to compare participant behaviors of interest (i.e., responses to baseline and different treatment phases) across study phases. The results will inform future break-taking intervention work and supply the initial dataset needed for more sophisticated future SAR system learning strategies.

2 SAR SYSTEM DESIGN

In the previous SAR deployment from [15], an initial robotic system prototype built around the Anki Cozmo robot helped users to take breaks at work over a one-day period. One challenge, as mentioned in the past paper, was occasional *losses in connection* between the system's base station hardware and the robot. This problem was manageable in a short-term deployment, but would cause major difficulties over long-term study designs. Among other products of that work (such as promising short-term social findings), we gleaned participant critiques of the SAR system. Most commonly, users mentioned:

- Confusion about when the robot cue started and stopped
- Feelings that the robot interrupted their work
- Desire for a more personalized robot
- Preferences for particular types of robot cues

Although promising, the past study clearly showed that the SAR system required revision. One key needed update was *improvements to the robot connectivity* to ensure robustness of operation. The previous base station computer in [15] was a Raspberry Pi 3 B+ running Ubuntu, which interfaced with an Android phone to control Cozmo. To achieve more flexibility and persistence in strategies for connecting with the Cozmo, we updated the central computer to a MinisForum U500-H Mini-PC with a Windows 10 operating system running both Android and Ubuntu emulation environments. The Android emulator, which runs on Windows, communicates between central logic in the Linux environment and the Cozmo robot. The new, more substantial central processor is able to persistently reconnect to the Cozmo robot and prevent interruptions to the robot function during future deployments. A recent system test demonstrated the robot's ability to stay connected for multiple weeks without interruption.

The other observed design flaws, while not as critical to the literal functioning of Cozmo, would interfere with the ongoing adoptability of the robot over extended deployments. Thus, we also prioritized addressing these shortcomings in the updated system. To reduce *confusion about robot cue bounds*, we updated the robot routine logic to stop prompting the user once the person stands up to take a break, rather than enforcing that the robot complete a long sequence of animations.

To address the feeling that the robot *interrupted the work of users*, we added to the sensing system that the Cozmo robot uses to gauge the person's work state. As in the previous system, we used a seat occupancy sensor connected to the rest of the system via a Bluetooth-enabled chair sensor module to track how long a user remained seated. However, unlike in the previous system, the robot does not deliver a prompt by default once a fixed amount of time seated has passed. The updated system uses two additional sensors–a LIS3DH accelerometer (connected to the base station via a Teensy 3.2 microcontroller) and a system "snooze" button–to enable more alignment between robot behaviors and user needs. The accelerometer is affixed to the participant's keyboard, and uses vibration to sense keystrokes, allowing for the selection of an opportune break prompt time (within a window around the desired break interval) when the user appears to be more interruptible; past work on work state detection supports the selection of typing as a proxy for work focus [9]. The snooze button is a large button the user can press to momentarily halt and delay the break-taking prompt. In this way, the system can still account for times the user may be mid-task even in the absence of active keyboard use. A final sensor in the system (residual from [15]) is a webcam for recording video of participant interactions with the system.

Past participants had also expressed a *desire for a more personalized robot*. To some extent, the improved user state sensing discussed above helps to support this need. Through early formal and informal deployments, we further noticed that although the state of the art in related literature is encouraging breaks every 30 minutes, individual users had different break-taking needs and goals. The updated system is designed to more easily accept personalized user requirements upon system initialization; for example, desired timing between break prompts is entered for each individual.

Lastly, users *desired differing types of robot behavior or motion*. We know from past work such as [12] that certain requirements (e.g., minimal sound) exist for robots in the workplace; however, it is not as clear how robot behaviors should personalize to the individual or even change over time during deployment with the same user. Our proposed long-term study will contribute to uncovering this information; during the robot-based treatment, Cozmo will perform a random walk of action styles when prompting participants to take breaks. These action styles are generally grouped into behaviors spanning the valence vs. energy level space proposed by Russel's circumplex model of affect [11]. These prompts include behaviors such as providing happy or unhappy facial cues, lifting and lowering the robot's forklift, spinning in circles on a user's desk, and charging forward a short distance. Information about participant work state before the prompt and subsequent response to the prompt will provide the information needed to train initial models that can support future personalized break-taking encouragement with this type of system.

3 METHODS

Our planned two-month exploratory study is between-subjects and single-case-style, i.e., following the example of [13]. We will include a treatment-free initial baseline phase (two weeks) and final retention phase (two weeks) in our study design. Users will additionally experience one of the following one-month treatment phases in between:

- **SAR system**: An Anki Cozmo tabletop robot-based (i.e., primarily hardware-mediated) interaction method for delivering break-taking prompts.
- **Non-Embodied system**: A phone application-based (i.e., primarily software-mediated) interaction method for delivering break-taking prompts, such as PomoDoneApp. Other than having a phone as the break prompt delivery mechanism (rather than the robot), this system uses the same logic and sensor hardware setup.

The conditions will be balanced and randomly assigned to participants. The presented methods are approved by the Oregon State University IRB under protocol #IRB-2019-0067.

3.1 Participants

The study will include ten participants who spend most of their workday sitting at a desk and working with a computer. Specifically, we will recruit individuals who use a computer while seated at least three hours a day.

3.2 Measures

We will use *surveys* to gauge participant opinions of and experience with robots, as well as feelings of satisfaction, overall mood, and productivity throughout each phase of the study. These surveys, as outlined below, will use 7-point Likert scales unless otherwise noted.

- A **pre-study survey** to capture participants' baseline perceptions of system performance expectancy, use effort expectancy, attitude toward using technology, self efficacy, and attachment using questions based on the Unified Theory of Acceptance and Use of Technology (UTAUT) [14]. We will also collect details on participants' robotics experience and health goals.
- A **weekly survey** to capture participants' experiences during each week, using questions adapted from the NASA Task Load Index (TLX) [6], the Self-Assessment Manikin (SAM) [2], the Working Alliance Inventory (WAI) [7], and questions about perceived work performance.
- A **post-study survey** including all UTAUT-based questions from the pre-study survey, Ten-Item Personality Inventory (TIPI) [5] questions, and demographic questions.

During the study, we will gather *behavioral* data from system logs and audiovisual recordings. Specifically, this includes:

- **Break-taking information** showing if and when the participant stands after being prompted.
- **Productivity levels** from keystroke activity.
- **System snooze inputs** from when the participant uses the snooze button to delay breaks.
- **User affect** from video recordings before, during, and after delivered prompts.

For reference during analysis, the logs also will record the timing of break prompts and the specific Cozmo behavior used for robot prompts.

The video recordings from the study will supply information for qualitative analysis. To gain further context for study observations, we will additionally ask participants to record a daily oral log of any notable events or challenges from each day and conduct a closing semi-structured interview with each participant.

3.3 Procedure

After consenting to participate in the study, the participant will complete the pre-study survey; then their workspace will be outfitted with the system hardware, which will initially be configured for the two-week *baseline phase*. During the baseline, the participant will complete weekly surveys and a brief daily oral log.

Next, the system will be reconfigured for the randomly assigned *treatment phase*. The participant's workspace will be equipped with either the Cozmo robot or non-embodied system (depending on the assigned condition). The treatment phase will last four weeks, during which the system will provide break prompts. The same weekly surveys and daily oral logs will occur as in the baseline phase.

The two-week *retention phase* has the same setup as the baseline phase. At the close of this phase, we will collect the system and administer the post-study survey and interview.

3.4 Hypotheses

Our hypotheses were guided by previous work from [15] showing evidence of more satisfaction using socially assistive robots over non-social systems in short term study and the work from [8] relating expressive systems and user responsiveness:

H1: Break adherence and productivity will be higher during either treatment compared to the baseline.
H2: Breaks will be more pleasant, relaxing, and re-energizing with the SAR system than with the non-embodied system.
H3: Responsiveness to break taking prompts will be faster and persist for longer with the SAR system compared to the non-embodied system.

3.5 Analysis

We will first analyze self-reported and behavioral data using visual inspection of observations over time during each study phase, as recommended by best practices in single-case studies [3]. The presence and repeatability of trends among users within and between study groups will help us understand the prospective reproducibility of findings. We will also consider whether statistically significant differences appear across phases by using repeated measures analysis of variance (rANOVA) and analysis of covariance (ANCOVA) tests. User affect will be extracted from the video recordings using OpenFace 2.2.0 [1]. We will use thematic analysis to better understand qualitative data from the study (e.g., to analyze user responses to/behaviors toward the robot from video).

4 PRELIMINARY RESULTS AND FUTURE WORK

Running the full proposed study is a future work step, but ongoing testing with a first test user reveals early anecdotal results of our system design updates. The system connection is more robust and reliable after the switch from phone to Android emulator control, which, paired with automatic re-connection scripts, allows the system to operate for weeks without interruption. The keystroke sensor shows promise as a beneficial automated fine-tuning adjustment on break-taking prompt timing to mitigate the frequency of mid-thought interruptions.

Our test user selected a one-hour break prompt spacing for his system and continues to take breaks with the robot after multiple weeks of use. A few of the robot prompt behaviors appear to be too subtle; for these, he failed to even notice that the robot left its base to deliver a cue. At the opposite end of the prompt disruptiveness spectrum, other behaviors seem to take up too much desk space or be excessively animated. This balance in initial findings supports our future plans of personalized learning for robot behavior selection; the robot should use information about its user and surroundings to effectively cue a break while avoiding unnecessary levels of intrusion.

Strengths of the proposed study include a relatively long-term deployment of a robotic system in a natural environment. Because of the large size of the population of interest, we will eventually be able to both achieve long-term deployments and gather enough data to achieve statistically significant results in follow-up work. Initial personalization features (i.e., user-selected settings) and potential future personalization strategies (i.e., online learning strategies) may support the enduring success of the robotic system.

Limitations of the proposed study will arise due to the small initial sample size and the potential homogeneity of the sample population, who will be recruited from a single university and region. Views about break-taking vary widely among work teams, larger organizations, and societies at large, but this early work will focus only on the cultural context of the United States.

Our main *future work* step will be running the proposed study, analyzing the results, and examining data about how users responded to different prompts from the personalized model-training perspective. The results of the study will guide future work for personalizing assistive robots, and further quantify how people emotionally delineate and respond to assistive robots, non-embodied systems, and other people.

ACKNOWLEDGMENTS

We thank Brian Zhang and Ameer Helmi for their contributions to the initial robotic system development and design.

REFERENCES

[1] Tadas Baltrušaitis, Marwa Mahmoud, and Peter Robinson. 2015. Cross-dataset learning and person-specific normalisation for automatic Action Unit detection. In *Proc. of the IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, Vol. 06. 1–6.

[2] Margaret M Bradley and Peter J Lang. 1994. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry* 25, 1 (1994), 49–59.

[3] Breanne J Byiers, Joe Reichle, and Frank J Symons. 2012. Single-subject experimental design for evidence-based practice. *American Journal of Speech-Language Pathology* 21, 4 (2012), 397–414.

[4] Charlotte L Edwardson, Tom Yates, Stuart J H Biddle, Melanie J Davies, David W Dunstan, Dale W Esliger, Laura J Gray, Benjamin Jackson, Sophie E O'Connell, Ghazala Waheed, and Fehmidah Munir. 2018. Effectiveness of the Stand More AT (SMArT) Work intervention: Cluster randomised controlled trial. *BMJ* 363 (2018).

[5] Samuel D Gosling, Peter J Rentfrow, and William B Swann. 2003. A very brief measure of the Big-Five personality domains. *Journal of Research in Personality* 37, 6 (2003), 504–528.

[6] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in Psychology*. Vol. 52. Elsevier Science & Technology, 139–183.

[7] A. O. Horvath and L. S. Greenberg. 1989. Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology* 36, 2 (1989), 223–233.

[8] Nassim Jafarinaimi, Jodi Forlizzi, Amy Hurst, and John Zimmerman. 2005. Breakaway: An ambient display designed to change human behavior. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems (CHI EA '05)*. ACM, 1945–1948.

[9] Harmanpreet Kaur, Alex C Williams, Daniel McDuff, Mary Czerwinski, Jaime Teevan, and Shamsi T Iqbal. 2020. Optimizing for happiness and productivity: Modeling opportune moments for transitions and breaks at work. In *Proc. of the CHI Conference on Human Factors in Computing Systems*. 1–15.

[10] Neville Owen, Geneviève N Healy, Charles E Matthews, and David W Dunstan. 2010. Too much sitting: The population health science of sedentary behavior. *Exercise and Sport Sciences Reviews* 38, 3 (2010), 105–113.

[11] James A Russell. 1980. A circumplex model of affect. *Journal of Personality and Social Psychology* 39, 6 (1980), 1161.

[12] Selma Sabanovic, Sarah Reeder, and Bobak Kechavarzi. 2014. Designing robots in the wild: In situ prototype evaluation for a break management robot. *Journal of Human-Robot Interaction* 3, 1 (2014), 70.

[13] Brian Scassellati, Laura Boccanfuso, Chien-Ming Huang, Marilena Mademtzi, Meiying Qin, Nicole Salomons, Pamela Ventola, and Frederick Shic. 2018. Improving social skills in children with ASD using a long-term, in-home social robot. *Science Robotics* 3, 21 (2018).

[14] A Weiss, R Bernhaupt, M Tscheligi, D Wollherr, K Kuhnlenz, and M Buss. 2008. A methodological variation for acceptance evaluation of human-robot interaction in public places. In *Proc. of the IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. 713–718.

[15] Brian J Zhang, Ryan Quick, Ameer Helmi, and Naomi T Fitter. 2020. Socially assistive robots at work: Making break-taking interventions more pleasant, enjoyable, and engaging. In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.